Applied Multivariate Statistical Analysis Solutions

Nathan Crouse

August 31, 2024

Contents

1 Chapter 1

2 Chapter 2

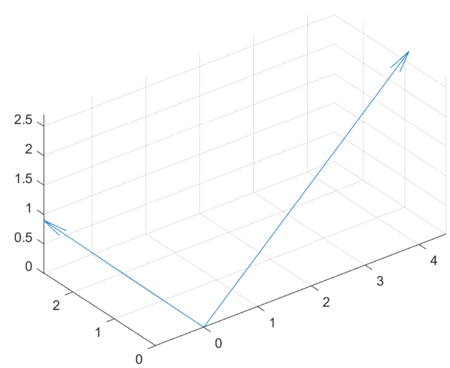
2.1

Let $\mathbf{x}' = [5, 1, 3]$ and $\mathbf{y}' = [-1, 3, 1]$.

(a) Graph the two vectors.

```
x = [5,1,3]'; y = [-1,3,1]';
starts = zeros(2,3);  % Starts at the origin.
ends = [x'; y'];  % Ends at the point.

% quiver3 args are x,y,z,u,v,w. x,y,z are the start positions and u,v,w are the end positions.
a = quiver3(starts(:,1), starts(:,2), starts(:,3), ends(:,1), ends(:,2), ends(:,3));
axis equal
saveas(a, '.\applied-multivariate-statistics\solutions\chapter-2\sol2.1a.png', 'png')
```



(b) Find (i) the length of $\mathbf{x}$, (ii) the angle between $\mathbf{x}$ and $\mathbf{y}$, and (iii) the projection of $\mathbf{y}$ onto $\mathbf{x}$.

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}'\mathbf{x}} = \sqrt{5^2 + 1^2 + 3^2} = 5.9161$$

$$\mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta \Rightarrow \theta = \cos^{-1}\left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}\right) = \cos^{-1}\left(\frac{1}{5.9161 \times 3.3166}\right) = 87.0787^\circ$$

MATLAB code `acosd((x'*y)/(norm(x)*norm(y)))` returns the angle in degrees.

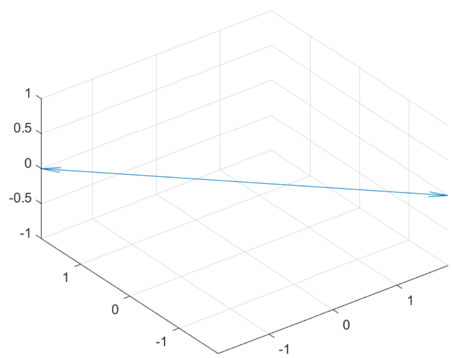
$$\text{comp}_{\mathbf{x}}\mathbf{y} = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|}$$

$$\text{proj}_{\mathbf{x}}\mathbf{y} = \text{comp}_{\mathbf{y}}\mathbf{x}\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) = \left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|}\right)\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) = \left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\right)\mathbf{x} = \left(\frac{1}{35}\right)\begin{bmatrix}5\\1\\3\end{bmatrix} = \begin{bmatrix}5/35\\1/35\\3/35\end{bmatrix}$$

(c) Since $\bar{x} = 3$ and $\bar{y} = 1$, graph $[5-3, 1-3, 3-3] = [2, -2, 0]$ and $[-1-1, 3-1, 1-1] = [-2, 2, 0]$.

```
starts = zeros(2,3);  % Starts at the origin.
ends = [(x-mean(x))'; (y-mean(y))'];  % Ends at the point. Subtract the mean values.

% quiver3 args are x,y,z,u,v,w. x,y,z are the start positions and u,v,w are
% the end positions.
b = quiver3(starts(:,1), starts(:,2), starts(:,3), ends(:,1), ends(:,2), ends(:,3));
axis equal
saveas(b, '.\applied-multivariate-statistics\solutions\chapter-2\sol2.1c.png', 'png')
```



After subtracting off the respective means from both $\mathbf{x}$ and $\mathbf{y}$ (centering), the results of both vectors exist on the same line through the origin, but point in different directions.

2.2

$$\mathbf{A} = \begin{bmatrix} -1 & 3 \\ 4 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 4 & -3 \\ 1 & -2 \\ -2 & 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} 5 \\ -4 \\ 2 \end{bmatrix}$$

(a) $5\mathbf{A}$

$$5\mathbf{A} = 5 \begin{bmatrix} -1 & 3 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} -5 & 15 \\ 20 & 10 \end{bmatrix}$$

(b) $\mathbf{BA}$

$$\mathbf{BA} = \begin{bmatrix} 4 & -3 \\ 1 & -2 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} -16 & 6 \\ -9 & -1 \\ 2 & -6 \end{bmatrix}$$

(c) $\mathbf{A}'\mathbf{B}'$

$$\mathbf{A}'\mathbf{B}' = \begin{bmatrix} -1 & 4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 4 & 1 & -2 \\ -3 & -2 & 0 \end{bmatrix} = \begin{bmatrix} -16 & -9 & 2 \\ 6 & -1 & -6 \end{bmatrix}$$

(d) $\mathbf{C}'\mathbf{B}$

$$\mathbf{C}'\mathbf{B} = \begin{bmatrix} 5 & -4 & 2 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 1 & -2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 12 & -7 \end{bmatrix}$$

(e) Is $\mathbf{AB}$ defined?

No, $\mathbf{A}$ is a 2×2 matrix and $\mathbf{B}$ is a 3×2 matrix, the number of columns in $\mathbf{A} = 2$ is not the same as the number of rows in $\mathbf{B} = 3$, so the two matrices are not conformable.

2.3

Verify the following properties of transpose when

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 4 & 2 \\ 5 & 0 & 3 \end{bmatrix}, \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix},$$

(a) $(\mathbf{A}')' = \mathbf{A}$

$$(\mathbf{A}')' = \left(\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}' \right)' = \left(\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \right)' = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} = \mathbf{A}$$

Matrix $\mathbf{A}$ is symmetric since $\mathbf{A} = \mathbf{A}'$.

(b) $(\mathbf{C}')^{-1} = \left(\mathbf{C}^{-1}\right)'$

$$(\mathbf{C}')^{-1} = \left(\begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}' \right)^{-1} = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}^{-1} = \frac{1}{2 - 12} \begin{bmatrix} 2 & -3 \\ -4 & 1 \end{bmatrix} = \begin{bmatrix} -2/10 & 3/10 \\ 4/10 & -1/10 \end{bmatrix}$$

$$\left(\mathbf{C}^{-1}\right)' = \left(\begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}^{-1} \right)' = \frac{1}{2 - 12} \begin{bmatrix} 2 & -4 \\ -3 & 1 \end{bmatrix}' = \begin{bmatrix} -2/10 & 3/10 \\ 4/10 & -1/10 \end{bmatrix}$$

So $(\mathbf{C}')^{-1} = \left(\mathbf{C}^{-1}\right)'$.

(c) $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$

$$(\mathbf{AB})' = \left(\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 4 & 2 \\ 5 & 0 & 3 \end{bmatrix}\right)' = \begin{bmatrix} 7 & 8 & 7 \\ 16 & 4 & 11 \end{bmatrix}' = \begin{bmatrix} 7 & 16 \\ 8 & 4 \\ 7 & 11 \end{bmatrix}$$

$$\mathbf{B}'\mathbf{A}' = \begin{bmatrix} 1 & 4 & 2 \\ 5 & 0 & 3 \end{bmatrix}'\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}' = \begin{bmatrix} 1 & 5 \\ 4 & 0 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 7 & 16 \\ 8 & 4 \\ 7 & 11 \end{bmatrix}$$

So $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.

(d) For general $\underset{(m\times k)}{\mathbf{A}}$ and $\underset{(k\times l)}{\mathbf{B}}$, $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.

$$(\mathbf{AB})' = \left(\begin{bmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mk} \end{bmatrix}\begin{bmatrix} b_{11} & \dots & b_{1\ell} \\ \vdots & \ddots & \vdots \\ b_{k1} & \dots & b_{k\ell} \end{bmatrix}\right)' = \left(\begin{bmatrix} \sum_{i=1}^k a_{1i}b_{i1} & \dots & \sum_{i=1}^k a_{1i}b_{i\ell} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^k a_{mi}b_{i1} & \dots & \sum_{i=1}^k a_{mi}b_{i\ell} \end{bmatrix}\right)' =$$

$$= \begin{bmatrix} \sum_{i=1}^k a_{1i}b_{i1} & \dots & \sum_{i=1}^k a_{mi}b_{i1} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^k a_{1i}b_{i\ell} & \dots & \sum_{i=1}^k a_{mi}b_{i\ell} \end{bmatrix}$$

$$\mathbf{B}'\mathbf{A}' = \begin{bmatrix} b_{11} & \dots & b_{1\ell} \\ \vdots & \ddots & \vdots \\ b_{k1} & \dots & b_{k\ell} \end{bmatrix}'\begin{bmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mk} \end{bmatrix}' = \begin{bmatrix} b_{11} & \dots & b_{k1} \\ \vdots & \ddots & \vdots \\ b_{1\ell} & \dots & b_{k\ell} \end{bmatrix}\begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1k} & \dots & a_{mk} \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{i=1}^k b_{i1}a_{1i} & \dots & \sum_{i=1}^k b_{i1}a_{mi} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^k b_{i\ell}a_{1i} & \dots & \sum_{i=1}^k b_{i\ell}a_{mi} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^k a_{1i}b_{i1} & \dots & \sum_{i=1}^k a_{mi}b_{i1} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^k a_{1i}b_{i\ell} & \dots & \sum_{i=1}^k a_{mi}b_{i\ell} \end{bmatrix}$$

So $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.

2.4

When $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ exist, prove each of the following.

(a) $(\mathbf{A}')^{-1} = \left(\mathbf{A}^{-1}\right)'$

$$\left(\mathbf{A}^{-1}\right)' = \left(\mathbf{A}^{-1}\right)'\mathbf{I} = \left(\mathbf{A}^{-1}\right)'\mathbf{A}'(\mathbf{A}')^{-1} = \left(\mathbf{A}\mathbf{A}^{-1}\right)'(\mathbf{A}')^{-1} = \mathbf{I}(\mathbf{A}')^{-1} = (\mathbf{A}')^{-1}$$

(b) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

If we define $\mathbf{C} = \mathbf{AB}$ and $\mathbf{D} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, and if **Def 2A.27** is satisifed ($\mathbf{CD} = \mathbf{DC} = \mathbf{I}$), then $\mathbf{D}$ is the inverse of $\mathbf{C}$.

$$\mathbf{CD} = (\mathbf{AB})\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right) = \mathbf{A}\left(\mathbf{BB}^{-1}\right)\mathbf{A}^{-1} = \mathbf{AIA}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}$$

$$\mathbf{DC} = \left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right)(\mathbf{AB}) = \mathbf{B}\left(\mathbf{AA}^{-1}\right)\mathbf{B}^{-1} = \mathbf{BIB}^{-1} = \mathbf{BB}^{-1} = \mathbf{I}$$

Now, **Def 2A.27** is satisifed ($\mathbf{CD} = \mathbf{DC} = \mathbf{I}$), so the inverse of $\mathbf{AB}$ is $\mathbf{B}^{-1}\mathbf{A}^{-1}$. That is, $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Hint: Part a can be proven noting that $\mathbf{AA}^{-1} = \mathbf{I}$, $\mathbf{I} = \mathbf{I}'$, and $\left(\mathbf{AA}^{-1}\right)' = \left(\mathbf{A}^{-1}\right)'\mathbf{A}'$. Part b follows from $\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right)\mathbf{AB} = \mathbf{B}^{-1}\left(\mathbf{A}^{-1}\mathbf{A}\right)\mathbf{B} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$.

2.5

Check that

$$\mathbf{Q} = \begin{bmatrix} \frac{5}{13} & \frac{12}{13} \\ \frac{-12}{13} & \frac{5}{13} \end{bmatrix}$$

is an orthogonal matrix.

If the conditions of **Result 2A.13**, ($\mathbf{AA}' = \mathbf{A}'\mathbf{A} = \mathbf{I}$) are true, then we have an orthogonal matrix.

$$\mathbf{Q}'\mathbf{Q} = \begin{bmatrix} \frac{5}{13} & \frac{-12}{13} \\ \frac{12}{13} & \frac{5}{13} \end{bmatrix} \begin{bmatrix} \frac{5}{13} & \frac{12}{13} \\ \frac{-12}{13} & \frac{5}{13} \end{bmatrix} = \begin{bmatrix} \frac{25}{169} + \frac{144}{169} & \frac{60}{169} - \frac{60}{169} \\ \frac{60}{169} - \frac{60}{169} & \frac{144}{169} + \frac{25}{169} \end{bmatrix} = \begin{bmatrix} \frac{169}{169} & \frac{0}{169} \\ \frac{0}{169} & \frac{169}{169} \end{bmatrix} = \mathbf{I}$$

$$\mathbf{QQ}' = \begin{bmatrix} \frac{5}{13} & \frac{12}{13} \\ \frac{-12}{13} & \frac{5}{13} \end{bmatrix} \begin{bmatrix} \frac{5}{13} & \frac{-12}{13} \\ \frac{12}{13} & \frac{5}{13} \end{bmatrix} = \begin{bmatrix} \frac{25}{169} + \frac{144}{169} & -\frac{60}{169} + \frac{60}{169} \\ -\frac{60}{169} + \frac{60}{169} & \frac{144}{169} + \frac{25}{169} \end{bmatrix} = \begin{bmatrix} \frac{169}{169} & \frac{0}{169} \\ \frac{0}{169} & \frac{169}{169} \end{bmatrix} = \mathbf{I}$$

Because $\mathbf{QQ}' = \mathbf{Q}'\mathbf{Q} = \mathbf{I}$, $\mathbf{Q}$ is an orthofonal matrix.

2.6

Let

$$\mathbf{A} = \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$$

(a) Is $\mathbf{A}$ symmetric?

Yes, $\mathbf{A}' = \mathbf{A}$, so $\mathbf{A}$ is symmetric.

(b) Show that **A** is positive definite.

$$0 = |\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 9-\lambda & -2 \\ -2 & 6-\lambda \end{vmatrix} = (9-\lambda)(6-\lambda) - 4 = \lambda^2 - 15\lambda + 54 - 4 = (\lambda - 5)(\lambda - 10)$$

The two eigenvalues of 5 and 10 are both positive, so from what's in **2.3** on page 63, **A** is positive definite.

2.7

Let **A** be as given in Exercise 2.6.

(a) Determine the eigenvalues and eigenvectors of **A**.

From problem 2.7, the eigenvalues are 5 and 10.To get the eigenvectors,
$\underline{\lambda_1 = 5}$:

$$\mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1 \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix} \Rightarrow \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 5 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$9x_1 - 2x_2 = 5x_1 \Rightarrow 4x_1 = 2x_2 \Rightarrow 2x_1 = x_2$$

and

$$-2x_1 + 6x_2 = 5x_1 \Rightarrow x_2 = 2x_1$$

So $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and normalizing, $\mathbf{e}_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$.
$\underline{\lambda_2 = 10}$:

$$\mathbf{A}\mathbf{x}_2 = \lambda_2\mathbf{x}_2 \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix} \Rightarrow \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 10 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$9x_1 - 2x_2 = 10x_1 \Rightarrow x_1 = -2x_2$$

and

$$-2x_1 + 6x_2 = 10x_1 \Rightarrow 12x_1 = -6x_2 \Rightarrow x_1 = -2x_2$$

So $\mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and normalizing, $\mathbf{e}_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$.

(b) Write the spectral decomposition of **A**.

The spectral decomposition would be,

$$\mathbf{A} = \sum_{k=1}^{2} \lambda_k \mathbf{e}_k \mathbf{e}_k' = 5 \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}' + 10 \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}' =$$

$$= 5 \begin{bmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{bmatrix} + 10 \begin{bmatrix} 4/5 & -2/5 \\ -2/5 & 1/5 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} + \begin{bmatrix} 8 & -4 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$$

(c) Find $\mathbf{A}^{-1}$.

Using the spectral decomposition for the inverse in **(2-21)** on page 66,

$$\mathbf{A}^{-1} = \left(\mathbf{P}\mathbf{\Lambda}\mathbf{P}'\right)^{-1} = \left(\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2\end{bmatrix}\mathbf{\Lambda}\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2\end{bmatrix}'\right)^{-1} = \begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2\end{bmatrix}\mathbf{\Lambda}^{-1}\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2\end{bmatrix}' =$$

$$= \begin{bmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix}\begin{bmatrix}1/5 & 0\\ 0 & 1/10\end{bmatrix}\begin{bmatrix}1/\sqrt{5} & 2/\sqrt{5}\\ -2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix} = \frac{1}{5}\begin{bmatrix}1/5 & -2/10\\ 2/5 & 1/10\end{bmatrix}\begin{bmatrix}1 & 2\\ -2 & 1\end{bmatrix} =$$

$$= \frac{1}{5}\begin{bmatrix}30/50 & 10/50\\ 10/50 & 45/50\end{bmatrix} = \frac{1}{50}\begin{bmatrix}6 & 2\\ 2 & 9\end{bmatrix}$$

By direct computation,

$$\mathbf{A}^{-1} = \begin{bmatrix}9 & -2\\ -2 & 6\end{bmatrix}^{-1} = \frac{1}{54-4}\begin{bmatrix}6 & 2\\ 2 & 9\end{bmatrix} = \frac{1}{50}\begin{bmatrix}6 & 2\\ 2 & 9\end{bmatrix}$$

(d) Find the eigenvalues and eigenvectors of $\mathbf{A}^{-1}$.

Using **2-21** on page 66 again, the eigenvalues of $\mathbf{A}^{-1}$ are the reciprocal of the eigenvalues of $\mathbf{A}$ and the eigenvectors are the same as those for $\mathbf{A}$.

$$\mathbf{\Lambda}^{-1} = \begin{bmatrix}1/5 & 0\\ 0 & 1/10\end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2\end{bmatrix} = \begin{bmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix}$$

If we did it by-hand, setting $\mathbf{B} = \mathbf{A}^{-1}$,

$$0 = |\mathbf{B} - \lambda\mathbf{I}| = \frac{1}{50}\begin{vmatrix}6-\lambda & 2\\ 2 & 9-\lambda\end{vmatrix} = \frac{1}{50}\left[(6-\lambda)(9-\lambda) - 54 + 4\right] = \frac{1}{50}\left[(6-\lambda)(9-\lambda) - 50\right] =$$

$$\frac{1}{50}\left(\lambda^2 - 15\lambda + 50\right) = \left(\lambda - \frac{5}{50}\right)\left(\lambda - \frac{10}{50}\right) = \left(\lambda - \frac{1}{10}\right)\left(\lambda - \frac{1}{5}\right)$$

These eigenvalues are the same as the reciprocal of the eigenvalues for $\mathbf{A}$.

2.8

Given the matrix

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

find the eigenvalues λ_1 and λ_2 and the associated eigenvectors $\mathbf{e}_1$ and $\mathbf{e}_2$. Determine the spectral decomposition **(2-16)** of $\mathbf{A}$.

$$0 = |\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix}1-\lambda & 2\\ 2 & -2-\lambda\end{vmatrix} = (1-\lambda)(-2-\lambda) - 4 = -2 + 2\lambda - \lambda + \lambda^2 - 4 =$$

$$= \lambda^2 + \lambda - 6 = (\lambda + 3)(\lambda - 2)$$

The eigenvalues are -3 and 2.
$\underline{\lambda_1 = -3}$:

$$\mathbf{A}\mathbf{x}_1 = \lambda \mathbf{x}_1 \Rightarrow \mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -3x_1 \\ -3x_2 \end{bmatrix}$$

$$x_1 + 2x_2 = -3x_1 \Rightarrow -4x_1 = 2x_2 \Rightarrow -2x_1 = x_2$$

and

$$2x_1 - 2x_2 = -3x_1 \Rightarrow 2x_1 = -x_2 \Rightarrow -2x_1 = x_2$$

So $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and normalizing $\mathbf{e}_1 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$.
$\underline{\lambda_2 = 2}$:

$$\mathbf{A}\mathbf{x}_2 = \lambda \mathbf{x}_2 \Rightarrow \mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$$

$$x_1 + 2x_2 = 2x_1 \Rightarrow 2x_2 = x_1$$

and

$$2x_1 - 2x_2 = 2x_1 \Rightarrow x_2 = 4x_1 \Rightarrow x_1 = 2x_2$$

So $\mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and normalizing $\mathbf{e}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$.

$$\mathbf{A} = \sum_{k=1}^{2} \lambda_k \mathbf{e}_k \mathbf{e}_k' = -3 \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} + 2 \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} =$$

$$= \frac{1}{5}\left(-3 \begin{bmatrix} 1 & -2 \\ 2 & 4 \end{bmatrix} + 2 \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \right) = \frac{1}{5}\left(- \begin{bmatrix} 3 & -6 \\ -6 & 12 \end{bmatrix} + \begin{bmatrix} 8 & 4 \\ 4 & 2 \end{bmatrix} \right) =$$

$$\frac{1}{5} \begin{bmatrix} 5 & 10 \\ 10 & -10 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$$

2.9

Let $\mathbf{A}$ be as in Exercise 2.8.

(a) Find $\mathbf{A}^{-1}$.

Using **(2-21)**,

$$\mathbf{A}^{-1} = \mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}' = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} -1/3 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} =$$

$$= \frac{1}{5} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} -1/3 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} -1/3 & 2/3 \\ 1 & 1/2 \end{bmatrix} =$$

$$= \frac{1}{5} \begin{bmatrix} 5/3 & 5/3 \\ 5/3 & -5/6 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 \\ 1/3 & -1/6 \end{bmatrix}$$

Using direct computation,

$$\mathbf{A}^{-1} = \frac{1}{-2-4}\begin{bmatrix} -2 & -2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 \\ 1/3 & -1/6 \end{bmatrix}$$

(b) Compute the eigenvalues and eigenvectors of $\mathbf{A}^{-1}$.

Also from **(2-21)** on page 66, we can see that the eigenvalues of $\mathbf{A}^{-1}$ are the reciprocal of the eigenvalues of $\mathbf{A}$, so

$$\mathbf{\Lambda}^{-1} = \begin{bmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_2 \end{bmatrix} = \begin{bmatrix} -1/3 & 0 \\ 0 & 1/2 \end{bmatrix}$$

and the eigenvectors for $\mathbf{A}^{-1}$ are the same as those for $\mathbf{A}$,

$$\mathbf{P} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

(c) Write the spectral decomposition of $\mathbf{A}^{-1}$, and compare it with that of $\mathbf{A}$ from Exercise 2.8.

$$\mathbf{A}^{-1} = \sum_{k=1}^{2} \frac{1}{\lambda_k}\mathbf{e}_k\mathbf{e}_k' = -\frac{1}{3}\begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}\begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} =$$

$$= \frac{1}{5}\left(-\frac{1}{3}\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}\right) = \frac{1}{30}\left(-\begin{bmatrix} 2 & -4 \\ -4 & 8 \end{bmatrix} + \begin{bmatrix} 12 & 6 \\ 6 & 3 \end{bmatrix}\right) =$$

$$= \frac{1}{30}\begin{bmatrix} 10 & 10 \\ 10 & -5 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 \\ 1/3 & -1/6 \end{bmatrix}$$

In the spectral decomposition of both $\mathbf{A}$ and $\mathbf{A}^{-1}$ the matrices created for all of the $\mathbf{e}_k\mathbf{e}_k'$ components are the same. The difference is in the eigenvalues. The eigenvalues for $\mathbf{A}$ are λ_k and the eigenvalues for $\mathbf{A}^{-1}$ are $1/\lambda_k$.

2.10

Consider the matrices

$$\mathbf{A} = \begin{bmatrix} 4 & 4.001 \\ 4.001 & 4.002 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 4 & 4.001 \\ 4.001 & 4.002001 \end{bmatrix}$$

These matrices are identical except for a small difference in the (2, 2) position. Moreover, the columns of $\mathbf{A}$ (and $\mathbf{B}$) are nearly dependent. Show that $\mathbf{A}^{-1} \doteq (-3)\mathbf{B}^{-1}$. Consequently, small changes — perhaps caused by rounding — can give substantially different inverses.

$$\mathbf{A}^{-1} = \frac{1}{16.008 - 16.008001}\begin{bmatrix} 4.002 & -4.001 \\ -4.001 & 4 \end{bmatrix} = \frac{1}{-0.000001}\begin{bmatrix} 4.002 & -4.001 \\ -4.001 & 4 \end{bmatrix} =$$

$$= -1000000 \begin{bmatrix} 4.002 & -4.001 \\ -4.001 & 4 \end{bmatrix} = \begin{bmatrix} -4002000 & 4001000 \\ 4001000 & -4000000 \end{bmatrix}$$

$$\mathbf{B}^{-1} = \frac{1}{16.008004 - 16.008001} \begin{bmatrix} 4.002001 & -4.001 \\ -4.001 & 4 \end{bmatrix} = \frac{1}{0.000003} \begin{bmatrix} 4.002001 & -4.001 \\ -4.001 & 4 \end{bmatrix} =$$

$$= 333333.\bar{3} \begin{bmatrix} 4.002001 & -4.001 \\ -4.001 & 4 \end{bmatrix} = \begin{bmatrix} 1334000.3333331999333 & -1333666.6666665333 \\ -1333666.6666665333 & 1333333.3333332 \end{bmatrix}$$

$$(-3)\mathbf{B}^{-1} = -3 \begin{bmatrix} 1334000.3333331999333 & -1333666.6666665333 \\ -1333666.6666665333 & 1333333.3333332 \end{bmatrix} =$$

$$= \begin{bmatrix} -4002000.9999995997999 & 4000999.9999995999 \\ 4000999.9999995999 & -3999999.9999996 \end{bmatrix} \doteq \mathbf{A}^{-1}$$

Used $\frac{1}{|B|} = \frac{1}{0.000003} = 333333.3333333$ for computation.

2.11

Show that the determinant of the $p \times p$ diagonal matrix $\mathbf{A} = a_{ij}$ with $a_{ij} = 0$, $i \neq j$, is given by the product of the diagonal elements; thus, $|\mathbf{A}| = a_{11}a_{22}, \ldots a_{pp}$. *Hint:* By **Definition 2A.24**, $|\mathbf{A}| = a_{11}\mathbf{A}_{11} + 0 + \cdots + 0$. Repeat for the submatrix $\mathbf{A}_{11}$ obtained by deleting the first row and first column of $\mathbf{A}$.

$$|\mathbf{A}| = \sum_{j=1}^{p} a_{1j} \left| A_{1j} \right| (-1)^{1+j} =$$

$$= a_{11} \left| \mathbf{A}_{11} \right| + 0 + \cdots + 0 =$$

$$= a_{11} \left(\sum_{j=2}^{p} a_{2j} \left| A_{2j} \right| (-1)^{2+j} \right) =$$

$$= a_{11} \left(a_{22} \left| \mathbf{A}_{22} \right| + 0 + \cdots + 0 \right) =$$

$$\vdots$$

$$= a_{11}a_{22} \ldots a_{p-2,p-2} \left| \mathbf{A}_{p-2,p-2} \right|$$

$$= a_{11}a_{22} \ldots a_{p-2,p-2} \left(a_{p-1,p-1}a_{pp} - 0 \right)$$

$$= a_{11}a_{22} \ldots a_{p-2,p-2}a_{p-1,p-1}a_{pp} = \prod_{i=1}^{p} a_{ii}$$

2.12

Show that the determinant of a square symmetric $p \times p$ matrix $\mathbf{A}$ can be expressed as the product of its eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$; that is, $|\mathbf{A}| = \prod_{i=1}^{p} \lambda_i$.
Hint: From **(2-16)** and **(2-20)**, $\mathbf{A} = \mathbf{P \Lambda P'}$ with $\mathbf{P'P} = \mathbf{I}$. From **Result 2A.ll(e)**, $|A| = |\mathbf{P\Lambda P'}| = |\mathbf{P}|\,|\mathbf{\Lambda P'}| = |\mathbf{P}|\,|\mathbf{\Lambda}|\,|\mathbf{P'}| = |\mathbf{\Lambda}|\,|\mathbf{I}|$,since $|\mathbf{I}| = |\mathbf{P'P}| = |\mathbf{P'}|\,|\mathbf{P}|$. Apply **Exercise 2.11**.

$$|\mathbf{A}| = |\mathbf{P\Lambda P'}| \overset{2A.11(e)}{=} |\mathbf{P}|\,|\mathbf{\Lambda P'}| \overset{2A.11(e)}{=} |\mathbf{P}|\,|\mathbf{\Lambda}|\,|\mathbf{P'}| = |\mathbf{\Lambda}|\,|\mathbf{P}|\,|\mathbf{P'}| =$$

$$= |\mathbf{\Lambda}|\,|\mathbf{PP'}| = |\mathbf{\Lambda}|\,|\mathbf{I}| = |\mathbf{\Lambda}|\,(1) = |\mathbf{\Lambda}| \overset{\text{Exercise 2.11}}{=} \prod_{i=1}^{p} \lambda_i$$

2.13

Show that $|\mathbf{Q}| = +1$ or -1 if $\mathbf{Q}$ is a $p \times p$ orthogonal matrix.
Hint: $|\mathbf{QQ'}| = |\mathbf{I}|$. Also, from **Result 2A.ll**, $|\mathbf{Q}|\,|\mathbf{Q'}| = |\mathbf{Q}|^2$. Thus, $|\mathbf{Q}|^2 = |\mathbf{I}|$. Now use **Exercise 2.11**.

We know $\mathbf{Q}$ is orthogonal iff $\mathbf{QQ'} = \mathbf{Q'Q} = \mathbf{I}$ from **Definition 2A.13**.

$$|\mathbf{Q'Q}| = |\mathbf{Q'}|\,|\mathbf{Q}| = |\mathbf{Q}|\,|\mathbf{Q}| = |\mathbf{Q}|^2 = |\mathbf{I}| \overset{\text{Exercise 2.11}}{=} 1$$

So

$$|\mathbf{Q}| = \left(|\mathbf{Q}|^2\right)^{1/2} = (|\mathbf{I}|)^{1/2} = \pm\sqrt{1} = \pm 1$$

2.14

Show that $\underset{(p\times p)}{\mathbf{Q'}}\underset{(p\times p)}{\mathbf{A}}\underset{(p\times p)}{\mathbf{Q}}$ and $\underset{p\times p}{\mathbf{A}}$ have the same eigenvalues if $\mathbf{Q}$ is orthogonal.
Hint: Let λ be an eigenvalue of $\mathbf{A}$. Then $0 = |\mathbf{A} - \lambda\mathbf{I}|$. By **Exercise 2.13** and **Result 2A.11(e)**, we can write $0 = |\mathbf{Q'}|\,|\mathbf{A} - \lambda\mathbf{I}|\,|\mathbf{Q}| = |\mathbf{Q'AQ} - \lambda\mathbf{I}|$, since $\mathbf{Q'Q} = \mathbf{I}$ from **Exercise 2.11**.

Answer is pretty much in the hint. We already know $|\mathbf{Q'Q}| = |\mathbf{I}| = 1$. We also already know $|\mathbf{Q'Q}| = |\mathbf{Q'}|\,|\mathbf{Q}|$ from **Result 2A.11(e)**. Not mentioned in the book, but the determinant is a scalar output, so it's also commutative, so $|\mathbf{Q'Q}| = |\mathbf{Q'}|\,|\mathbf{Q}| = |\mathbf{Q}|\,|\mathbf{Q'}|$.

$$0 = |\mathbf{A} - \lambda\mathbf{I}| = |\mathbf{Q'Q}|\,|\mathbf{A} - \lambda\mathbf{I}| = |\mathbf{Q'}|\,|\mathbf{Q}|\,|\mathbf{A} - \lambda\mathbf{I}| = |\mathbf{Q'}|\,|\mathbf{A} - \lambda\mathbf{I}|\,|\mathbf{Q}| =$$

$$= |\mathbf{Q'AQ} - \lambda\mathbf{Q'IQ}| = |\mathbf{Q'AQ} - \lambda\mathbf{Q'Q}| = |\mathbf{Q'AQ} - \lambda\mathbf{I}|$$

There it is, $0 = |\mathbf{A} - \lambda\mathbf{I}| = |\mathbf{Q'AQ} - \lambda\mathbf{I}|$, so the eigenvalues of $\mathbf{A}$ are the same as those for $\mathbf{Q'AQ}$.

2.15

A quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x}$ is said to be positive definite if the matrix $\mathbf{A}$ is positive definite. Is the quadratic form $3x_1^2 + 3x_2^2 - 2x_1x_2$ positive definite?

Converting $3x_1^2 + 3x_2^2 - 2x_1x_2$ to a matrix, we'd have,

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 (3x_1 - x_2) + x_2 (-x_1 + 3x_2) = 3x_1^2 + 3x_2^2 - 2x_1x_2$$

So now,

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

Now check if $\mathbf{A}$ has positive eigenvalues,

$$0 = |\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 1 = \lambda^2 - 6\lambda + 9 - 1 = (\lambda - 2)(\lambda - 4)$$

The eigenvalues are $\{\lambda_1, \lambda_2\} = \{2, 4\}$, so since $\lambda_i > 0$, i.e., the eigenvalues are all positive, the matrix $\mathbf{A}$ is positive definite.

2.16

Consider an arbitrary $n \times p$ matrix $\mathbf{A}$. Then $\mathbf{A}'\mathbf{A}$ is a symmetric $p \times p$ matrix. Show that $\mathbf{A}'\mathbf{A}$ is necessarily nonnegative definite.
Hint: Set $\mathbf{y} = \mathbf{A}\mathbf{x}$ so that $\mathbf{y}'\mathbf{y} = \mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x}$.

Ignoring the hint, could use the Singular-Value Decomposition from **Result 2A.15** for nonsquare matrices,

$$\mathbf{A}'\mathbf{A} = (\mathbf{U}\mathbf{\Lambda}\mathbf{V}')' (\mathbf{U}\mathbf{\Lambda}\mathbf{V}') = \mathbf{V}\mathbf{\Lambda}\mathbf{U}'\mathbf{U}\mathbf{\Lambda}\mathbf{V}' = \mathbf{V}\mathbf{\Lambda}\mathbf{I}\mathbf{\Lambda}\mathbf{V}' = \mathbf{V}\mathbf{\Lambda}\mathbf{\Lambda}\mathbf{V}' = \mathbf{V}\mathbf{\Lambda}^2\mathbf{V}'$$

The eigenvalues in $\mathbf{\Lambda}^2$ are all squared values of the eigenvalues of $\mathbf{\Lambda}$, so they are either zero or a positive value. Thus, $\mathbf{A}'\mathbf{A}$ is nonnegative definite.

Using the hint, $\mathbf{y} = \mathbf{A}\mathbf{x}$,

$$\mathbf{y}'\mathbf{y} = (\mathbf{A}\mathbf{x})' (\mathbf{A}\mathbf{x}) = \mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{B}\mathbf{x}$$

As explained on page 62 **(2-17)**, this is in quadratic form. The matrix, $\mathbf{B} = \mathbf{A}'\mathbf{A}$, is $p \times p$ and is nonnegative definite. This could also be explained as, $\mathbf{y}'\mathbf{y} = y_1^2 + \cdots + y_p^2 = \|\mathbf{y}\|^2$, the sum of squared values, and so the sum cannot be negative, so must be at least zero (nonnegative definite).

2.17

Prove that every eigenvalue of a $k \times k$ positive definite matrix $\mathbf{A}$ is positive.
Hint: Consider the definition of an eigenvalue, where $\mathbf{A}\mathbf{e} = \lambda\mathbf{e}$. Multiply on the left by $\mathbf{e}'$ so that $\mathbf{e}'\mathbf{A}\mathbf{e} = \lambda\mathbf{e}'\mathbf{e}$.

Using the hint,

$$\mathbf{A}\mathbf{e} = \lambda\mathbf{e}$$
$$\Rightarrow \mathbf{e}'\mathbf{A}\mathbf{e} = \mathbf{e}'\lambda\mathbf{e}$$
$$\Rightarrow \mathbf{e}'\mathbf{A}\mathbf{e} = \lambda\mathbf{e}'\mathbf{e}$$
$$\Rightarrow \mathbf{e}'\mathbf{A}\mathbf{e} = \lambda(1)$$
$$\Rightarrow \mathbf{e}'\mathbf{A}\mathbf{e} = \lambda$$

We now have a positive definite form, like in the definition of positive definiteness **(2-18)**, $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$, so all of the eigenvalues nust be positive, that is,

$$\Rightarrow \mathbf{e}'\mathbf{A}\mathbf{e} = \lambda > 0$$

2.18

Consider the sets of points (x_1, x_2) whose 'distances' from the origin are given by

$$c^2 = 4x_1^2 + 3x_2^2 - 2\sqrt{2}x_1x_2$$

for $c^2 = 1$ and for $c^2 = 4$. Determine the major and minor axes of the ellipses of constant distances and their associated lengths. Sketch the ellipses of constant distances and comment on their positions. What will happen as c^2 increases?

Converting the quadratic polynomial to a matrix,

$$c^2 = 4x_1^2 + 3x_2^2 - 2\sqrt{2}x_1x_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & -\sqrt{2} \\ -\sqrt{2} & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}'\mathbf{A}\mathbf{x}$$

Finding the eigenvalues and eigenvectors,

$$0 = |\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 4-\lambda & -\sqrt{2} \\ -\sqrt{2} & 3-\lambda \end{vmatrix} = (4-\lambda)(3-\lambda) - 2 = \lambda^2 - 7\lambda + 12 - 2 =$$
$$= (\lambda - 2)(\lambda - 5)$$

The eigenvalues are $\{\lambda_1, \lambda_2\} = \{2, 5\}$. Finding the eigenvectors,
<u>For $\lambda_1 = 2$:</u>

$$\mathbf{A}\mathbf{x}_1 - \lambda_1\mathbf{x}_1 =$$
$$\begin{bmatrix} 4 & -\sqrt{2} \\ -\sqrt{2} & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & -\sqrt{2} \\ -\sqrt{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} 2 & -\sqrt{2} \\ -\sqrt{2} & 1 \end{bmatrix} \xrightarrow{\text{Row } 2 + (\sqrt{2}/2)\text{Row } 1} \begin{bmatrix} 2 & -\sqrt{2} \\ 0 & 0 \end{bmatrix}$$

So $2x_1 - \sqrt{2}x_2 = 0 \Rightarrow 2x_1 = \sqrt{2}x_2 \Rightarrow x_1 = \frac{\sqrt{2}}{2}x_2$. Pick,

$$\mathbf{x}_1 = \begin{bmatrix} \sqrt{2}/2 \\ 1 \end{bmatrix} \Rightarrow \mathbf{e}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}\|} = \frac{1}{\sqrt{3/2}} \begin{bmatrix} \sqrt{2}/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{bmatrix}$$

<u>For $\lambda_1 = 5$:</u>

$$\mathbf{A}\mathbf{x}_2 - \lambda_2\mathbf{x}_2 =$$

$$\begin{bmatrix} 4 & -\sqrt{2} \\ -\sqrt{2} & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5x_1 \\ 5x_2 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & -\sqrt{2} \\ -\sqrt{2} & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 2 \end{bmatrix} \xrightarrow{\text{Row 2}-(\sqrt{2})\text{Row 1}} \begin{bmatrix} 1 & \sqrt{2} \\ 0 & 0 \end{bmatrix}$$

So $x_1 + \sqrt{2}x_2 = 0 \Rightarrow x_1 = -\sqrt{2}x_2 \rightarrow x_2 = -\frac{\sqrt{2}}{2}x_1$. Pick,

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -\sqrt{2}/2 \end{bmatrix} \Rightarrow \mathbf{e}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}\|} = \frac{1}{\sqrt{3/2}} \begin{bmatrix} 1 \\ -\sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix}$$

We now have all the parts,

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$$

and

$$\mathbf{P} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & \sqrt{2}/\sqrt{3} \\ \sqrt{2}/\sqrt{3} & -1/\sqrt{3} \end{bmatrix}$$

<u>As an aside:</u>

$$\mathbf{A} = \lambda_1\mathbf{e}_1\mathbf{e}_1' + \lambda_2\mathbf{e}_2\mathbf{e}_2' \Rightarrow$$

$$\Rightarrow \mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\left(\lambda_1\mathbf{e}_1\mathbf{e}_1' + \lambda_2\mathbf{e}_2\mathbf{e}_2'\right)\mathbf{x} = \lambda_1\mathbf{x}'\mathbf{e}_1\mathbf{e}_1'\mathbf{x} + \lambda_2\mathbf{x}'\mathbf{e}_2\mathbf{e}_2'\mathbf{x} =$$

$$= \lambda_1\mathbf{x}'\mathbf{e}_1(\mathbf{x}'\mathbf{e}_1)' + \lambda_2\mathbf{x}'\mathbf{e}_2(\mathbf{x}'\mathbf{e}_2)' = \lambda_1(\mathbf{x}'\mathbf{e}_1)^2 + \lambda_2(\mathbf{x}'\mathbf{e}_2)^2 = \lambda_1 y_1^2 + \lambda_2 y_2^2 =$$

$$\Rightarrow c^2 = \mathbf{x}'\mathbf{A}\mathbf{x} = \lambda_1 y_1^2 + \lambda_2 y_2^2$$

$$\Rightarrow \frac{\lambda_1 y_1^2}{c^2} + \frac{\lambda_2 y_2^2}{c^2} = 1$$

$$\Rightarrow \frac{y_1^2}{\left(c/\sqrt{\lambda_1}\right)^2} + \frac{y_2^2}{\left(c/\sqrt{\lambda_2}\right)^2} = 1$$

The formula for a horizontal ellipse is,

$$\frac{(y_1 - h)^2}{b^2} + \frac{(y_2 - k)^2}{a^2} = 1$$

Where the ellipse is centered at $(y_1, y_2) = (h, k)$. The major axis are at $\pm b$ and the minor axis are at $\pm a$. Here, for quadratic form we are centered at the origin, so $h = k = 0$ the major axis are in the direction of $\mathbf{e}_1$ with length $b = \pm c/\sqrt{\lambda_1}$, and the minor axis are in the direction of $\mathbf{e}_2$ with length $\pm a = c/\sqrt{\lambda_2}$. Note that $0 \leq \lambda_1 \leq \lambda_2$.

For $c^2 = 1$ with our data, the major axis are in the direction of

$$\mathbf{e}_1 = \begin{bmatrix} 1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{bmatrix}$$

with length $\pm 1/\sqrt{2}$. These are the red circles in figure 1.
The minor axis are in the direction of

$$\mathbf{e}_2 = \begin{bmatrix} \sqrt{2}/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix}$$

with length $\pm 1/\sqrt{5}$. These are the green circles in figure 1. When $c^2 = 4$, the direction vectors stay the same, but the lengths change to $\pm 2/\sqrt{2}$ in the $\mathbf{e}_1$ direction and $\pm 2/\sqrt{5}$ in the $\mathbf{e}_2$ direction. The major and minor axis are represented in figure 2, where the red circles are the major axis and the green circles are the minor axis. The plot of the $c^2 = 1$ ellipse is in figure 1, and the plot of the ellipse for $c^2 = 4$ is in figure 2. As c^2 increases the ellipse grows larger maintaining the same shape.

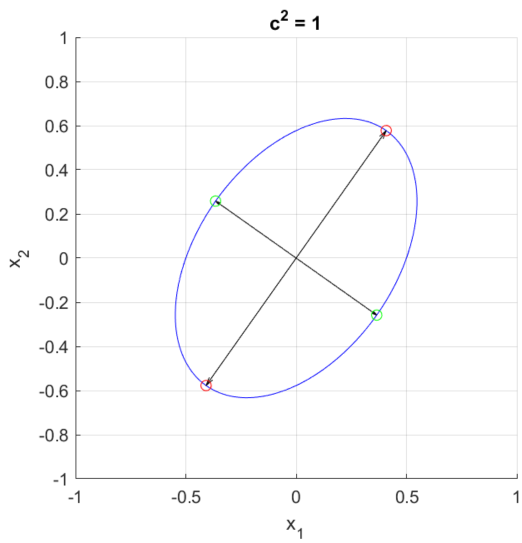


Figure 1: When $c^2 = 1$

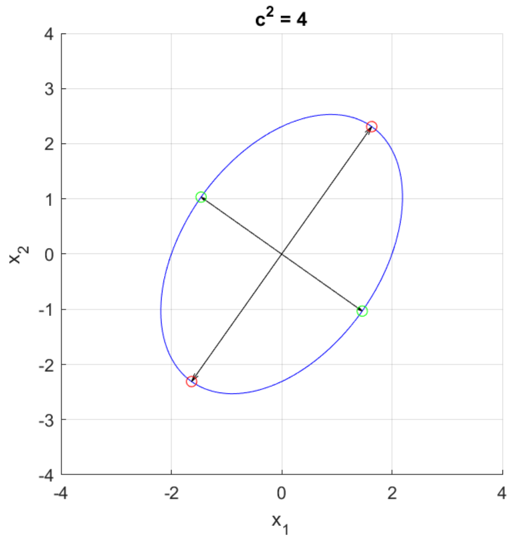


Figure 2: When $c^2 = 4$

MATLAB code:

```
A = [4 -sqrt(2); -sqrt(2) 3];
[V,D] = eig(A);
rref(A - D(1,1)*eye(width(A)))
rref(A - D(2,2)*eye(width(A)))
MyPlotEllipse(V,D,1,'sol2.18.c1')
MyPlotEllipse(V,D,4,'sol2.18.c4')
```

```
function [] = MyPlotEllipse(V, D, c, fName
    )
% Compute points corresponding to axis-oriented
    ellipse.
% Where to center the ellipse.
xc = 0;
yc = 0;
% The length in the major and minor axis.
b = c/sqrt(D(1,1));
a = c/sqrt(D(2,2));
theta = acos(-V(:,1)'*[1 ; 0]); % acos(1/sqrt(3));

t = linspace(0, 2*pi, 200);
xt = b * cos(t) + xc;
yt = a * sin(t) + yc;

% Apply rotation by angle theta (in radians).
```

```
    cot = cos(theta); sit = sin(theta);
    x = xt * cot - yt * sit;
    y = xt * sit + yt * cot;

    hold on
        % Plot the ellipse.
        p=plot(x, y, '-', 'Color', 'blue');

        % Plot the vector for the major axis.
        quiver(0, 0, V(1,1), V(2,1), c, 'color', 'k');
        quiver(0, 0, -V(1,1), -V(2,1), c, 'color', 'k');

        % Plot the vector for the minor axis.
        quiver(0, 0, V(1,2), V(2,2), c, 'color', 'k');
        quiver(0, 0, -V(1,2), -V(2,2), c, 'color', 'k');

        % Plot red point for major axis.
        plot((c/sqrt(D(1,1)))*V(1,1), (c/sqrt(D(1,1)))*V(2,1), 'o', 'Color', 'red');
        plot(-(c/sqrt(D(1,1)))*V(1,1), -(c/sqrt(D(1,1)))*V(2,1), 'o', 'Color', 'red');

        % Plot green point for minor axis.
        plot((c/sqrt(D(2,2)))*V(1,2), (c/sqrt(D(2,2)))*V(2,2), 'o', 'Color', 'green');
        plot(-(c/sqrt(D(2,2)))*V(1,2), -(c/sqrt(D(2,2)))*V(2,2), 'o', 'Color', 'green');
        title(append('c^2 = ', num2str(c)))
        grid on
        pbaspect([1 1 1])
    hold off
    saved_file = append('.\applied-multivariate-statistics\solutions\chapter-2\', fName, '.png');
    saveas(p, saved_file, 'png')
end
```

2.19

Let $\underset{(m\times m)}{\mathbf{A}^{1/2}} = \sum_{i=1}^{m} \sqrt{\lambda_i}\mathbf{e}_i\mathbf{e}_i' = \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}'$, where $\mathbf{P}\mathbf{P}' = \mathbf{P}'\mathbf{P} = \mathbf{I}$. (The λ_i's and the $\mathbf{e}$'s are the eigenvalues and associated normalized eigenvectors of the matrix $\mathbf{A}$.) Show Properties **(1)-(4)** of the square-root matrix in **(2-22)**.

(1) $\left(\mathbf{A}^{1/2}\right)' = \mathbf{A}^{1/2}$ (that is, $\mathbf{A}$ is symmetric).

$$\left(\mathbf{A}^{1/2}\right)' = \left(\mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}'\right)' \overset{\text{Exercise 2.3(c)}}{=} (\mathbf{P}')'\left(\mathbf{\Lambda}^{1/2}\right)'(\mathbf{P})' =$$

$$= \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}' = \mathbf{A}^{1/2}$$

(2) $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$.

$$\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \left(\mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}'\right)\left(\mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}'\right) = \mathbf{P}\mathbf{\Lambda}^{1/2}\left(\mathbf{P}'\mathbf{P}\right)\mathbf{\Lambda}^{1/2}\mathbf{P}' =$$

$$= \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{I}\mathbf{\Lambda}^{1/2}\mathbf{P}' = \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{P}' =$$

$$= \mathbf{P}\begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_m} \end{bmatrix}\begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_m} \end{bmatrix}\mathbf{P}' =$$

$$\mathbf{P}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix}\mathbf{P}' = \mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \mathbf{A}$$

(3) $\left(\mathbf{A}^{1/2}\right)^{-1} = \sum_{i=1}^{k} \frac{1}{\sqrt{\lambda_i}}\mathbf{e}_i\mathbf{e}_i' = \mathbf{P}\mathbf{\Lambda}^{-1/2}\mathbf{P}'$, where $\mathbf{\Lambda}^{-1/2}$ is a diagonal matrix with $1/\sqrt{\lambda_i}$ as the ith diagonal element.

First off, something useful,

$$\mathbf{\Lambda}^{-1/2}\mathbf{\Lambda}^{1/2} = \begin{bmatrix} 1/\sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & 1/\sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sqrt{\lambda_m} \end{bmatrix}\begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_m} \end{bmatrix} = \mathbf{I}$$

$$\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{-1/2} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_m} \end{bmatrix}\begin{bmatrix} 1/\sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & 1/\sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sqrt{\lambda_m} \end{bmatrix} = \mathbf{I}$$

We have $\mathbf{\Lambda}^{-1/2}\mathbf{\Lambda}^{1/2} = \mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{-1/2} = \mathbf{I}$, so $\mathbf{\Lambda}^{-1/2} = \left(\mathbf{\Lambda}^{1/2}\right)^{-1}$ is the inverse of $\mathbf{\Lambda}^{1/2}$ by **Definition 2A.27**.

$$\left(\mathbf{A}^{1/2}\right)^{-1} = \left(\mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}'\right)^{-1} \overset{Exercise 2.4(b)}{=} (\mathbf{P}')^{-1}\left(\mathbf{\Lambda}^{1/2}\right)^{-1}\mathbf{P}^{-1} =$$

$$\overset{\mathbf{P}'=\mathbf{P}^{-1}}{=} (\mathbf{P}')'\left(\mathbf{\Lambda}^{1/2}\right)^{-1}\mathbf{P}' = \mathbf{P}\left(\mathbf{\Lambda}^{1/2}\right)^{-1}\mathbf{P}'$$

$$= \mathbf{P}\mathbf{\Lambda}^{-1/2}\mathbf{P}' = \sum_{i=1}^{m} \frac{1}{\sqrt{\lambda_i}}\mathbf{e}_i\mathbf{e}_i'$$

(4) $\mathbf{A}^{1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1/2}\mathbf{A}^{1/2} = \mathbf{I}$, and $\mathbf{A}^{-1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1}$, where $\mathbf{A}^{-1/2} = \left(\mathbf{A}^{1/2}\right)^{-1}$.

Above in **(3)** it was shown that $\mathbf{A}^{1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1/2}\mathbf{A}^{1/2} = \mathbf{I}$, so $\mathbf{A}^{-1/2} = \left(\mathbf{A}^{1/2}\right)^{-1}$.

$$\mathbf{A}^{-1/2}\mathbf{A}^{-1/2} = \begin{bmatrix} 1/\sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & 1/\sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sqrt{\lambda_m} \end{bmatrix} \begin{bmatrix} 1/\sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & 1/\sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sqrt{\lambda_m} \end{bmatrix} =$$

$$= \begin{bmatrix} 1/\lambda_1 & 0 & \cdots & 0 \\ 0 & 1/\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\lambda_m \end{bmatrix} = \mathbf{A}^{-1}$$

2.20

Determine the square-root matrix $\mathbf{A}^{1/2}$, using the matrix $\mathbf{A}$ in **Exercise 2.3**. Also, determine $\mathbf{A}^{-1/2}$, and show that $\mathbf{A}^{1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1/2}\mathbf{A}^{1/2} = \mathbf{I}$.

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$

$$0 = |\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = (2-\lambda)(3-\lambda) - 1 = \lambda^2 - 5\lambda + 6 - 1 =$$

$$= \left(\lambda - \frac{5-\sqrt{5}}{2}\right)\left(\lambda - \frac{5+\sqrt{5}}{2}\right)$$

The eigenvalues are $\{\lambda_1, \lambda_2\} = \left\{\frac{5-\sqrt{5}}{2}, \frac{5+\sqrt{5}}{2}\right\}$.
To sumplify things, define variables a and b,

$$a = (5 - \sqrt{(}5))^{1/2}$$

$$b = (5 + \sqrt{(}5))^{1/2}$$

Using a few useful facts,

$$a^{1/2}(1 + 5^{1/2}) = 2b^{1/2}$$

and

$$b^{1/2}(1-5^{1/2}) = -2a^{1/2}$$

$\underline{\lambda_1 = \frac{5-\sqrt{5}}{2}}$:

$$\mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1 \Rightarrow \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \left(\frac{5-\sqrt{5}}{2}\right) x_1 \\ \left(\frac{5-\sqrt{5}}{2}\right) x_2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \left(\frac{-1+\sqrt{5}}{2}\right) & 1 \\ 1 & \left(\frac{1+\sqrt{5}}{2}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \left(\frac{-1+\sqrt{5}}{2}\right) & 1 \\ 1 & \left(\frac{1+\sqrt{5}}{2}\right) \end{bmatrix} \xrightarrow{\text{Row } 2-\left(\frac{2}{\sqrt{5}-1}\right)\text{Row } 1} \begin{bmatrix} \left(\frac{-1+\sqrt{5}}{2}\right) & 1 \\ 0 & 0 \end{bmatrix}$$

So $\left(\frac{-1+5^{1/2}}{2}\right) x_1 + x_2 = 0 \Rightarrow x_2 = -\left(\frac{-1+5^{1/2}}{2}\right) x_1 = \left(\frac{1-5^{1/2}}{2}\right) x_1$
We have $x_2 = \left(\frac{1-5^{1/2}}{2}\right) x_1$, so when $x = -1$, $x_2 = -\left(\frac{1-5^{1/2}}{2}\right) > 0$.

$$\|\mathbf{x}_1\| = \frac{\left(5-5^{1/2}\right)^{1/2}}{2^{1/2}} = \frac{a^{1/2}}{2^{1/2}}$$

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ -\left(\frac{1-5^{1/2}}{2}\right) \end{bmatrix} \Rightarrow \mathbf{e}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \frac{2^{1/2}}{a^{1/2}} \begin{bmatrix} -1 \\ -\left(\frac{1-5^{1/2}}{2}\right) \end{bmatrix} = \begin{bmatrix} -\frac{2^{1/2}}{a^{1/2}} \\ -\left(\frac{1-5^{1/2}}{2^{1/2}a^{1/2}}\right)\left(\frac{1+5^{1/2}}{1+5^{1/2}}\right) \end{bmatrix} = \begin{bmatrix} -(2/a)^{1/2} \\ (2/b)^{1/2} \end{bmatrix}$$

$\underline{\lambda_2 = \frac{5+\sqrt{5}}{2}}$:

$$\mathbf{A}\mathbf{x}_2 = \lambda_2\mathbf{x}_2 \Rightarrow \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \left(\frac{5+\sqrt{5}}{2}\right) x_1 \\ \left(\frac{5+\sqrt{5}}{2}\right) x_2 \end{bmatrix}$$

$$= \begin{bmatrix} \left(\frac{-1-\sqrt{5}}{2}\right) & 1 \\ 1 & \left(\frac{1-\sqrt{5}}{2}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \left(\frac{-1-\sqrt{5}}{2}\right) & 1 \\ 1 & \left(\frac{1-\sqrt{5}}{2}\right) \end{bmatrix} \xrightarrow{\text{Row } 2+\left(\frac{2}{1+\sqrt{5}}\right)\text{Row } 1} \begin{bmatrix} \left(\frac{-1-\sqrt{5}}{2}\right) & 1 \\ 0 & 0 \end{bmatrix}$$

So $\left(\frac{-1-\sqrt{5}}{2}\right) x_1 + x_2 = 0 \Rightarrow x_2 = \left(\frac{1+\sqrt{5}}{2}\right) x_1$
We have $x_2 = \left(\frac{1+5^{1/2}}{2}\right) x_1$, so when $x = 1$, $x_2 = \left(\frac{1+5^{1/2}}{2}\right) > 0$.

$$\|\mathbf{x}_2\| = \frac{\left(5+5^{1/2}\right)^{1/2}}{2^{1/2}} = \frac{b^{1/2}}{2^{1/2}}$$

$$\mathbf{x}_2 = \begin{bmatrix} 1 \\ \left(\frac{1+5^{1/2}}{2}\right) \end{bmatrix} \Rightarrow \mathbf{e}_1 = \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} = \frac{2^{1/2}}{b^{1/2}} \begin{bmatrix} 1 \\ \left(\frac{1+5^{1/2}}{2}\right) \end{bmatrix} = \begin{bmatrix} \frac{2^{1/2}}{b^{1/2}} \\ \left(\frac{1+5^{1/2}}{2^{1/2}b^{1/2}}\right)\left(\frac{1-5^{1/2}}{1-5^{1/2}}\right) \end{bmatrix} = \begin{bmatrix} (2/b)^{1/2} \\ (2/a)^{1/2} \end{bmatrix}$$

We finally have all of the eigenvalues and eigenvectors,

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} a/2 & 0 \\ 0 & b/2 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} (2/a)^{1/2} & (2/b)^{1/2} \\ -(2/b)^{1/2} & (2/a)^{1/2} \end{bmatrix}$$

Now to find $\mathbf{A}^{1/2}$. Rewrite a few things,

$$\mathbf{A}^{1/2} = 2^{1/2} \begin{bmatrix} \left(b/a^2\right)^{1/2} & \left(a/b^2\right)^{1/2} \\ \left(a/b^2\right)^{1/2} & \left(2^2/a\right)^{1/2} \end{bmatrix}$$

2.21

(See **Result 2A.15**) Using the matrix

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

(a) Calculate $\mathbf{A}'\mathbf{A}$ and obtain the eigenvalues and eigenvectors.

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & -2 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 9 & 1 \\ 1 & 9 \end{bmatrix}$$

$$0 = |\mathbf{A}'\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 9-\lambda & 1 \\ 1 & 9-\lambda \end{vmatrix} = (9-\lambda)^2 - 1 = \lambda^2 - 18\lambda + 81 - 1 = (\lambda - 8)(\lambda - 10)$$

$$\left(\lambda_1^2, \lambda_1^2\right) = (8, 10)$$

$\underline{\lambda_1^2 = 8}$:

$$\mathbf{A}'\mathbf{A}\mathbf{x}_1 = \lambda_1^2\mathbf{x}_1 \Rightarrow \begin{bmatrix} 9 & 1 \\ 1 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8x_1 \\ 8x_2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \xrightarrow{\text{Row 2 - Row 1}} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So $x_1 + x_2 = 0 \Rightarrow x_1 = -x_2$. Pick,

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Rightarrow \mathbf{e}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

$\underline{\lambda_2^2 = 10}$:

$$\mathbf{A}'\mathbf{A}\mathbf{x}_1 = \lambda_1^2\mathbf{x}_1 \Rightarrow \begin{bmatrix} 9 & 1 \\ 1 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10x_1 \\ 10x_2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \xrightarrow{\text{Row 2 + Row 1}} \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So $-x_1 + x_2 = 0 \Rightarrow x_1 = x_2$. Pick,

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \mathbf{e}_2 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

The eigenvectors are,

$$\mathbf{V} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

(b) Calculate $\mathbf{AA'}$ and obtain the eigenvalues and eigenvectors. Check that the nonzero eigenvalues are the same as those in part a.

$$\mathbf{AA'} = \begin{bmatrix} 1 & 1 \\ 2 & -2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 1 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 8 \end{bmatrix}$$

$$0 = |\mathbf{AA'} - \lambda \mathbf{I}| = \begin{vmatrix} 2-\lambda & 0 & 4 \\ 0 & 8-\lambda & 0 \\ 4 & 0 & 8-\lambda \end{vmatrix} = (2-\lambda)(8-\lambda)^2 - 4(4(8-\lambda)) =$$

$$= (8-\lambda)[(2-\lambda)(8-\lambda) - 16] = (8-\lambda)\left(\lambda^2 - 10\lambda\right) = \lambda(8-\lambda)(\lambda - 10)$$

$$\left(\lambda_1^2, \lambda_2^2, \lambda_2^2\right) = (0, 8, 10)$$

Yes, these nonzero eigenvales are the same as those in part a.

$\underline{\lambda_1^2 = 0}$:

$$\mathbf{AA'x}_1 = \lambda_1^2 \mathbf{x}_1 \Rightarrow$$

$$\begin{bmatrix} 2 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{\text{Row 3 - 2Row 1}} \begin{bmatrix} 2 & 0 & 4 \\ 0 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $2x_1 + 4x_3 = 0 \Rightarrow x_1 = -2x_3$ and $8x_2 = 0 \Rightarrow x_2 = 0$. Pick,

$$\mathbf{x}_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \Rightarrow \mathbf{e}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \begin{bmatrix} -2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{bmatrix}$$

$\underline{\lambda_2^2 = 8}$:

$$\mathbf{AA'x}_2 = \lambda_2^2 \mathbf{x}_2 \Rightarrow$$

$$\begin{bmatrix} 2 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8x_1 \\ 8x_2 \\ 8x_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -6 & 0 & 4 \\ 0 & 0 & 0 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $-6x_1 + 4x_3 = 0 \Rightarrow x_1 = 2/3x_3$ and $4x_1 = 0 \Rightarrow x_1 = 0$. Pick,

$$\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \Rightarrow \mathbf{e}_1 = \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$\underline{\lambda_3^2 = 10}$:

$$\mathbf{A}\mathbf{A}'\mathbf{x}_2 = \lambda_2^2\mathbf{x}_2 \Rightarrow$$

$$\begin{bmatrix} 2 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10x_1 \\ 10x_2 \\ 10x_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -8 & 0 & 4 \\ 0 & -2 & 0 \\ 4 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{\text{Row 3 + 1/2 Row 1}} \begin{bmatrix} -8 & 0 & 4 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $-8x_1 + 4x_3 = 0 \Rightarrow x_3 = 2x_1$ and $-2x_2 = 0 \Rightarrow x_3 = 0$. Pick,

$$\mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \Rightarrow \mathbf{e}_1 = \frac{\mathbf{x}_3}{\|\mathbf{x}_3\|} = \begin{bmatrix} 1/\sqrt{5} \\ 0 \\ 2/\sqrt{5} \end{bmatrix}$$

The eigenvectors are,

$$\mathbf{U} = \begin{bmatrix} 0 & 1/\sqrt{5} \\ 1 & 0 \\ 0 & 2/\sqrt{5} \end{bmatrix}$$

(c) Obtain the sigular-value decomposition of $\mathbf{A}$.

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}' = \begin{bmatrix} 0 & 1/\sqrt{5} \\ 1 & 0 \\ 0 & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{10} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -2 \\ 2 & 2 \end{bmatrix} = \mathbf{A}$$

$$\mathbf{A} = \sum_{k=1}^{2} \lambda_k \mathbf{e}_k \mathbf{e}_k' = \sqrt{8} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} + \sqrt{10} \begin{bmatrix} 1/\sqrt{5} \\ 0 \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} =$$

$$= \sqrt{8} \begin{bmatrix} 0 & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 \end{bmatrix} + \sqrt{10} \begin{bmatrix} 1/\sqrt{10} & 1/\sqrt{10} \\ 0 & 0 \\ 2/\sqrt{10} & 2/\sqrt{10} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 2 & -2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 2 & 2 \end{bmatrix} =$$

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

2.22

(See **Result 2A.15**) Using the matrix

$$\mathbf{A} = \begin{bmatrix} 4 & 8 & 8 \\ 3 & 6 & -9 \end{bmatrix}$$

(a) Calculate $\mathbf{AA'}$ and obtain its eigenvalues and eigenvectors.

$$\mathbf{AA'} = \begin{bmatrix} 4 & 8 & 8 \\ 3 & 6 & -9 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 8 & 6 \\ 8 & -9 \end{bmatrix} = \begin{bmatrix} 144 & -12 \\ -12 & 126 \end{bmatrix} = 6 \times \begin{bmatrix} 24 & -2 \\ -2 & 21 \end{bmatrix}$$

$$0 = \left|\mathbf{AA'} - \lambda^2 \mathbf{I}\right| = |\mathbf{AA'} - \gamma \mathbf{I}| = \begin{vmatrix} 24-\gamma & -2 \\ -2 & 21-\gamma \end{vmatrix} =$$

$$= (24-\gamma)(21-\gamma)-4 = 504-45\gamma+\gamma^2-4 = \gamma^2-45\gamma+500 = (\gamma-25)(\gamma-20)$$

The two eigenvalues are:
$\gamma_1 = (1/6)\lambda_1^2 = 20 \Rightarrow \lambda_1 = \sqrt{6 \times 20} = \sqrt{120}$
$\gamma_2 = (1/6)\lambda_2^2 = 25 \Rightarrow \lambda_2 = \sqrt{6 \times 25} = \sqrt{150}$

<u>For $\gamma_1 = 20$:</u>

$$\mathbf{AA'x}_1 = \gamma_1 \mathbf{x} \Rightarrow \begin{bmatrix} 24 & -2 \\ -2 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 20x_1 \\ 20x_2 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \xrightarrow{\text{Row 2 + (1/2) Row 1}} \begin{bmatrix} 4 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So $4x_1 - 2x_2 = 0 \Rightarrow x_2 = 2x_1$

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow \mathbf{u}_1 = \frac{\mathbf{x}_1}{||\mathbf{x}_1||} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

<u>For $\gamma_1 = 25$:</u>

$$\mathbf{AA'x}_2 = \gamma_2 \mathbf{x} \Rightarrow \begin{bmatrix} 24 & -2 \\ -2 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 25x_1 \\ 25x_2 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \xrightarrow{\text{Row 2 - 2 Row 1}} \begin{bmatrix} -1 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So $-x_1 - 2x_2 = 0 \Rightarrow x_1 = -2x_2$

$$\mathbf{x}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \Rightarrow \mathbf{u}_1 = \frac{\mathbf{x}_1}{||\mathbf{x}_1||} = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} \sqrt{120} & 0 \\ 0 & \sqrt{150} \end{bmatrix} \quad \text{and} \quad \mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

(b) Calculate $\mathbf{A}'\mathbf{A}$ and obtain its eigenvalues and eigenvectors. Check that the nonzero eigenvalues are the same as those in part a.

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} 4 & 3 \\ 8 & 6 \\ 8 & -9 \end{bmatrix} \begin{bmatrix} 4 & 8 & 8 \\ 3 & 6 & -9 \end{bmatrix} = \begin{bmatrix} 25 & 50 & 5 \\ 50 & 100 & 10 \\ 5 & 10 & 145 \end{bmatrix} = 5 \times \begin{bmatrix} 5 & 10 & 1 \\ 10 & 20 & 2 \\ 1 & 2 & 29 \end{bmatrix}$$

$$0 = \left|\mathbf{A}'\mathbf{A} - \lambda^2\mathbf{I}\right| = \left|\mathbf{A}'\mathbf{A} - \gamma\mathbf{I}\right| = \begin{vmatrix} 5-\gamma & 10 & 1 \\ 10 & 20-\gamma & 2 \\ 1 & 2 & 29-\gamma \end{vmatrix} =$$

$$= (5-\gamma)[(20-\gamma)(29-\gamma)-4] - 10[10(29-\gamma)-2] + [20-(20-\gamma)] =$$

$$= (29\gamma^2 - 725\gamma + 2900) + (-\gamma^3 + 25\gamma^2 - 100\gamma) - 40 + 5\gamma - 2900 + 100\gamma + 40 =$$

$$= -\gamma(\gamma-24)(\gamma-30)$$

The two nonzero eigenvalues are:
$\gamma_1 = (1/5)\lambda_1^2 = 24 \Rightarrow \lambda_1 = \sqrt{5 \times 24} = \sqrt{120}$
$\gamma_2 = (1/5)\lambda_2^2 = 30 \Rightarrow \lambda_2 = \sqrt{5 \times 30} = \sqrt{150}$
These are the same two eigenvalues as in part (a).

<u>For $\gamma_1 = 24$:</u>

$$\mathbf{A}'\mathbf{A}\mathbf{x}_1 = \gamma_1\mathbf{x} \Rightarrow \begin{bmatrix} 5 & 10 & 1 \\ 10 & 20 & 2 \\ 1 & 2 & 29 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 24x_1 \\ 24x_2 \\ 24x_3 \end{bmatrix} \Rightarrow \begin{bmatrix} -19 & 10 & 1 \\ 10 & -4 & 2 \\ 1 & 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -19 & 10 & 1 \\ 10 & -4 & 2 \\ 1 & 2 & 5 \end{bmatrix} \xrightarrow{\text{Swap Row 1 with Row 3}} \begin{bmatrix} 1 & 2 & 5 \\ 10 & -4 & 2 \\ -19 & 10 & 1 \end{bmatrix} \xrightarrow{\text{Row 3 + 19 Row 1}} \begin{bmatrix} 1 & 2 & 5 \\ 10 & -4 & 2 \\ 0 & 48 & 96 \end{bmatrix}$$

$$\xrightarrow{\text{Row 2 - 10 Row 1}} \begin{bmatrix} 1 & 2 & 5 \\ 0 & -24 & -48 \\ 0 & 48 & 96 \end{bmatrix} \xrightarrow{\text{Simplify rows}} \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{\text{Row 3 - Row 2}}$$

$$\begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Row 1 - 2 Row 2}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

So $x_1 + x_3 = 0 \Rightarrow x_1 = -x_3$ and $x_2 + 2x_3 = 0 \Rightarrow x_2 = -2x_3$.

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \Rightarrow \mathbf{v}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \begin{bmatrix} -1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

<u>For $\gamma_2 = 30$:</u>

$$\mathbf{A}'\mathbf{A}\mathbf{x}_1 = \gamma_1\mathbf{x} \Rightarrow \begin{bmatrix} 5 & 10 & 1 \\ 10 & 20 & 2 \\ 1 & 2 & 29 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 30x_1 \\ 30x_2 \\ 30x_3 \end{bmatrix} \Rightarrow \begin{bmatrix} -25 & 10 & 1 \\ 10 & -10 & 2 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix}-25 & 10 & 1\\ 10 & -10 & 2\\ 1 & 2 & -1\end{bmatrix} \xrightarrow{\text{Swap Row 1 with Row 3}} \begin{bmatrix}1 & 2 & -1\\ 10 & -10 & 2\\ -25 & 10 & 1\end{bmatrix} \xrightarrow{\text{Row 3 + 25 Row 1}} \begin{bmatrix}1 & 2 & -1\\ 10 & -10 & 2\\ 0 & 60 & -24\end{bmatrix}$$

$$\xrightarrow{\text{Row 2 - 10 Row 1}} \begin{bmatrix}1 & 2 & -1\\ 0 & -30 & 12\\ 0 & 48 & 96\end{bmatrix} \xrightarrow{\text{Simplify rows}} \begin{bmatrix}1 & 2 & -1\\ 0 & -5 & 2\\ 0 & 5 & -2\end{bmatrix} \xrightarrow{\text{Row 3 + Row 2}}$$

$$\begin{bmatrix}1 & 2 & -1\\ 0 & -5 & 2\\ 0 & 0 & 0\end{bmatrix} \xrightarrow{\text{Row 1 - (2/5) Row 2}} \begin{bmatrix}1 & 0 & -1/5\\ 0 & -5 & 2\\ 0 & 0 & 0\end{bmatrix}$$

So $x_1 - (1/5)x_3 = 0 \Rightarrow x_3 = 5x_1$ and $-5x_2 + 2x_3 = 0 \Rightarrow x_2 = (2/5)x_3$.

$$\mathbf{x}_2 = \begin{bmatrix}1\\2\\5\end{bmatrix} \Rightarrow \mathbf{v}_2 = \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} = \begin{bmatrix}1/\sqrt{30}\\ 2/\sqrt{30}\\ 5/\sqrt{30}\end{bmatrix}$$

$$\Lambda = \begin{bmatrix}\sqrt{120} & 0\\ 0 & \sqrt{150}\end{bmatrix} \quad \text{and} \quad \mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} = \begin{bmatrix}-1/\sqrt{6} & 1/\sqrt{30}\\ -2/\sqrt{6} & 2/\sqrt{30}\\ 2/\sqrt{6} & 5/\sqrt{30}\end{bmatrix}$$

(c) Obtain the Singular value decomposition of $\mathbf{A}$.

Need to multiply SVD by -1, so we get the right result. Back when computing the eigenvalues notice the negative in front.

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}' = (-1)\begin{bmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix}\begin{bmatrix}\sqrt{120} & 0\\ 0 & \sqrt{150}\end{bmatrix}\begin{bmatrix}-1/\sqrt{6} & -2/\sqrt{6} & 2/\sqrt{6}\\ 1/\sqrt{30} & 2/\sqrt{30} & 5/\sqrt{30}\end{bmatrix} =$$

$$= \left(\frac{1}{\sqrt{5}}\right)\begin{bmatrix}-1 & 2\\ -2 & -1\end{bmatrix}\left(\sqrt{5}\sqrt{6}\right)\begin{bmatrix}\sqrt{4} & 0\\ 0 & \sqrt{5}\end{bmatrix}\left(\frac{1}{\sqrt{6}}\right)\begin{bmatrix}-1 & -2 & 1\\ 1/\sqrt{5} & 2/\sqrt{5} & 5/\sqrt{5}\end{bmatrix} =$$

$$= \begin{bmatrix}-1 & 2\\ -2 & -1\end{bmatrix}\begin{bmatrix}-2 & -4 & 2\\ 1 & 2 & 5\end{bmatrix} = \begin{bmatrix}4 & 8 & 8\\ 3 & 6 & -9\end{bmatrix} = \mathbf{A}$$

2.23

Verify the relationships $\mathbf{V}^{1/2}\boldsymbol{\rho}\mathbf{V}^{1/2} = \boldsymbol{\Sigma}$ and $\boldsymbol{\rho} = \left(\mathbf{V}^{1/2}\right)^{-1}\boldsymbol{\Sigma}\left(\mathbf{V}^{1/2}\right)^{-1}$, where $\boldsymbol{\Sigma}$ is the $p \times p$ population covariance matrix [**Equation (2-32)**], $\boldsymbol{\rho}$ is the $p \times p$ population correlation matrix [**Equation (2-34)**], and $\mathbf{V}^{1/2}$ is the population standard deviation matrix [**Equation (2-35)**].

$$\mathbf{V}^{1/2}\boldsymbol{\rho}\mathbf{V}^{1/2} = \begin{bmatrix}\sqrt{\sigma_{11}} & & 0\\ & \ddots & \\ 0 & & \sqrt{\sigma_{pp}}\end{bmatrix}\begin{bmatrix}\frac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{11}}} & \cdots & \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}}\\ \vdots & \ddots & \vdots\\ \frac{\sigma_{p1}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{11}}} & \cdots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{pp}}}\end{bmatrix}\begin{bmatrix}\sqrt{\sigma_{11}} & & 0\\ & \ddots & \\ 0 & & \sqrt{\sigma_{pp}}\end{bmatrix} =$$

$$= \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}} & \cdots & \frac{\sigma_{1p}}{\sqrt{\sigma_{pp}}} \\ \vdots & \ddots & \vdots \\ \frac{\sigma_{p1}}{\sqrt{\sigma_{11}}} & \cdots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}}} \end{bmatrix} \begin{bmatrix} \sqrt{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\sigma_{pp}} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix} = \mathbf{\Sigma}$$

$$\left(\mathbf{V}^{1/2}\right)^{-1} \mathbf{\Sigma} \left(\mathbf{V}^{1/2}\right)^{-1} =$$

$$= \left(\begin{bmatrix} \sqrt{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\sigma_{pp}} \end{bmatrix} \right)^{-1} \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix} \left(\begin{bmatrix} \sqrt{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\sigma_{pp}} \end{bmatrix} \right)^{-1} =$$

$$= \begin{bmatrix} 1/\sqrt{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & 1/\sqrt{\sigma_{pp}} \end{bmatrix} \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix} \begin{bmatrix} 1/\sqrt{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & 1/\sqrt{\sigma_{pp}} \end{bmatrix} =$$

$$= \begin{bmatrix} 1/\sqrt{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & 1/\sqrt{\sigma_{pp}} \end{bmatrix} \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}} & \cdots & \frac{\sigma_{1p}}{\sqrt{\sigma_{pp}}} \\ \vdots & \ddots & \vdots \\ \frac{\sigma_{p1}}{\sqrt{\sigma_{11}}} & \cdots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}}} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{11}}} & \cdots & \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} \\ \vdots & \ddots & \vdots \\ \frac{\sigma_{p1}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{11}}} & \cdots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{pp}}} \end{bmatrix} = \boldsymbol{\rho}$$

2.24

Let $\mathbf{X}$ have covariance matrix

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

Find

(a) $\mathbf{\Sigma}^{-1}$

Using what's on page 59, the inverse of a diagonal matrix is the reciprocal of the elements.

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/9 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(b) The eigenvalues and eigenvectors of $\mathbf{\Sigma}$.

$$0 = |\mathbf{\Sigma} - \lambda \mathbf{I}| = \begin{bmatrix} 4-\lambda & 0 & 0 \\ 0 & 9-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{bmatrix} = (4-\lambda)(9-\lambda)(1-\lambda)$$

The eigenvalues are simply the diagonal elements of $\mathbf{\Sigma}$, $(\lambda_1, \lambda_2, \lambda_3) = (1, 4, 9)$.

$\underline{\lambda_1 = 1}$:

$$\mathbf{\Sigma x}_1 = \lambda_1 \mathbf{x}_1 \Rightarrow \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $x_1 = x_2 = 0$ and x_3 is free.

$$\mathbf{x}_1 = \mathbf{e}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$\underline{\lambda_2 = 4}$:

$$\mathbf{\Sigma x}_2 = \lambda_2 \mathbf{x}_2 \Rightarrow \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4x_1 \\ 4x_2 \\ 4x_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $x_2 = x_3 = 0$ and x_1 is free.

$$\mathbf{x}_2 = \mathbf{e}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$\underline{\lambda_3 = 9}$:

$$\mathbf{\Sigma x}_3 = \lambda_3 \mathbf{x}_3 \Rightarrow \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9x_1 \\ 9x_2 \\ 9x_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $x_1 = x_3 = 0$ and x_2 is free.

$$\mathbf{x}_3 = \mathbf{e}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Putting it all together,

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

The eigenvalues are the elements in the diagonal matrix and the eigenvectors are the identity matrix.

(c) The eigenvalues and eigenvectors of $\mathbf{\Sigma}^{-1}$.

$$0 = |\mathbf{\Sigma} - \lambda\mathbf{I}| = \begin{bmatrix} 1/4-\lambda & 0 & 0 \\ 0 & 1/9-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{bmatrix} = (1/4-\lambda)\,(1/9-\lambda)\,(1/9-\lambda)$$

The eigenvalues are simply the diagonal elements of $\mathbf{\Sigma}$, $(\lambda_1, \lambda_2, \lambda_3) = (1/9, 1/4, 1)$.

$\underline{\lambda_1 = 1/9:}$

$$\mathbf{\Sigma}\mathbf{x}_1 = \lambda_1\mathbf{x}_1 \Rightarrow \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/9 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} (1/9)x_1 \\ (1/9)x_2 \\ (1/9)x_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 5/36 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 8/9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $x_1 = x_3 = 0$ and x_2 is free.

$$\mathbf{x}_1 = \mathbf{e}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$\underline{\lambda_2 = 1/4:}$

$$\mathbf{\Sigma}\mathbf{x}_2 = \lambda_1\mathbf{x}_2 \Rightarrow \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/9 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} (1/4)x_1 \\ (1/4)x_2 \\ (1/4)x_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & -5/36 & 0 \\ 0 & 0 & 3/4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $x_2 = x_3 = 0$ and x_1 is free.

$$\mathbf{x}_2 = \mathbf{e}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$\underline{\lambda_3 = 1:}$

$$\mathbf{\Sigma}\mathbf{x}_3 = \lambda_3\mathbf{x}_3 \Rightarrow \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/9 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -3/4 & 0 & 0 \\ 0 & -8/9 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $x_1 = x_2 = 0$ and x_3 is free.

$$\mathbf{x}_3 = \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Putting it all together,

$$\mathbf{\Lambda} = \begin{bmatrix} 1/9 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The eigenvalues are the elements in the diagonal matrix and the eigenvectors are the identity matrix.

2.25

Let $\mathbf{X}$ have covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} 25 & -2 & 4 \\ -2 & 4 & 1 \\ 4 & 1 & 9 \end{bmatrix}$$

(a) Determine $\boldsymbol{\rho}$ and $\mathbf{V}^{1/2}$.

$$\boldsymbol{\rho} = \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{11}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \frac{\sigma_{13}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{33}}} \\ \frac{\sigma_{21}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{11}}} & \frac{\sigma_{22}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{22}}} & \frac{\sigma_{23}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{33}}} \\ \frac{\sigma_{31}}{\sqrt{\sigma_{33}}\sqrt{\sigma_{11}}} & \frac{\sigma_{32}}{\sqrt{\sigma_{33}}\sqrt{\sigma_{22}}} & \frac{\sigma_{33}}{\sqrt{\sigma_{33}}\sqrt{\sigma_{33}}} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{25}{\sqrt{25}\sqrt{25}} & \frac{-2}{\sqrt{25}\sqrt{4}} & \frac{4}{\sqrt{25}\sqrt{9}} \\ \frac{-2}{\sqrt{4}\sqrt{25}} & \frac{4}{\sqrt{4}\sqrt{4}} & \frac{1}{\sqrt{4}\sqrt{9}} \\ \frac{4}{\sqrt{9}\sqrt{25}} & \frac{1}{\sqrt{9}\sqrt{4}} & \frac{9}{\sqrt{9}\sqrt{9}} \end{bmatrix} = \begin{bmatrix} 1 & -1/5 & 4/15 \\ -1/5 & 1 & 1/6 \\ 4/15 & 1/6 & 1 \end{bmatrix}$$

(b) Multiply your matrices to check the relation $\mathbf{V}^{1/2}\boldsymbol{\rho}\mathbf{V}^{1/2} = \mathbf{\Sigma}$.

$$\mathbf{V}^{1/2}\boldsymbol{\rho}\mathbf{V}^{1/2} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1/5 & 4/15 \\ -1/5 & 1 & 1/6 \\ 4/15 & 1/6 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} =$$

$$= \begin{bmatrix} 5 & -1 & 4/3 \\ -2/5 & 2 & 1/3 \\ 4/5 & 1/2 & 3 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 25 & -2 & 4 \\ -2 & 4 & 1 \\ 4 & 1 & 9 \end{bmatrix}$$

2.26

Use $\mathbf{\Sigma}$ as given in Exercise 2.25.

(a) Find ρ_{13}.

We can pick ρ_{13} off the result from Exercise 2.25 (a),

$$\rho_{13} = 4/15$$

(b) Find the correlation between X_1 and $\frac{1}{2}X_2 + \frac{1}{2}X_3$.

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \hdashline \mathbf{X}^{(2)} \end{bmatrix}$$

$$\mathbf{X}^{(1)} = \begin{bmatrix} X_1 \end{bmatrix} \quad \text{and} \quad \mathbf{X}^{(2)} = \begin{bmatrix} X_2 \\ X_3 \end{bmatrix}$$

$$Y = \frac{1}{2}X_1 + \frac{1}{2}X_2 = \begin{bmatrix} 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \mathbf{c}'\mathbf{X}^{(2)}$$

So $\mathbf{c} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$.

$$\mathbf{\Sigma} = \left[\begin{array}{c:cc} 25 & -2 & 4 \\ \hdashline -2 & 4 & 1 \\ 4 & 1 & 9 \end{array}\right] = \left[\begin{array}{c:c} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \hdashline \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{array}\right] =$$

$$\mathrm{Cov}\left(\mathbf{X}^{(1)}, Y\right) = \mathrm{Cov}\left(\mathbf{X}^{(1)}, \mathbf{c}'\mathbf{X}^{(2)}\right) =$$

$$= E\left[\left(\mathbf{X}^{(1)} - E\left[\mathbf{X}^{(1)}\right]\right)\left(\mathbf{c}'\mathbf{X}^{(2)} - E\left[\mathbf{c}'\mathbf{X}^{(2)}\right]\right)'\right] =$$

$$= E\left[\left(\mathbf{X}^{(1)} - E\left[\mathbf{X}^{(1)}\right]\right)\left(\mathbf{c}'\mathbf{X}^{(2)} - \mathbf{c}'E\left[\mathbf{X}^{(2)}\right]\right)'\right] =$$

$$= E\left[\left(\mathbf{X}^{(1)} - E\left[\mathbf{X}^{(1)}\right]\right)\left\{\mathbf{c}'\left(\mathbf{X}^{(2)} - E\left[\mathbf{X}^{(2)}\right]\right)\right\}'\right] =$$

$$= E\left[\left(\mathbf{X}^{(1)} - E\left[\mathbf{X}^{(1)}\right]\right)\left\{\left(\mathbf{X}^{(2)} - E\left[\mathbf{X}^{(2)}\right]\right)\right\}'\mathbf{c}\right] =$$

$$= E\left[\left(\mathbf{X}^{(1)} - E\left[\mathbf{X}^{(1)}\right]\right)\left\{\left(\mathbf{X}^{(2)} - E\left[\mathbf{X}^{(2)}\right]\right)\right\}'\right]\mathbf{c} =$$

$$= \mathrm{Cov}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}\right)\mathbf{c} = \begin{bmatrix} -2 & 4 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = 1$$

2.27

Derive expressions for the mean and variances of the following linear combinations in terms of the means and covariances of the random variables X_1, X_2, and X_3.

(a) $X_1 - 2X_2$

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$$

$$E\left[\mathbf{c}'\mathbf{X}\right] = E\left[\begin{bmatrix} 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] = \begin{bmatrix} 1 & -2 & 0 \end{bmatrix} E\left[\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] =$$

$$= \begin{bmatrix} 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} E\left[X_1\right] \\ E\left[X_2\right] \\ E\left[X_3\right] \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = 1 \times \mu_1 - 2 \times \mu_2 + 0 \times \mu_3 = \mu_1 - 2\mu_2$$

$$\mathrm{V}\left(\mathbf{c}'\mathbf{X}\right) = \mathrm{Cov}\left(\mathbf{c}'\mathbf{X}, \mathbf{c}'\mathbf{X}\right) = E\left[\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)'\right] =$$

$$= \mathbf{c}' E\left[\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)'\right]\mathbf{c} = \mathbf{c}'\mathbf{\Sigma}\mathbf{c} =$$

$$\begin{bmatrix} 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} (\sigma_{11} + \sigma_{21}) & (\sigma_{12} + \sigma_{22}) & (\sigma_{13} + \sigma_{23}) \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} =$$

$$= 1 \times (\sigma_{11} + \sigma_{21}) - 2 \times (\sigma_{12} + \sigma_{22}) + 0 \times (\sigma_{13} + \sigma_{23}) = (\sigma_{11} - 2\sigma_{12}) + (\sigma_{21} - 2\sigma_{22})$$

(b) $-X_1 + 3X_2$

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}$$

$$E\left[\mathbf{c}'\mathbf{X}\right] = E\left[\begin{bmatrix} -1 & 3 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] = \begin{bmatrix} -1 & 3 & 0 \end{bmatrix} E\left[\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] =$$

$$= \begin{bmatrix} -1 & 3 & 0 \end{bmatrix} \begin{bmatrix} E\left[X_1\right] \\ E\left[X_2\right] \\ E\left[X_3\right] \end{bmatrix} = \begin{bmatrix} -1 & 3 & 0 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = -1 \times \mu_1 3 \times \mu_2 + 0 \times \mu_3 = -\mu_1 + 3\mu_2$$

$$\mathrm{V}\left(\mathbf{c}'\mathbf{X}\right) = \mathrm{Cov}\left(\mathbf{c}'\mathbf{X}, \mathbf{c}'\mathbf{X}\right) = E\left[\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)'\right] =$$

$$= \mathbf{c}' E\left[\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)'\right]\mathbf{c} = \mathbf{c}'\mathbf{\Sigma}\mathbf{c} =$$

$$\begin{bmatrix} -1 & 3 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} (-\sigma_{11} + 3\sigma_{21}) & (-\sigma_{12} + 3\sigma_{22}) & (-\sigma_{13} + 3\sigma_{23}) \end{bmatrix} \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} =$$

$$= -1 \times (-\sigma_{11} + 3\sigma_{21}) + 3 \times (-\sigma_{12} + 3\sigma_{22}) + 0 \times (-\sigma_{13} + 3\sigma_{23}) = \sigma_{11} + 3(\sigma_{21} - \sigma_{12}) + 6\sigma_{22}$$

(c) $X_1 + X_2 + X_3$

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$E\left[\mathbf{c}'\mathbf{X}\right] = E\left[\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} E\left[\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] =$$

$$= \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} E\left[X_1\right] \\ E\left[X_2\right] \\ E\left[X_3\right] \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = 1\times\mu_1 + 1\times\mu_2 + 1\times\mu_3 = \mu_1 + \mu_2 + \mu_3$$

$$\mathrm{V}\left(\mathbf{c}'\mathbf{X}\right) = \mathrm{Cov}\left(\mathbf{c}'\mathbf{X}, \mathbf{c}'\mathbf{X}\right) = E\left[\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)'\right] =$$

$$= \mathbf{c}'E\left[\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)'\right]\mathbf{c} = \mathbf{c}'\mathbf{\Sigma}\mathbf{c} =$$

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} (\sigma_{11} + \sigma_{21} + \sigma_{31}) & (\sigma_{12} + \sigma_{22} + \sigma_{32}) & (\sigma_{13} + \sigma_{23} + \sigma_{33}) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} =$$

$$= 1 \times (\sigma_{11} + \sigma_{21} + \sigma_{31}) + 1 \times (\sigma_{12} + \sigma_{22} + \sigma_{32}) + 1 \times (\sigma_{13} + \sigma_{23} + \sigma_{33}) =$$

$$= \sum_{i=1}^{3} \sigma_{i1} + \sum_{i=1}^{3} \sigma_{i3} + \sum_{i=1}^{3} \sigma_{i3} = \sum_{i=1}^{3} \sum_{j=1}^{3} \sigma_{ij}$$

(d) $X_1 + 2X_2 - X_3$

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$E\left[\mathbf{c}'\mathbf{X}\right] = E\left[\begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} E\left[\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] =$$

$$= \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} E\left[X_1\right] \\ E\left[X_2\right] \\ E\left[X_3\right] \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = 1\times\mu_1 + 2\times\mu_2 - 1\times\mu_3 = \mu_1 + 2\mu_2 - \mu_3$$

$$\mathrm{V}\left(\mathbf{c}'\mathbf{X}\right) = \mathrm{Cov}\left(\mathbf{c}'\mathbf{X}, \mathbf{c}'\mathbf{X}\right) = E\left[\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)'\right] =$$

$$= \mathbf{c}'E\left[\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)'\right]\mathbf{c} = \mathbf{c}'\mathbf{\Sigma}\mathbf{c} =$$

$$\begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} =$$

$$= \begin{bmatrix}(\sigma_{11}+2\sigma_{21}-\sigma_{31}) & (\sigma_{12}+2\sigma_{22}-\sigma_{32}) & (\sigma_{13}+2\sigma_{23}-\sigma_{33})\end{bmatrix}\begin{bmatrix}1\\2\\-1\end{bmatrix} =$$

$$= 1\times(\sigma_{11}+2\sigma_{21}-\sigma_{31})+2\times(\sigma_{12}+2\sigma_{22}-\sigma_{32})-1\times(\sigma_{13}+2\sigma_{23}-\sigma_{33}) =$$

$$= \sigma_{11}+2\sigma_{12}-\sigma_{13}+2(\sigma_{21}+\sigma_{22}-\sigma_{23})-(\sigma_{31}+\sigma_{32}-\sigma_{33})$$

(e) $3X_1 - 4X_2$ if X_1 and X_2 are independent random variables.

First off, if $X_1 \perp\!\!\!\perp X_2$, then $\sigma_{12} = \sigma_{21} = 0$.

$$\mathbf{X} = \begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix}3\\-4\\0\end{bmatrix}$$

$$E\left[\mathbf{c}'\mathbf{X}\right] = E\left[\begin{bmatrix}3 & -4 & 0\end{bmatrix}\begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix}\right] = \begin{bmatrix}3 & -4 & 0\end{bmatrix} E\left[\begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix}\right] =$$

$$= \begin{bmatrix}3 & -4 & 0\end{bmatrix}\begin{bmatrix}E[X_1]\\E[X_2]\\E[X_3]\end{bmatrix} = \begin{bmatrix}3 & -4 & 0\end{bmatrix}\begin{bmatrix}\mu_1\\\mu_2\\\mu_3\end{bmatrix} = 3\times\mu_1-4\times\mu_2+0\times\mu_3 = 3\mu_1-4\mu_2$$

$$\mathrm{V}\left(\mathbf{c}'\mathbf{X}\right) = \mathrm{Cov}\left(\mathbf{c}'\mathbf{X}, \mathbf{c}'\mathbf{X}\right) = E\left[\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)\left(\mathbf{c}'\mathbf{X} - E\left[\mathbf{c}'\mathbf{X}\right]\right)'\right] =$$

$$= \mathbf{c}'E\left[\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)\left(\mathbf{X} - E\left[\mathbf{X}\right]\right)'\right]\mathbf{c} = \mathbf{c}'\boldsymbol{\Sigma}\mathbf{c} =$$

$$\begin{bmatrix}3 & -4 & 0\end{bmatrix}\begin{bmatrix}\sigma_{11} & \sigma_{12} & \sigma_{13}\\\sigma_{21} & \sigma_{22} & \sigma_{23}\\\sigma_{31} & \sigma_{32} & \sigma_{33}\end{bmatrix}\begin{bmatrix}3\\-4\\0\end{bmatrix} =$$

$$= \begin{bmatrix}(3\sigma_{11}-4\sigma_{21}) & (3\sigma_{12}-4\sigma_{22}) & (3\sigma_{13}-4\sigma_{23})\end{bmatrix}\begin{bmatrix}3\\-4\\0\end{bmatrix} =$$

$$= 3\times(3\sigma_{11}-4\sigma_{21}) - 4\times(3\sigma_{12}-4\sigma_{22}) + 0\times(3\sigma_{13}-4\sigma_{23}) =$$

$$= 9\sigma_{11}-12\sigma_{21}-12\sigma_{12}-16\sigma_{22} = 9\sigma_{11}-0-0-16\sigma_{22} = 9\sigma_{11}-16\sigma_{22}$$

2.28

Show that

$$\mathrm{Cov}\left(c_{11}X_1 + c_{12}X_2 + \cdots + c_{1p}X_p, c_{21}X_1 + c_{22}X_2 + \cdots + c_{2p}X_p\right) = \mathbf{c}_1'\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{c}_2$$

where $\mathbf{c}_1' = \begin{bmatrix}c_{11}, c_{12}, \ldots, c_{1p}\end{bmatrix}$ and $\mathbf{c}_2' = \begin{bmatrix}c_{21}, c_{22}, \ldots, c_{2p}\end{bmatrix}$. This verifies the off-diagonal elements $\mathbf{C}\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{C}'$ in (2-45) or diagonal elements if $\mathbf{c}_1 = \mathbf{c}_2$.
Hint: By **(2-45)**, $Z_1 - E(Z_1) = c_{11}(X_1-\mu_1) + \cdots + c_{1p}(X_p-\mu_p)$ and $Z_2 - E(Z_2) = c_{21}(X_1-\mu_1) + \cdots + c_{2p}(X_p-\mu_p)$. So $\mathrm{Cov}(Z_1, Z_2) = E[(Z_1 -$

$E(Z_1))(Z_2 - E(Z_2))] = E[(c_{11}(X_1 - \mu_1) + \cdots + c_{1p}(X_p - \mu_p))(c_{21}(X_1 - \mu_1) + \cdots + c_{2p}(X_p - \mu_p))]$. The product

$$(c_{11}(X_1 - \mu_1) + \cdots + c_{1p}(X_p - \mu_p))(c_{21}(X_1 - \mu_1) + \cdots + c_{2p}(X_p - \mu_p)) =$$

$$= \left(\sum_{\ell=1}^{p} c_{1\ell}(X_\ell - \mu_\ell)\right)\left(\sum_{m=1}^{p} c_{2m}(X_m - \mu_m)\right)$$

$$= \sum_{\ell=1}^{p}\sum_{\ell=1}^{p} c_{1\ell}c_{2m}(X_\ell - \mu_\ell)(X_m - \mu_m)$$

has expected value

$$= \sum_{\ell=1}^{p}\sum_{\ell=1}^{p} c_{1\ell}c_{2m}\sigma_{\ell m} = [c_{11}, \ldots, c_{1p}]\,\mathbf{\Sigma}\,[c_{21}, \ldots, c_{2p}]'.$$

Verify the last step by the definition of matrix multiplication. The same steps should hold for all elements.

Ignoring the hint, using the definition,

$$Z_1 = \mathbf{c}_1'\mathbf{X} = \sum_{i=1}^{p} c_{1i}X_i$$

$$Z_2 = \mathbf{c}_2'\mathbf{X} = \sum_{i=1}^{p} c_{2i}X_i$$

$$\begin{aligned}
&\mathrm{Cov}(Z_1, Z_2) = \\
&= E\left[(Z_1 - E[Z_1])(Z_2 - E[Z_2])'\right] = \\
&= E\left[(\mathbf{c}_1'\mathbf{X} - E[\mathbf{c}_1'\mathbf{X}])(\mathbf{c}_2'\mathbf{X} - E[\mathbf{c}_2'\mathbf{X}])'\right] = \\
&= E\left[\mathbf{c}_1'(\mathbf{X} - E[\mathbf{X}])(\mathbf{c}_2'(\mathbf{X} - E[\mathbf{X}]))'\right] = \\
&= E\left[\mathbf{c}_1'(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])'(\mathbf{c}_2')'\right] = \\
&= E\left[\mathbf{c}_1'(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])'\mathbf{c}_2\right] = \\
&= \mathbf{c}_1'E\left[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])'\right]\mathbf{c}_2 = \\
&= \mathbf{c}_1'\mathrm{Cov}(\mathbf{X})\,\mathbf{c}_2 = \\
&= \mathbf{c}_1'\mathbf{\Sigma}\mathbf{c}_2
\end{aligned}$$

Another way, using the hint and multiplying everything out,

$$\mathrm{Cov}(Z_1, Z_2) = E\left[(Z_1 - E[Z_2])(Z_1 - E[Z_2])'\right] =$$

$$E\left[\left\{\sum_{\ell=1}^{p} c_{1\ell}\left(X_\ell = \mu_\ell\right)\right\}\left\{\sum_{m=1}^{p} c_{2m}\left(X_m = \mu_m\right)\right\}\right] =$$

$$= E[c_{11}c_{21}(X_1-\mu_1)(X_1-\mu_1)+c_{11}c_{22}(X_1-\mu_1)(X_2-\mu_2)+\cdots+c_{11}c_{2p}(X_1-\mu_1)(X_p-\mu_p)+$$
$$+c_{12}c_{21}(X_2-\mu_2)(X_1-\mu_1)+c_{12}c_{22}(X_2-\mu_2)(X_2-\mu_2)+\cdots+c_{12}c_{2p}(X_2-\mu_2)(X_p-\mu_p)+$$
$$\cdots$$
$$+c_{1p}c_{21}(X_p-\mu_p)(X_1-\mu_1)+c_{1p}c_{22}(X_p-\mu_p)(X_2-\mu_2)+\cdots+c_{1p}c_{2p}(X_p-\mu_p)(X_p-\mu_p)] =$$

$$E\left[\sum_{\ell=1}^{p}\sum_{m=1}^{p} c_{1\ell}c_{2m}\left(X_\ell-\mu_\ell\right)\left(X_m-\mu_m\right)\right] = \sum_{\ell=1}^{p}\sum_{m=1}^{p} c_{1\ell}c_{2m}E\left[\left(X_\ell-\mu_\ell\right)\left(X_m-\mu_m\right)\right] =$$

$$= \sum_{\ell=1}^{p}\sum_{m=1}^{p} c_{1\ell}c_{2m}\sigma_{\ell m}$$

$$\mathbf{c}'\boldsymbol{\Sigma}\mathbf{c} = \begin{bmatrix} c_{11}, & \cdots, & c_{1p}\end{bmatrix}\begin{bmatrix}\sigma_{11} & \cdots & \sigma_{1p}\\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp}\end{bmatrix}\begin{bmatrix} c_{21}\\ \vdots \\ c_{2p}\end{bmatrix} =$$

$$= \begin{bmatrix}\sum_{\ell=1}^{p} c_{1\ell}\sigma_{\ell 1}, & \cdots, & \sum_{\ell=1}^{p} c_{1\ell}\sigma_{\ell p}\end{bmatrix}\begin{bmatrix} c_{21}\\ \vdots \\ c_{2p}\end{bmatrix} =$$

$$= c_{21}\sum_{\ell=1}^{p} c_{1\ell}\sigma_{\ell 1} + \cdots + c_{2p}\sum_{\ell=1}^{p} c_{1\ell}\sigma_{\ell p} = \sum_{\ell=1}^{p}\sum_{m=1}^{p} c_{1\ell}c_{2m}\sigma_{\ell m}$$

Now we finally have,

$$\operatorname{Cov}\left(Z_1, Z_2\right) = \sum_{\ell=1}^{p}\sum_{m=1}^{p} c_{1\ell}c_{2m}\sigma_{\ell m} = \mathbf{c}'\boldsymbol{\Sigma}\mathbf{c}$$

2.29

Consider the arbitrary random vector $\mathbf{X}' = \begin{bmatrix} X_1, & X_2, & X_3, & X_4, & X_5\end{bmatrix}$ with mean vector $\boldsymbol{\mu}' = \begin{bmatrix}\mu_1, & \mu_2, & \mu_3, & \mu_4, & \mu_5\end{bmatrix}$. Partition $\mathbf{X}$ into

$$\mathbf{X} = \begin{bmatrix}\mathbf{X}^{(1)}\\ \hdashline \mathbf{X}^{(2)}\end{bmatrix}$$

where

$$\mathbf{X}^{(1)} = \begin{bmatrix} X_1\\ X_2\end{bmatrix} \quad \text{and} \quad \mathbf{X}^{(2)} = \begin{bmatrix} X_3\\ X_4\\ X_5\end{bmatrix}$$

Let $\mathbf{\Sigma}$ be the covariance matrix of $\mathbf{X}$ with general element σ_{ik}. Partition $\mathbf{\Sigma}$ into the covariance matrices of $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ and the covariance matrix of an element of $\mathbf{X}^{(1)}$ and an element of $\mathbf{X}^{(2)}$.

$$\mathbf{\Sigma} = \left[\begin{array}{cc:ccc} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \hdashline \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} & \sigma_{35} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} & \sigma_{45} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_{55} \end{array}\right] = \left[\begin{array}{c:c} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \hdashline \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{array}\right]$$

$$\mathbf{\Sigma}_{11} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}, \quad \mathbf{\Sigma}_{12} = \begin{bmatrix} \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{23} & \sigma_{24} & \sigma_{25} \end{bmatrix}, \quad \mathbf{\Sigma}_{21} = \begin{bmatrix} \sigma_{31} & \sigma_{32} \\ \sigma_{41} & \sigma_{42} \\ \sigma_{51} & \sigma_{52} \end{bmatrix}, \quad \mathbf{\Sigma}_{22} \begin{bmatrix} \sigma_{33} & \sigma_{34} & \sigma_{35} \\ \sigma_{43} & \sigma_{44} & \sigma_{45} \\ \sigma_{53} & \sigma_{54} & \sigma_{55} \end{bmatrix}$$

where $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{21}'$.

2.30

You are given the random vector $\mathbf{X}' = \begin{bmatrix} X_1, & X_2, & X_3, & X_4 \end{bmatrix}$ with mean vector $\boldsymbol{\mu}_{\mathbf{x}}' = \begin{bmatrix} 4, & 3, & 2, & 1 \end{bmatrix}$ and the variance-covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 9 & -2 \\ 2 & 0 & -2 & 4 \end{bmatrix}$$

Partition $\mathbf{X}$ as

$$\mathbf{X} = \left[\begin{array}{c} X_1 \\ X_2 \\ \hdashline X_3 \\ X_4 \end{array}\right] = \left[\begin{array}{c} \mathbf{X}^{(1)} \\ \hdashline \mathbf{X}^{(2)} \end{array}\right]$$

Let

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & -2 \\ 2 & -1 \end{bmatrix}$$

and consider the linear combinations $\mathbf{AX}^{(1)}$ and $\mathbf{AX}^{(2)}$. Find

(a) $E\left(\mathbf{X^{(1)}}\right)$

$$E\left[\mathbf{X}^{(1)}\right] = E\left[\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} E[X_1] \\ E[X_2] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

(b) $E\left(\mathbf{AX^{(1)}}\right)$

$$E\left[\mathbf{AX}^{(1)}\right] = E\left[\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} 1 & 2 \end{bmatrix} E\left[\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = 4+6 = 10$$

(c) $\text{Cov}\left(\mathbf{X^{(1)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(1)}}\right) = \mathbf{\Sigma}_{11} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

(d) $\text{Cov}\left(\mathbf{AX^{(1)}}\right)$

$$\text{Cov}\left(\mathbf{AX^{(1)}}\right) = \mathbf{A}\text{Cov}\left(\mathbf{X^{(1)}}\right)\mathbf{A'} = \mathbf{A\Sigma}_{11}\mathbf{A'} = \begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}3 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix} =$$

$$= \begin{bmatrix}3 & 2\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix} = 7$$

(e) $E\left(\mathbf{X^{(2)}}\right)$

$$E\left[\mathbf{X}^{(2)}\right] = E\left[\begin{bmatrix}X_3\\X_4\end{bmatrix}\right] = \begin{bmatrix}E[X_3]\\E[X_4]\end{bmatrix} = \begin{bmatrix}\mu_3\\\mu_4\end{bmatrix} = \begin{bmatrix}2\\1\end{bmatrix}$$

(f) $E\left(\mathbf{BX^{(2)}}\right)$

$$E\left[\mathbf{BX}^{(2)}\right] = E\left[\begin{bmatrix}1 & -2\\2 & -1\end{bmatrix}\begin{bmatrix}X_3\\X_4\end{bmatrix}\right] = \begin{bmatrix}1 & -2\\2 & -1\end{bmatrix}E\left[\begin{bmatrix}X_3\\X_4\end{bmatrix}\right] = \begin{bmatrix}1 & -2\\2 & -1\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}0\\3\end{bmatrix}$$

(g) $\text{Cov}\left(\mathbf{X^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(2)}}\right) = \mathbf{\Sigma}_{22} = \begin{bmatrix}9 & -2\\-2 & 4\end{bmatrix}$$

(h) $\text{Cov}\left(\mathbf{BX^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{BX^{(2)}}\right) = \mathbf{B}\text{Cov}\left(\mathbf{X^{(2)}}\right)\mathbf{B'} = \mathbf{B\Sigma}_{22}\mathbf{B'} = \begin{bmatrix}1 & -2\\2 & -1\end{bmatrix}\begin{bmatrix}9 & -2\\-2 & 4\end{bmatrix}\begin{bmatrix}1 & 2\\-2 & -1\end{bmatrix} =$$

$$= \begin{bmatrix}13 & -10\\20 & -8\end{bmatrix}\begin{bmatrix}1 & 2\\-2 & -1\end{bmatrix} = \begin{bmatrix}33 & 36\\36 & 48\end{bmatrix}$$

(i) $\text{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right) = \mathbf{\Sigma}_{12} = \begin{bmatrix}2 & 2\\1 & 0\end{bmatrix}$$

(j) $\text{Cov}\left(\mathbf{AX^{(1)}}, \mathbf{BX^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{AX^{(1)}}, \mathbf{BX^{(2)}}\right) = \mathbf{A}\text{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right)\mathbf{B'} = \mathbf{A\Sigma}_{12}\mathbf{B'} =$$

$$= \begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}2 & 2\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 2\\-2 & -1\end{bmatrix} = \begin{bmatrix}4 & 2\end{bmatrix}\begin{bmatrix}1 & 2\\-2 & -1\end{bmatrix} = \begin{bmatrix}0 & 6\end{bmatrix}$$

2.31

Repeat **Exercise 2.30**, but with $\mathbf{A}$ and $\mathbf{B}$ replaced by

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$$

(a) $E\left(\mathbf{X^{(1)}}\right)$

$$E\left[\mathbf{X}^{(1)}\right] = E\left[\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} E[X_1] \\ E[X_2] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

(b) $E\left(\mathbf{AX^{(1)}}\right)$

$$E\left[\mathbf{AX}^{(1)}\right] = E\left[\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} 1 & -1 \end{bmatrix} E\left[\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = 4 - 3 = 1$$

(c) $\text{Cov}\left(\mathbf{X^{(1)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(1)}}\right) = \mathbf{\Sigma}_{11} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

(d) $\text{Cov}\left(\mathbf{AX^{(1)}}\right)$

$$\text{Cov}\left(\mathbf{AX^{(1)}}\right) = \mathbf{A}\text{Cov}\left(\mathbf{X^{(1)}}\right)\mathbf{A}' = \mathbf{A\Sigma}_{11}\mathbf{A}' = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} =$$

$$= \begin{bmatrix} 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 4$$

(e) $E\left(\mathbf{X^{(2)}}\right)$

$$E\left[\mathbf{X}^{(2)}\right] = E\left[\begin{bmatrix} X_3 \\ X_4 \end{bmatrix}\right] = \begin{bmatrix} E[X_3] \\ E[X_4] \end{bmatrix} = \begin{bmatrix} \mu_3 \\ \mu_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(f) $E\left(\mathbf{BX^{(2)}}\right)$

$$E\left[\mathbf{BX}^{(2)}\right] = E\left[\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_3 \\ X_4 \end{bmatrix}\right] = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} E\left[\begin{bmatrix} X_3 \\ X_4 \end{bmatrix}\right] = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

(g) $\text{Cov}\left(\mathbf{X^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(2)}}\right) = \mathbf{\Sigma}_{22} = \begin{bmatrix} 9 & -2 \\ -2 & 4 \end{bmatrix}$$

(h) $\text{Cov}\left(\mathbf{BX^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{BX^{(2)}}\right) = \mathbf{B}\text{Cov}\left(\mathbf{X^{(2)}}\right)\mathbf{B}' = \mathbf{B\Sigma}_{22}\mathbf{B}' = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 9 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 20 & -8 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 48 & -8 \\ -8 & 4 \end{bmatrix}$$

(i) $\text{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right) = \mathbf{\Sigma}_{12} = \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}$$

(j) $\text{Cov}\left(\mathbf{AX^{(1)}}, \mathbf{BX^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{AX^{(1)}}, \mathbf{BX^{(2)}}\right) = \mathbf{A}\text{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right)\mathbf{B}' = \mathbf{A\Sigma}_{12}\mathbf{B}' =$$

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

2.32

You are given the random vector $\mathbf{X}' = \begin{bmatrix} X_1, & X_2, & \cdots, & X_5 \end{bmatrix}$ with the mean vector $\boldsymbol{\mu}'_{\mathbf{X}} = \begin{bmatrix} 2, & 4, & -1, & 3, & 0 \end{bmatrix}$ and the variance-covariance matrix

$$\mathbf{\Sigma_X} = \begin{bmatrix} 4 & -1 & \frac{1}{2} & -\frac{1}{2} & 0 \\ -1 & 3 & 1 & -1 & 0 \\ \frac{1}{2} & 1 & 6 & 1 & -1 \\ -\frac{1}{2} & -1 & 1 & 4 & 0 \\ 0 & 0 & -1 & 0 & 2 \end{bmatrix}$$

Partition $\mathbf{X}$ as

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \hdashline X_3 \\ X_4 \\ X_5 \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \hdashline \mathbf{X}^{(2)} \end{bmatrix}$$

Let

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

(a) $E\left(\mathbf{X^{(1)}}\right)$

$$E\left[\mathbf{X}^{(1)}\right] = E\left[\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} E[X_1] \\ E[X_2] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

(b) $E\left(\mathbf{AX^{(1)}}\right)$

$$E\left[\mathbf{AX}^{(1)}\right] = E\left[\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} E\left[\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \end{bmatrix}$$

(c) $\text{Cov}\left(\mathbf{X^{(1)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(1)}}\right) = \mathbf{\Sigma}_{11} = \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix}$$

(d) $\mathrm{Cov}\left(\mathbf{AX^{(1)}}\right)$

$$\mathrm{Cov}\left(\mathbf{AX^{(1)}}\right) = \mathbf{A}\mathrm{Cov}\left(\mathbf{X^{(1)}}\right)\mathbf{A}' = \mathbf{A}\boldsymbol{\Sigma}_{11}\mathbf{A}' = \begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}4 & -1\\ -1 & 3\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} =$$

$$= \begin{bmatrix}5 & -4\\ 3 & 2\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} = \begin{bmatrix}9 & 1\\ 1 & 5\end{bmatrix}$$

(e) $E\left(\mathbf{X^{(2)}}\right)$

$$E\left[\mathbf{X}^{(2)}\right] = E\left[\begin{bmatrix}X_3\\ X_4\\ X_5\end{bmatrix}\right] = \begin{bmatrix}E[X_3]\\ E[X_4]\\ E[X_5]\end{bmatrix} = \begin{bmatrix}\mu_3\\ \mu_4\\ \mu_5\end{bmatrix} = \begin{bmatrix}-1\\ 3\\ 0\end{bmatrix}$$

(f) $E\left(\mathbf{BX^{(2)}}\right)$

$$E\left[\mathbf{BX}^{(2)}\right] = E\left[\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & -2\end{bmatrix}\begin{bmatrix}X_3\\ X_4\\ X_5\end{bmatrix}\right] = \begin{bmatrix}1 & 1 & 1\\ 1 & 1 & -2\end{bmatrix}E\left[\begin{bmatrix}X_3\\ X_4\\ X_5\end{bmatrix}\right] =$$

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

(g) $\mathrm{Cov}\left(\mathbf{X^{(2)}}\right)$

$$\mathrm{Cov}\left(\mathbf{X^{(2)}}\right) = \boldsymbol{\Sigma}_{22} = \begin{bmatrix}6 & 1 & -1\\ 1 & 4 & 0\\ -1 & 0 & 2\end{bmatrix}$$

(h) $\mathrm{Cov}\left(\mathbf{BX^{(2)}}\right)$

$$\mathrm{Cov}\left(\mathbf{BX^{(2)}}\right) = \mathbf{B}\mathrm{Cov}\left(\mathbf{X^{(2)}}\right)\mathbf{B}' = \mathbf{B}\boldsymbol{\Sigma}_{22}\mathbf{B}' =$$

$$= \begin{bmatrix}1 & 1 & 1\\ 1 & 1 & -2\end{bmatrix}\begin{bmatrix}6 & 1 & -1\\ 1 & 4 & 0\\ -1 & 0 & 2\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & -2\end{bmatrix} = \begin{bmatrix}6 & 5 & 1\\ 9 & 5 & -5\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & -2\end{bmatrix} =$$

$$= \begin{bmatrix}12 & 9\\ 9 & 24\end{bmatrix}$$

(i) $\mathrm{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right)$

$$\mathrm{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right) = \boldsymbol{\Sigma}_{12} = \begin{bmatrix}\frac{1}{2} & -\frac{1}{2} & 0\\ 1 & -1 & 0\end{bmatrix}$$

(j) Cov $\left(\mathbf{AX^{(1)}}, \mathbf{BX^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{AX^{(1)}}, \mathbf{BX^{(2)}}\right) = \mathbf{A}\text{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right)\mathbf{B}' = \mathbf{A}\mathbf{\Sigma}_{12}\mathbf{B}' =$$

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

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}$$

2.33

Repeat Exercise 2.32, but with $\mathbf{X}$ partitioned as

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \hdashline X_4 \\ X_5 \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \hdashline \mathbf{X}^{(2)} \end{bmatrix}$$

and with $\mathbf{A}$ and $\mathbf{B}$ replaced by

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 1 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$$

(a) $E\left(\mathbf{X^{(1)}}\right)$

$$E\left[\mathbf{X}^{(1)}\right] = E\left[\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ E[X_3] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ -1 \end{bmatrix}$$

(b) $E\left(\mathbf{AX^{(1)}}\right)$

$$E\left[\mathbf{AX}^{(1)}\right] = E\left[\begin{bmatrix} 2 & -1 & 0 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 1 & 3 \end{bmatrix} E\left[\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right] =$$

$$= \begin{bmatrix} 2 & -1 & 0 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

(c) Cov $\left(\mathbf{X^{(1)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(1)}}\right) = \mathbf{\Sigma}_{11} = \begin{bmatrix} 4 & -1 & \frac{1}{2} \\ -1 & 3 & 1 \\ \frac{1}{2} & 1 & 6 \end{bmatrix}$$

(d) Cov $\left(\mathbf{AX^{(1)}}\right)$

$$\text{Cov}\left(\mathbf{AX^{(1)}}\right) = \mathbf{A}\text{Cov}\left(\mathbf{X^{(1)}}\right)\mathbf{A}' = \mathbf{A\Sigma}_{11}\mathbf{A}' =$$

$$= \begin{bmatrix} 2 & -1 & 0 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 4 & -1 & \frac{1}{2} \\ -1 & 3 & 1 \\ \frac{1}{2} & 1 & 6 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 9 & -5 & 0 \\ \frac{9}{2} & 5 & \frac{39}{2} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 23 & 4 \\ 4 & 68 \end{bmatrix}$$

(e) $E\left(\mathbf{X^{(2)}}\right)$

$$E\left[\mathbf{X}^{(2)}\right] = E\left[\begin{bmatrix} X_4 \\ X_5 \end{bmatrix}\right] = \begin{bmatrix} E[X_4] \\ E[X_5] \end{bmatrix} = \begin{bmatrix} \mu_4 \\ \mu_5 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

(f) $E\left(\mathbf{BX^{(2)}}\right)$

$$E\left[\mathbf{BX}^{(2)}\right] = E\left[\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} X_4 \\ X_5 \end{bmatrix}\right] = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} E\left[\begin{bmatrix} X_4 \\ X_5 \end{bmatrix}\right] =$$

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

(g) Cov $\left(\mathbf{X^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(2)}}\right) = \mathbf{\Sigma}_{22} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

(h) Cov $\left(\mathbf{BX^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{BX^{(2)}}\right) = \mathbf{B}\text{Cov}\left(\mathbf{X^{(2)}}\right)\mathbf{B}' = \mathbf{B\Sigma}_{22}\mathbf{B}' =$$

$$= \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} =$$

$$= \begin{bmatrix} 12 & 0 \\ 0 & 6 \end{bmatrix}$$

(i) Cov $\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right) = \mathbf{\Sigma}_{12} = \begin{bmatrix} -\frac{1}{2} & 0 \\ -1 & 0 \\ 1 & -1 \end{bmatrix}$$

(j) Cov $\left(\mathbf{AX^{(1)}}, \mathbf{BX^{(2)}}\right)$

$$\text{Cov}\left(\mathbf{AX^{(1)}}, \mathbf{BX^{(2)}}\right) = \mathbf{A}\text{Cov}\left(\mathbf{X^{(1)}}, \mathbf{X^{(2)}}\right)\mathbf{B}' = \mathbf{A\Sigma}_{12}\mathbf{B}' =$$

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

$$= \begin{bmatrix} 0 & 0 \\ -\frac{9}{2} & \frac{9}{2} \end{bmatrix}$$

2.34

Consider the vectors $\mathbf{b}' = \begin{bmatrix}2, & -1, & 4, & 0\end{bmatrix}$ and $\mathbf{d}' = \begin{bmatrix}-1, & 3, & -2, & 1\end{bmatrix}$. Verify the Cauchy-schwarz inequality $(\mathbf{b}'\mathbf{d})^2 \leq (\mathbf{b}'\mathbf{b})\,(\mathbf{d}'\mathbf{d})$.

$$\left(\mathbf{b}'\mathbf{d}\right)^2 = \left(\begin{bmatrix}2 & -1 & 4 & 0\end{bmatrix}\begin{bmatrix}-1\\3\\-2\\1\end{bmatrix}\right)^2 = (-13)^2 = 169$$

$$\left(\mathbf{b}'\mathbf{b}\right) = \|\mathbf{b}\|^2 = \begin{bmatrix}2 & -1 & 4 & 0\end{bmatrix}\begin{bmatrix}2\\-1\\4\\0\end{bmatrix} = 21$$

$$\left(\mathbf{d}'\mathbf{d}\right) = \|\mathbf{d}\|^2 = \begin{bmatrix}-1 & 3 & -2 & 1\end{bmatrix}\begin{bmatrix}-1\\3\\-2\\1\end{bmatrix} = 15$$

$$\|\mathbf{b}\|^2\|\mathbf{d}\|^2 = (21)(15) = 315$$

$$169 = \left(\mathbf{b}'\mathbf{d}\right)^2 < \|\mathbf{b}\|^2\|\mathbf{d}\|^2 = 315$$

2.35

Using the vector $\mathbf{b}' = \begin{bmatrix}-4, & 3\end{bmatrix}$ and $\mathbf{d}' = \begin{bmatrix}1, & 1\end{bmatrix}$, verify the extended Cauchy-schwarz inequality $(\mathbf{b}'\mathbf{d})^2 \leq (\mathbf{b}'\mathbf{B}\mathbf{b})\left(\mathbf{d}'\mathbf{B}^{-1}\mathbf{d}\right)$ if

$$\mathbf{B} = \begin{bmatrix}2 & -2\\-2 & 5\end{bmatrix}$$

$$\left(\mathbf{b}'\mathbf{d}\right)^2 = \left(\begin{bmatrix}-4 & 3\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\right)^2 = (-1)^2 = 1$$

$$\left(\mathbf{b}'\mathbf{B}\mathbf{b}\right) = \begin{bmatrix}-4 & 3\end{bmatrix}\begin{bmatrix}2 & -2\\-2 & 5\end{bmatrix}\begin{bmatrix}-4\\3\end{bmatrix} = \begin{bmatrix}-14 & 23\end{bmatrix}\begin{bmatrix}-4\\3\end{bmatrix} = 125$$

$$\mathbf{B}^{-1} = \frac{1}{6}\begin{bmatrix}5 & 2\\2 & 5\end{bmatrix}$$

$$\left(\mathbf{d}'\mathbf{B}^{-1}\mathbf{d}\right) = \begin{bmatrix}1 & 1\end{bmatrix}\left(\frac{1}{6}\begin{bmatrix}5 & 2\\2 & 5\end{bmatrix}\right)\begin{bmatrix}1\\1\end{bmatrix} = \left(\frac{1}{6}\right)\begin{bmatrix}7 & 4\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \frac{11}{6}$$

$$\left(\mathbf{b}'\mathbf{B}\mathbf{b}\right)\left(\mathbf{d}'\mathbf{B}^{-1}\mathbf{d}\right) = 125 \times \frac{11}{6} = \frac{1375}{6} = 229.16\bar{6}$$

$$1 = \left(\mathbf{b}'\mathbf{d}\right)^2 < \left(\mathbf{b}'\mathbf{B}\mathbf{b}\right)\left(\mathbf{d}'\mathbf{B}^{-1}\mathbf{d}\right) = 229.16\bar{6}$$

2.36

Find the maximum and minimum values of the quadratic form $4x_1^2+4x_2^2+6x_1x_2$ for all points $\mathbf{x}' = \begin{bmatrix} x_1, & x_2 \end{bmatrix}$ such that $\mathbf{x}'\mathbf{x}' = 1$.

Convert into matrix form

$$4x_1^2 + 4x_2^2 + 6x_1x_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}'\mathbf{A}\mathbf{x}$$

so we have that

$$\mathbf{A} = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$$

Find the eigenvalues of $\mathbf{A}$,

$$0 = |\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 4-\lambda & 3 \\ 3 & 4-\lambda \end{vmatrix} = (4-\lambda)^2 - 9 = \lambda^2 - 8\lambda + 16 - 9 =$$

$$= \lambda^2 - 8\lambda + 7 = (\lambda - 7)(\lambda - 1)$$

The eigenvalues are $(\lambda_1, \lambda_2) = (7, 1)$. The matrix $\mathbf{A}$ is positive definite, since $\lambda_1 > \lambda_2 > 0$. By **(2-51)** on page 80, $\lambda_1 = 7$ is the maximum and $\lambda_2 = 1$ is the minimum.

2.37

With $\mathbf{A}$ as given in Exercise 2.6, find the maximum value of $\mathbf{x}'\mathbf{A}\mathbf{x}$ for $\mathbf{x}'\mathbf{x} = 1$.

From Exercise 2.6, the eigenvalues are $(\lambda_1, \lambda_2) = (10, 5)$. The matrix $\mathbf{A}$ is positive definite, since $\lambda_1 > \lambda_2 > 0$. By **(2-51)** on page 80, $\lambda_1 = 10$ is the maximum and $\lambda_2 = 5$ is the minimum.

2.38

Find the maximum and minimum values of the ratio $\mathbf{x}'\mathbf{A}\mathbf{x}/\mathbf{x}'\mathbf{x}$ for any nonzero vector $\mathbf{x}' = \begin{bmatrix} x_1, & x_2, & x_3 \end{bmatrix}$ if

$$\mathbf{A} = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}$$

Finding the eigenvalues of $\mathbf{A}$,

$$0 = |\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{vmatrix} =$$

$$= (13-\lambda)\left[(13-\lambda)(10-\lambda) - 4\right] + 4\left[-4(10-\lambda) + 4\right] + 2\left[8 - 2(13-\lambda)\right] =$$

$$= (13-\lambda)^2(10-\lambda) - 4(13-\lambda) - 16(10-\lambda) + 16 + 16 - 4(13-\lambda)$$

$$= \left(\lambda^2 - 26\lambda + 169\right)(10-\lambda) - 52 + 4\lambda - 160 + 16\lambda + 32 - 52 + 4\lambda =$$

$$= \left(10\lambda^2 - 260\lambda + 1690\right) + \left(-\lambda^3 + 26\lambda^2 - 169\lambda\right) 24\lambda - 232 =$$

$$= 1458 - 405\lambda + 36\lambda^2 - \lambda^3 =$$

$$= -\left(\lambda^3 - 36\lambda^2 + 405\lambda - 1458\right) =$$

$$= -\left(\lambda^2(\lambda-9) - 27\lambda(\lambda-9) + 162(\lambda-9)\right) =$$

$$= -(\lambda-9)\left(\lambda^2 - 27\lambda + 162\right) =$$

$$= -(\lambda-9)(\lambda-18)(\lambda-9)$$

The eigenvalues are $(\lambda_1, \lambda_2, \lambda_3) = (18, 9, 9)$. The matrix $\mathbf{A}$ is positive definite, since $\lambda_1 > \lambda_2 > \lambda_3 > 0$. By **(2-51)** on page 80, $\lambda_1 = 18$ is the maximum and $\lambda_2 = \lambda_3 = 9$ is the minimum.

2.39

Show that

$$\underset{(r\times s)}{\mathbf{A}}\underset{(s\times t))}{\mathbf{B}}\underset{(t\times v)}{C} \text{ has } (i,j)\text{th entry } \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{i\ell}b_{\ell k}c_{kj}$$

Hint: $\mathbf{BC}$ has (ℓ, j)th entry $\sum_{k=1}^{t} b_{\ell k}c_{kj} = d_{\ell j}$. So $\mathbf{A}(\mathbf{BC})$ has (i,j)th element

$$a_{i1}d_{1j} + a_{i2}d_{2j} + \cdots + a_{is}d_{sj} = \sum_{\ell=1}^{s} a_{i\ell}\left(\sum_{k=1}^{t} b_{\ell k}c_{kj}\right) = \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{i\ell}b_{\ell k}c_{kj}$$

$$\mathbf{ABC} = \mathbf{A}(\mathbf{BC}) = \mathbf{A}\left(\begin{bmatrix} b_{11} & \cdots & b_{1t} \\ \vdots & \ddots & \vdots \\ b_{s1} & \cdots & b_{st} \end{bmatrix}\begin{bmatrix} c_{11} & \cdots & c_{1v} \\ \vdots & \ddots & \vdots \\ c_{t1} & \cdots & c_{tv} \end{bmatrix}\right) =$$

$$= \mathbf{A}\begin{bmatrix} \sum_{k=1}^{t} b_{1k}c_{k1} & \cdots & \sum_{k=1}^{t} b_{1k}c_{kv} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{t} b_{sk}c_{k1} & \cdots & \sum_{k=1}^{t} b_{sk}c_{kv} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1s} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rs} \end{bmatrix}\begin{bmatrix} \sum_{k=1}^{t} b_{1k}c_{k1} & \cdots & \sum_{k=1}^{t} b_{1k}c_{kv} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{t} b_{sk}c_{k1} & \cdots & \sum_{k=1}^{t} b_{sk}c_{kv} \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{\ell=1}^{s} a_{1\ell}\left(\sum_{k=1}^{t} b_{\ell k}c_{k1}\right) & \cdots & \sum_{\ell=1}^{s} a_{1\ell}\left(\sum_{k=1}^{t} b_{\ell k}c_{kv}\right) \\ \vdots & \ddots & \vdots \\ \sum_{\ell=1}^{s} a_{r\ell}\left(\sum_{k=1}^{t} b_{\ell k}c_{k1}\right) & \cdots & \sum_{\ell=1}^{s} a_{r\ell}\left(\sum_{k=1}^{t} b_{\ell k}c_{kv}\right) \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{1\ell}b_{\ell k}c_{k1} & \cdots & \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{1\ell}b_{\ell k}c_{kv} \\ \vdots & \ddots & \vdots \\ \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{r\ell}b_{\ell k}c_{k1} & \cdots & \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{r\ell}b_{\ell k}c_{kv} \end{bmatrix}$$

The output of $\mathbf{ABC}$ is a $r \times v$ matrix, whose elements are

$$\sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{i\ell}b_{\ell k}c_{kj}$$

where $i = 1, \ldots, r$ and $j = 1, \ldots, v$.

2.40

Verify **(2-24)**: $E(\mathbf{X}+\mathbf{Y}) = E(\mathbf{X}) + E(\mathbf{Y})$ and $E(\mathbf{AXB}) = \mathbf{A}E(\mathbf{X})\mathbf{B}$.
Hint: $\mathbf{X}+\mathbf{Y}$ has $X_{ij}+Y_{ij}$ as its (i,j)th element. Now, $E(X_{ij}+Y_{ij}) = E(X_{ij}) + E(Y_{ij})$ by the univariate property of expectation, and this last quantity is the (i,j)th element of $E(\mathbf{X}) + E(\mathbf{Y})$. Next (see Exercise 2.39), $\mathbf{AXB}$ has (i,j)th entry $\sum_\ell \sum_k a_{i\ell}X_{\ell k}b_{kj}$, and by the additive property of expectation,

$$E\left(\sum_\ell \sum_k a_{i\ell}X_{\ell k}b_{kj}\right) = \sum_\ell \sum_k a_{i\ell}E(X_{\ell k})b_{kj}$$

which is the (i,j)th element of $\mathbf{A}E(\mathbf{X})\mathbf{B}$.

Let both $\mathbf{X}$ and $\mathbf{Y}$ be $n \times p$ matrices, so

$$E[\mathbf{X}+\mathbf{Y}] = E\left[\begin{bmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{bmatrix} + \begin{bmatrix} Y_{11} & \cdots & Y_{1p} \\ \vdots & \ddots & \vdots \\ Y_{n1} & \cdots & Y_{np} \end{bmatrix}\right] =$$

$$= E\left[\begin{bmatrix} X_{11}+Y_{11} & \cdots & X_{1p}+Y_{1p} \\ \vdots & \ddots & \vdots \\ Y_{n1}+X_{n1} & \cdots & X_{np}+Y_{np} \end{bmatrix}\right] =$$

$$= \begin{bmatrix} E[X_{11}+Y_{11}] & \cdots & E[X_{1p}+Y_{1p}] \\ \vdots & \ddots & \vdots \\ E[X_{n1}+Y_{n1}] & \cdots & E[X_{np}+Y_{np}] \end{bmatrix} =$$

$$= \begin{bmatrix} E[X_{11}]+E[Y_{11}] & \cdots & E[X_{1p}]+E[Y_{1p}] \\ \vdots & \ddots & \vdots \\ E[X_{n1}]+E[Y_{n1}] & \cdots & E[X_{np}]+E[Y_{np}] \end{bmatrix} =$$

$$= \begin{bmatrix} E[X_{11}] & \cdots & E[X_{1p}] \\ \vdots & \ddots & \vdots \\ E[X_{n1}] & \cdots & E[X_{np}] \end{bmatrix} + \begin{bmatrix} E[Y_{11}] & \cdots & E[Y_{1p}] \\ \vdots & \ddots & \vdots \\ E[Y_{n1}] & \cdots & E[Y_{np}] \end{bmatrix} =$$

$$= E\left[\begin{bmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{bmatrix}\right] + E\left[\begin{bmatrix} Y_{11} & \cdots & Y_{1p} \\ \vdots & \ddots & \vdots \\ Y_{n1} & \cdots & Y_{np} \end{bmatrix}\right] =$$

$$= E\left[\mathbf{X}\right] + E\left[\mathbf{Y}\right]$$

Using Exercise 2.39

$$E\left(\mathbf{AXB}\right) =$$

$$= E\left[\begin{bmatrix} \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{1\ell}X_{\ell k}c_{k1} & \cdots & \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{1\ell}X_{\ell k}c_{kv} \\ \vdots & \ddots & \vdots \\ \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{r\ell}X_{\ell k}c_{k1} & \cdots & \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{r\ell}X_{\ell k}c_{kv} \end{bmatrix}\right] =$$

$$= \begin{bmatrix} E\left[\sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{1\ell}X_{\ell k}c_{k1}\right] & \cdots & E\left[\sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{1\ell}X_{\ell k}c_{kv}\right] \\ \vdots & \ddots & \vdots \\ E\left[\sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{r\ell}X_{\ell k}c_{k1}\right] & \cdots & E\left[\sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{r\ell}X_{\ell k}c_{kv}\right] \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{\ell=1}^{s}\sum_{k=1}^{t} E\left[a_{1\ell}X_{\ell k}c_{k1}\right] & \cdots & \sum_{\ell=1}^{s}\sum_{k=1}^{t} E\left[a_{1\ell}X_{\ell k}c_{kv}\right] \\ \vdots & \ddots & \vdots \\ \sum_{\ell=1}^{s}\sum_{k=1}^{t} E\left[a_{r\ell}X_{\ell k}c_{k1}\right] & \cdots & \sum_{\ell=1}^{s}\sum_{k=1}^{t} E\left[a_{r\ell}X_{\ell k}c_{kv}\right] \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{1\ell}E\left[X_{\ell k}\right]c_{k1} & \cdots & \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{1\ell}E\left[X_{\ell k}\right]c_{kv} \\ \vdots & \ddots & \vdots \\ \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{r\ell}E\left[X_{\ell k}\right]c_{k1} & \cdots & \sum_{\ell=1}^{s}\sum_{k=1}^{t} a_{r\ell}E\left[X_{\ell k}\right]c_{kv} \end{bmatrix} =$$

$$= \mathbf{A}E(\mathbf{X})\mathbf{B}$$

2.41

You are given the random vector $\mathbf{X}' = \begin{bmatrix} X_1, & X_2, & X_3, & X_4 \end{bmatrix}$ with the mean vector $\boldsymbol{\mu}_{\mathbf{X}}' = \begin{bmatrix} 3, & 2, & -2, & 0 \end{bmatrix}$ and variance-covariance matrix

$$\boldsymbol{\Sigma}_{\mathbf{X}} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Let

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix}$$

(a) Find $E(\mathbf{AX})$, the mean of $\mathbf{AX}$.

$$E[\mathbf{AX}] = \mathbf{A}E[\mathbf{X}] = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ 3 \end{bmatrix}$$

(b) Find $\text{Cov}(\mathbf{AX})$, the variances and covariances of $\mathbf{AX}$.

$$\text{Cov}(\mathbf{AX}) = \mathbf{A}\text{Cov}(\mathbf{X})\,\mathbf{A}' = \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{A}' =$$

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

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

$$= 3 \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & 36 \end{bmatrix}$$

(c) Which pairs of linear combinations have zero covariances?

$$\text{Cov}(X_1, X_2) = \text{Cov}(X_1, X_3) = \text{Cov}(X_2, X_3) = 0$$

2.42

Repeat Exercise 2.41, but with

$$\boldsymbol{\Sigma}_{\mathbf{X}} = \begin{bmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix}$$

(a) Find $E(\mathbf{AX})$, the mean of $\mathbf{AX}$.

$$E[\mathbf{AX}] = \mathbf{A}E[\mathbf{X}] = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ 3 \end{bmatrix}$$

(b) Find $\text{Cov}(\mathbf{AX})$, the variances and covariances of $\mathbf{AX}$.

$$\text{Cov}(\mathbf{AX}) = \mathbf{A}\text{Cov}(\mathbf{X})\,\mathbf{A}' = \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{A}' =$$

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

$$= \begin{bmatrix} 2 & -2 & 0 & 0 \\ 2 & 2 & -4 & 0 \\ 2 & 2 & 2 & -6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{bmatrix} =$$

$$= \begin{bmatrix} 4 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 24 \end{bmatrix}$$

(c) Which pairs of linear combinations have zero covariances?

$$\mathrm{Cov}\,(X_1, X_2) = \mathrm{Cov}\,(X_1, X_3) = \mathrm{Cov}\,(X_2, X_3) = 0$$

3 Chapter 3

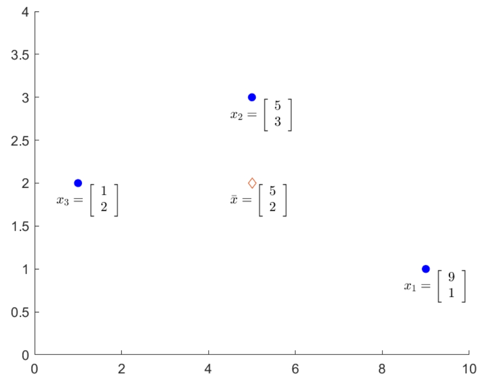
3.1

Given the data matrix

$$\mathbf{X} = \begin{bmatrix} 9 & 1 \\ 5 & 3 \\ 1 & 2 \end{bmatrix}$$

(a) Graph the scatter plot in $p = 2$ dimensions. Locate the sample mean on the diagram.

```
X = [9 1; 5 3; 1 2];
mean_pt = mean(X);
hold on
    % Plot data.
    scatter(X(:,1),X(:,2),'blue','filled')
    % Plot mean.
    plot(mean_pt(1),mean_pt(2),'d')
    % Text for mean.
    text(mean_pt(1)-0.5,mean_pt(2)-0.2, ...
        join(["$$\bar{x}=\left[\begin{array}{c}",mean_pt(1),"\\",mean_pt(2),"\end{array}\right]$$"],''), ...
        'interpreter','latex')
    % Text for data.
    for r = 1:height(X)
        anno = join(["$$x_",r," = \left[\begin{array}{c}",X(r,1),"\\",X(r,2),"\end{array}\right]$$"],'');
        text(X(r,1)-0.5,X(r,2)-0.2,anno,'interpreter','latex');
    end
    xlim([0 10])
    ylim([0 4])
    saveas(gcf,'sol3.1a.png')
hold off
```



(b) Sketch the $n = 3$-dimensional representation of the data, and plot the deviation vectors $\mathbf{y}_1 - \bar{x}_1\mathbf{1}$ and $\mathbf{y}_2 - \bar{x}_2\mathbf{1}$.

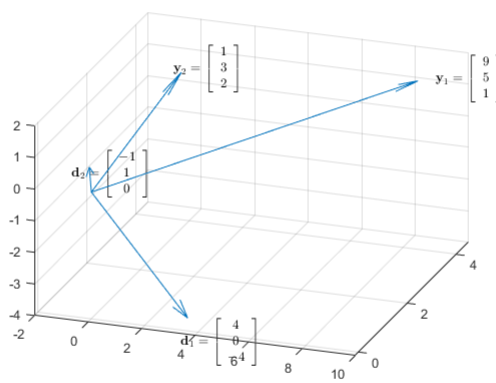
$$\mathbf{y}_1 - \bar{x}_1\mathbf{1} = \begin{bmatrix} 9 \\ 5 \\ 1 \end{bmatrix} - \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ -4 \end{bmatrix}$$

$$\mathbf{y}_2 - \bar{x}_2\mathbf{1} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

```
% Continuing from part (a)...
% Compute the deviation vectors.
d1 = X(:,1) - mean_pt(1)*ones([3,1]);
d2 = X(:,2) - mean_pt(2)*ones([3,1]);
% Combine the data with the deviation vectors. First two rows are data,
% second two are the deviation vectors.
D = [X d1 d2]';
start = zeros(size(D));

% Plot the y_1 and y_2 vectors and the d_1, d_2 deviation vectors.
quiver3(start(:,1), start(:,2), start(:,3), D(:,1), D(:,2), D(:,3));
% Text for data.
for r = 1:height(D)
    if r < 3
        % Labels for the data, y_1 and y_2.
        anno = join(["$$\textbf{y}_",r," = \left[\begin{array}{c}",D(r,1)
```

```
                    ,"\\",D(r,2),"\\",D(r,3),"\end{
                    array}\right]$$"],'');
                text(D(r,1)-0.5,D(r,2)-0.2,D(r,3)
                    -0.2,anno,'interpreter','latex'
                    );
            else
                % Labels for the deviation vectors
                    , d_1 and d_2.
                anno = join(["$$\textbf{d}_",r-2,"
                     = \left[\begin{array}{c}",D(r
                    ,1),"\\",D(r,2),"\\",D(r,3),"\
                    end{array}\right]$$"],'');
                text(D(r,1)-0.5,D(r,2)-0.2,D(r,3)
                    -0.2,anno,'interpreter','latex'
                    );
            end
        end
```



(c) Sketch the deviation vectors in (b) emanating from the origin. Calculate the lengths of these vectors and the cosine of the angle between them. Relate the quantities to $\mathbf{S}_n$ and $\mathbf{R}$.

The sketch of $\mathbf{d}_1$ and $\mathbf{d}_2$ are in part (b). The lengths are

$$\|\mathbf{d}_1\| = \sqrt{(4)^2 + (0)^2 + (-4)^2} = \sqrt{16+16} = \sqrt{32}$$

$$\|\mathbf{d}_2\| = \sqrt{(-1)^2 + (1)^2 + (0)^2} = \sqrt{2}$$

$$\mathbf{d}_1 \cdot \mathbf{d}_2 = \begin{bmatrix}4 & 0 & -4\end{bmatrix} \begin{bmatrix}-1\\1\\0\end{bmatrix} = -4$$

$$\cos\theta = \frac{\mathbf{d}_1 \cdot \mathbf{d}_2}{\|\mathbf{d}_1\|\,\|\mathbf{d}_2\|} = \frac{-4}{\sqrt{32}\sqrt{2}}$$

$$\Rightarrow \theta = \arccos\left(\frac{-4}{\sqrt{32}\sqrt{2}}\right) = \arccos\left(\frac{-4}{\sqrt{64}}\right) = \arccos\left(\frac{-4}{8}\right) = \arccos\left(\frac{-1}{2}\right) = 120^\circ$$

For $\mathbf{S}_n$, element $s_{12} = \frac{1}{n}(\mathbf{y}_1 - \bar{x}_1\mathbf{1})'(\mathbf{y}_2 - \bar{x}_2\mathbf{1}) = \frac{1}{n}\mathbf{d}_1 \cdot \mathbf{d}_2$, so what we computed for $\mathbf{d}_1\mathbf{d}_2 = n \times s_{12}$. For $\mathbf{R}$, when using $\mathbf{S}_n$, element $r_{12} = \frac{s_{12}}{\sqrt{s_{22}}\sqrt{s_{22}}} = \frac{(\mathbf{d}_1 \cdot \mathbf{d}_2/n)}{\sqrt{\mathbf{d}_1 \cdot \mathbf{d}_1/n}\sqrt{\mathbf{d}_2 \cdot \mathbf{d}_2/n}} = \frac{(\mathbf{d}_1 \cdot \mathbf{d}_2)}{\sqrt{\mathbf{d}_1 \cdot \mathbf{d}_1}\sqrt{\mathbf{d}_2 \cdot \mathbf{d}_2}}$

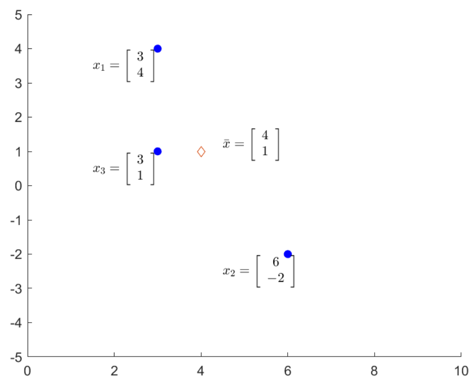
3.2

Given the data matrix

$$\mathbf{X} = \begin{bmatrix} 3 & 4 \\ 6 & -2 \\ 3 & 1 \end{bmatrix}$$

(a) Graph the scatter plot in $p = 2$ dimensions. Locate the sample mean on the diagram.

```
        X = [3 4; 6 -2; 3 1];
        mean_pt = mean(X);
        hold on
            % Plot data.
            scatter(X(:,1),X(:,2),'blue','filled')
            % Plot mean.
            plot(mean_pt(1),mean_pt(2),'d')
            % Text for mean.
            text(mean_pt(1)+0.5,mean_pt(2)+0.2, ...
                join(["$$\bar{x}=\left[\begin{array}{c}",mean_pt(1),"\\",mean_pt(2),"\end{array}\right]$$"],''), ...
                'interpreter','latex')
            % Text for data.
            for r = 1:height(X)
                anno = join(["$$x_",r," = \left[\begin{array}{c}",X(r,1),"\\",X(r,2),"\end{array}\right]$$"],'');
                text(X(r,1)-1.5,X(r,2)-0.5,anno,'interpreter','latex');
            end
            xlim([0 10])
            ylim([-5 5])
            saveas(gcf,'sol3.2a.png')
        hold off
```



(b) Sketch the $n = 3$-dimensional representation of the data, and plot the deviation vectors $\mathbf{y}_1 - \bar{x}_1\mathbf{1}$ and $\mathbf{y}_2 - \bar{x}_2\mathbf{1}$.

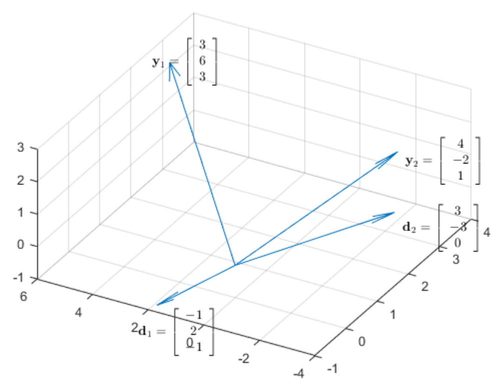
$$\mathbf{y}_1 - \bar{x}_1\mathbf{1} = \begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

$$\mathbf{y}_2 - \bar{x}_2\mathbf{1} = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}$$

```
        % Continuing from part (a)...
        % Compute the deviation vectors.
        d1 = X(:,1) - mean_pt(1)*ones([3,1]);
        d2 = X(:,2) - mean_pt(2)*ones([3,1]);
        % Combine the data with the deviation vectors. First two rows are data,
        % second two are the deviation vectors.
        D = [X d1 d2]';
        start = zeros(size(D));

        % Plot the y_1 and y_2 vectors and the d_1, d_2 deviation vectors.
        quiver3(start(:,1), start(:,2), start(:,3), D(:,1), D(:,2), D(:,3));
        % Text for data.
        for r = 1:height(D)
            if r < 3
                % Labels for the data, y_1 and y_2.
                anno = join(["$$\textbf{y}_",r," = \left[\begin{array}{c}",D(r,1)
```

```
                    ,"\\",D(r,2),"\\",D(r,3),"\end{array}\right]$$"],'');
                text(D(r,1)-0.5,D(r,2)-0.2,D(r,3)-0.2,anno,'interpreter','latex');
            else
                % Labels for the deviation vectors, d_1 and d_2.
                anno = join(["$$\textbf{d}_",r-2," = \left[\begin{array}{c}",D(r,1),"\\",D(r,2),"\\",D(r,3),"\end{array}\right]$$"],'');
                text(D(r,1)-0.5,D(r,2)-0.2,D(r,3)-0.2,anno,'interpreter','latex');
            end
        end
```



(c) Sketch the deviation vectors in (b) emanating from the origin. Calculate the lengths of these vectors and the cosine of the angle between them. Relate the quantities to $\mathbf{S}_n$ and $\mathbf{R}$.

The sketch of $\mathbf{d}_1$ and $\mathbf{d}_2$ are in part (b). The lengths are

$$\|\mathbf{d}_1\| = \sqrt{(-1)^2 + (2)^2 + (-1)^2} = \sqrt{1+4+1} = \sqrt{6}$$

$$\|\mathbf{d}_2\| = \sqrt{(3)^2 + (-3)^2 + (0)^2} = \sqrt{18}$$

$$\mathbf{d}_1 \cdot \mathbf{d}_2 = \begin{bmatrix} -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix} = -9$$

$$\cos\theta = \frac{\mathbf{d}_1 \cdot \mathbf{d}_2}{\|\mathbf{d}_1\|\,\|\mathbf{d}_2\|} = \frac{-9}{\sqrt{6}\sqrt{18}}$$

$$\Rightarrow \theta = \cos^{-1}\left(\frac{-9}{\sqrt{32}\sqrt{2}}\right) = \cos^{-1}\left(\frac{-9}{\sqrt{108}}\right) = \cos^{-1}\left(\frac{-3}{\sqrt{12}}\right) = \cos^{-1}\left(\frac{-\sqrt{3}}{2}\right) = 150°$$

For $\mathbf{S}_n$, element $s_{12} = \frac{1}{n}(\mathbf{y}_1 - \bar{x}_1\mathbf{1})'(\mathbf{y}_2 - \bar{x}_2\mathbf{1}) = \frac{1}{n}\mathbf{d}_1 \cdot \mathbf{d}_2$, so what we computed for $\mathbf{d}_1\mathbf{d}_2 = n \times s_{12}$. For $\mathbf{R}$, when using $\mathbf{S}_n$, element $r_{12} = \frac{s_{12}}{\sqrt{s_{22}}\sqrt{s_{22}}} = \frac{(\mathbf{d}_1 \cdot \mathbf{d}_2/n)}{\sqrt{\mathbf{d}_1 \cdot \mathbf{d}_1/n}\sqrt{\mathbf{d}_2 \cdot \mathbf{d}_2/n}} = \frac{(\mathbf{d}_1 \cdot \mathbf{d}_2)}{\sqrt{\mathbf{d}_1 \cdot \mathbf{d}_1}\sqrt{\mathbf{d}_2 \cdot \mathbf{d}_2}}$

3.3

Perform the decomposition of $\mathbf{y}_1$ into $\bar{x}_1\mathbf{1}$ and $\mathbf{y}_1 - \bar{x}_1\mathbf{1}$ using the first column of the data matrix in Example 3.9.

$$X = \begin{bmatrix} 1 & 2 & 5 \\ 4 & 1 & 6 \\ 4 & 0 & 4 \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 \end{bmatrix}$$

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} (1+4+4)/3 \\ (2+1+0)/3 \\ (5+6+4)/3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}$$

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 4 \\ 4 \end{bmatrix}, \qquad \bar{x}_1\mathbf{1}_3 = 3\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

$$\mathbf{y}_1 - \bar{x}_1\mathbf{1}_3 = \begin{bmatrix} 1 \\ 4 \\ 4 \end{bmatrix} - \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

3.4

Use the six observations on the variable X_1 in units of millions, from Table 1.1.

(a) Find the projection on $\mathbf{1}' = [1, 1, 1, 1, 1, 1]$.

Convert X_1 to units of millions

$$\mathbf{y}_1 = \left(\frac{1}{1000000}\right)\begin{bmatrix} 3497900 \\ 2485475 \\ 1782875 \\ 1725450 \\ 1645575 \\ 1469800 \end{bmatrix} = \begin{bmatrix} 3.4979 \\ 2.4855 \\ 1.7829 \\ 1.7254 \\ 1.6456 \\ 1.4698 \end{bmatrix}$$

$$\text{Proj}_{\mathbf{1}_6}\mathbf{y}_1 = \left(\frac{\mathbf{1}_6 \cdot \mathbf{y}_1}{\|\mathbf{1}_6\|}\right)\frac{\mathbf{1}_6}{\|\mathbf{1}_6\|} = \bar{x}_1\mathbf{1}_6 = 2.1012\begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}2.1012\\2.1012\\2.1012\\2.1012\\2.1012\\2.1012\end{bmatrix}$$

(b) Calculate the deviation vectot $\mathbf{y}_1 - \bar{x}_1\mathbf{1}$. Relate its length to the sample standard deviation.

$$\mathbf{y}_1 - \bar{x}_1\mathbf{1} = \begin{bmatrix}3.4979\\2.4855\\1.7829\\1.7254\\1.6456\\1.4698\end{bmatrix} - \begin{bmatrix}2.1012\\2.1012\\2.1012\\2.1012\\2.1012\\2.1012\end{bmatrix} = \begin{bmatrix}1.3967\\0.3843\\-0.3183\\-0.3757\\-0.4556\\-0.6314\end{bmatrix}$$

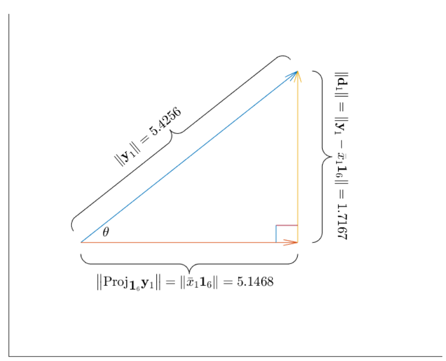
The sample variance for the first variable (feature, predictor, or whatever) is

$$\sqrt{s_{11}} = \sqrt{\frac{\sum_{i=1}^{6}(x_{i1} - \bar{x}_1)^2}{n-1}} = \sqrt{\frac{(\mathbf{y}_1 - \bar{x}_1\mathbf{1})'(\mathbf{y}_1 - \bar{x}_1\mathbf{1})}{n-1}} = \frac{\|\mathbf{y}_1 - \bar{x}_1\mathbf{1}\|}{\sqrt{n-1}}$$

$$\Rightarrow \|\mathbf{y}_1 - \bar{x}_1\mathbf{1}\| = \sqrt{s_{11}}\sqrt{n-1}$$

The length of $\mathbf{y}_1 - \bar{x}_1\mathbf{1}$ is the sample standard deviation times the square root of $n - 1$. For our data, $\|\mathbf{y}_1 - \bar{x}_1\mathbf{1}\| = 1.7167$.

(c) Graph (to scale) the triangle formed by $\mathbf{y}_1$, $\bar{x}_1\mathbf{1}$, and $\mathbf{y}_1 - \bar{x}_1\mathbf{1}$. Identify the length of each component on your graph.




```
        % Divide original units by 1,000,000 so we're in units of millions.
```

```
y1 = [3497900 2485475 1782875 1725450 1645575 1469800]'/1000000;
a1 = ones(height(y1),1);
% The projection of y1 onto a1.
x1bar1 = ((y1'*a1)/(norm(a1)*norm(a1)))*a1;

% Exercise 3.4 (b)
d1 = y1 - x1bar1;

% Exercise 3.4 (c)
clear clf
hold on
start = zeros(1,2);
quiver(start(:,1), start(:,2), 1, 1, 3);
quiver(start(:,1), start(:,2), 1, 0, 3);
quiver(3, 0, 0, 1, 3);
xlim([-1,5])
ylim([-2,4])
% Use the DRAWBRACE created by Pal Naverlid Savik
drawbrace([3.2, -0], [3.2, -3],10,1,'Color','k') % Draws a curly brace for deviance vector.
% Text for the norm of deviance vector d1, norm(d1).
anno = join(["$$\left\|\textbf{d}_1\right\|"," = \left\|\textbf{y}_1 - \bar{x}_1 \textbf{1}_6 \right\| = 1.7167$$"],'');
text(3.2+0.4,1.5+1.5,anno,'Rotation',270,'interpreter','latex');

drawbrace([0, 0.2], [3, 0.2],10,1,'Color','k') % Draws a curly brace for projection vector
% % Text for the norm of the projection vector, norm(x1bar1).
anno = join(["$$\left\|\rm{Proj}_{\textbf{1}_6}\textbf{y}_1 \right\| = \left\|\bar{x}_1\textbf{1}_6\right\|"," = 5.1468$$"],'');
text(0.2,-0.7,anno,'interpreter','latex');

drawbrace([-0.1, 0.2], [2.9, 3.2],10,0,'Color','k') % Draws a curly brace for y vector.
% Text for the norm of y2, norm(y1).
anno = join(["$$\left\|\textbf{y}_1 \right\| = "," 5.4256$$"],'');
text(0.5,1.4,anno,'Rotation',40,'interpreter','latex');

```

```
% Include text for angle theta.
text(0.3,0.2,'\theta')

% Include the right-angle symbol on plot.
plot([3,2.7],[0.3,0.3])
plot([2.7,2.7],[0,0.3])
set(gca,'xtick',[],'ytick',[]);
hold off
saveas(gcf,'sol3.4c.png')
```

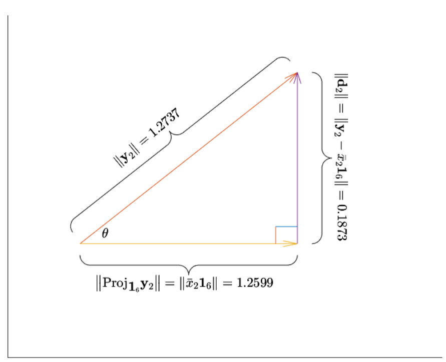
(d) Repeat Parts a-c for the variable X_2 in Table 1.1.

$$\mathbf{y}_2 = \begin{bmatrix} 0.623 \\ 0.593 \\ 0.512 \\ 0.500 \\ 0.463 \\ 0.395 \end{bmatrix}$$

$$\text{Proj}_{\mathbf{1}_6}\mathbf{y}_2 = \left(\frac{\mathbf{1}_6 \cdot \mathbf{y}_2}{\|\mathbf{1}_6\|}\right)\frac{\mathbf{1}_6}{\|\mathbf{1}_6\|} = \bar{x}_2\mathbf{1}_6 = 0.5143\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5143 \\ 0.5143 \\ 0.5143 \\ 0.5143 \\ 0.5143 \\ 0.5143 \end{bmatrix}$$

$$\mathbf{y}_2 - \bar{x}_2\mathbf{1} = \begin{bmatrix} 0.623 \\ 0.593 \\ 0.512 \\ 0.500 \\ 0.463 \\ 0.395 \end{bmatrix} - \begin{bmatrix} 0.5143 \\ 0.5143 \\ 0.5143 \\ 0.5143 \\ 0.5143 \\ 0.5143 \end{bmatrix} = \begin{bmatrix} 1.3967 \\ 0.3843 \\ -0.3183 \\ -0.3757 \\ -0.4556 \\ -0.6314 \end{bmatrix}$$

For our data, $\|\mathbf{y}_2 - \bar{x}_2\mathbf{1}\| = 0.1873$.



```
        y2 = [0.623 0.593 0.512 0.500 0.463 0.395]';
        % The projection of y2 onto a1.
        x2bar1 = ((y2'*a1)/(norm(a1)*norm(a1)))*a1;

        % Compute the deviance vector for y2.
        d2 = y2 - x2bar1;

        clear clf
        hold on
        start = zeros(1,2);
        quiver(start(:,1), start(:,2), 1, 1, 3);
        quiver(start(:,1), start(:,2), 1, 0, 3);
        quiver(3, 0, 0, 1, 3);
        xlim([-1,5])
        ylim([-2,4])
        % Use the DRAWBRACE created by Pal Naverlid Savik
        drawbrace([3.2, -0], [3.2, -3],10,1,'Color','k') % Draws a curly brace for deviance vector.
        % Text for the norm of deviance vector d2, norm(d2).
        anno = join(["$$\left\|\textbf{d}_2\right\|"," = \left\|\textbf{y}_2 - \bar{x}_2 \textbf{1}_6 \right\| = 0.1873$$"],'');
        text(3.2+0.4,1.5+1.5,anno,'Rotation',270,'interpreter','latex');

        drawbrace([0, 0.2], [3, 0.2],10,1,'Color','k') % Draws a curly brace for projection vector.
        % Text for the norm of the projection vector, norm(x2bar1).
        anno = join(["$$\left\|\rm{Proj}_{\textbf{1}_6}\textbf{y}_2 \right\| = \left\|\bar{x}_2\textbf{1}_6\right\|"," = 1.2599$$"],'');
        text(0.2,-0.7,anno,'interpreter','latex');

        drawbrace([-0.1, 0.2], [2.9, 3.2],10,0,'Color','k') % Draws a curly brace for y vector.
```

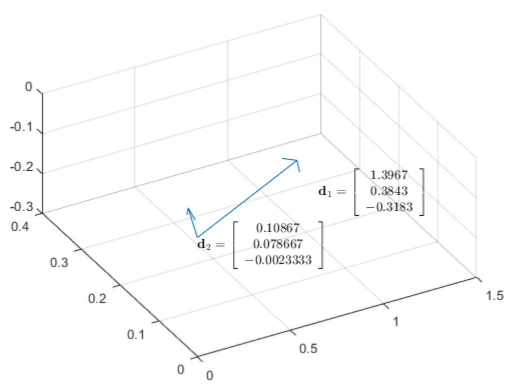
```
        % Text for the norm of y2, norm(y2).
        anno = join(["$$\left\|\textbf{y}_2 \right\|  = "," 1.2737$$"],'');
        text(0.5,1.4,anno,'Rotation',40,'interpreter','latex');

        % Include text for angle theta.
        text(0.3,0.2,'\theta')

        % Include the right-angle symbol on plot.
        plot([3,2.7],[0.3,0.3])
        plot([2.7,2.7],[0,0.3])
        set(gca,'xtick',[],'ytick',[]);
        hold off
        saveas(gcf,'sol3.4d.png')
```

(e) Graph (to scale) the two deviation vectors $\mathbf{y}_1 - \bar{x}_1\mathbf{1}$ and $\mathbf{y}_2 - \bar{x}_2\mathbf{1}$. Calculate the value of the angle between them.

The deviation vectors are in $\mathbb{R}^6$, so plotting that isn't feasible so here's the plot of the first 3 dimensions.



$$\cos\theta = \frac{\mathbf{d}_1 \cdot \mathbf{d}_2}{\|\mathbf{d}_1\| \, \|\mathbf{d}_2\|} = \frac{1}{1.7167 \times 0.1873} \begin{bmatrix} 1.3967 \\ 0.3843 \\ -0.3183 \\ -0.3757 \\ -0.4556 \\ -0.6314 \end{bmatrix}' \begin{bmatrix} 1.3967 \\ 0.3843 \\ -0.3183 \\ -0.3757 \\ -0.4556 \\ -0.6314 \end{bmatrix} = 0.8921$$

$$\Rightarrow \theta = \cos^{-1}(0.8921) = 26.8583^\circ$$

Also could do

$$\cos\theta = \frac{\mathbf{d}_1 \cdot \mathbf{d}_2}{\|\mathbf{d}_1\|\,\|\mathbf{d}_2\|} = \frac{(n-1)s_{12}}{\sqrt{(n-1)s_{11}}\sqrt{(n-1)s_{22}}} = \frac{s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}} = r_{12}$$

$$\Rightarrow \theta = \cos^{-1} r_{12} = \cos^{-1}(0.8921) = 26.8583^\circ$$

```
% Combine the deviation vectors.
D = [d1 d2]';
start = zeros(size(D));

clear clf
% Plot the d_1, d_2 deviation vectors.
quiver3(start(:,1), start(:,2), start(:,3), D(:,1), D(:,2), D(:,3));
% Text for data.
for r = 1:height(D)
    % Labels for the deviation vectors, d_1 and d_2.
    anno = join(["$$\textbf{d}_",r," = \left[\begin{array}{c}",D(r,1),"\\",D(r,2),"\\",D(r,3),"\end{array}\right]$$"],'');
    text(D(r,1)-0.05,D(r,2)-0.05,D(r,3)-0.05,anno,'interpreter','latex');
end

% Compute the angle between d1 and d2.
acosd(d1'*d2/(norm(d1)*norm(d2)))
% Same answer using (2-36) on page 72, R = V^{-1/2} \Sigma V^{-1/2}.
acosd(diag(sqrt(inv(diag(diag(cov(y1,y2)))))*cov(y1,y2)*sqrt(inv(diag(diag(cov(y1,y2))))),1))
```

3.5

Calculate the generalized sample variance $|\mathbf{S}|$ for (a) the data matrix $\mathbf{X}$ in Exercise 3.1 and (b) the data matrix $\mathbf{X}$ in Exercise 3.2.

(a)

$$\mathbf{X} = \begin{bmatrix} 9 & 1 \\ 5 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 \end{bmatrix}, \qquad \bar{\mathbf{x}} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

$$\mathbf{d}_1 = \mathbf{y}_1 - \bar{x}_1 \mathbf{1}_3 = \begin{bmatrix} 9 \\ 5 \\ 1 \end{bmatrix} - \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ -4 \end{bmatrix}$$

$$\mathbf{d}_2 = \mathbf{y}_2 - \bar{x}_2\mathbf{1}_3 = \begin{bmatrix}1\\3\\2\end{bmatrix} - \begin{bmatrix}2\\2\\2\end{bmatrix} = \begin{bmatrix}-1\\1\\0\end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix}\mathbf{d}_1 & \mathbf{d}_2\end{bmatrix} = \begin{bmatrix}4 & -1\\0 & 1\\-4 & 0\end{bmatrix}$$

$$\mathbf{S} = \left(\frac{1}{n-1}\right)\mathbf{D}'\mathbf{D} = \left(\frac{1}{3-1}\right)\begin{bmatrix}4 & 0 & -4\\-1 & 1 & 0\end{bmatrix}\begin{bmatrix}4 & -1\\0 & 1\\-4 & 0\end{bmatrix} =$$

$$\frac{1}{2}\begin{bmatrix}32 & -4\\-4 & 2\end{bmatrix} = \begin{bmatrix}16 & -2\\-2 & 1\end{bmatrix}$$

$$|\mathbf{S}| = \begin{vmatrix}16 & -2\\-2 & 1\end{vmatrix} = 16 - 4 = 12$$

(b)

$$\mathbf{X} = \begin{bmatrix}3 & 4\\6 & -2\\3 & 1\end{bmatrix} = \begin{bmatrix}\mathbf{y}_1 & \mathbf{y}_2\end{bmatrix}, \qquad \bar{\mathbf{x}} = \begin{bmatrix}4\\1\end{bmatrix}$$

$$\mathbf{d}_1 = \mathbf{y}_1 - \bar{x}_1\mathbf{1}_3 = \begin{bmatrix}3\\6\\3\end{bmatrix} - \begin{bmatrix}4\\4\\4\end{bmatrix} = \begin{bmatrix}-1\\2\\-1\end{bmatrix}$$

$$\mathbf{d}_2 = \mathbf{y}_2 - \bar{x}_2\mathbf{1}_3 = \begin{bmatrix}4\\-2\\1\end{bmatrix} - \begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}3\\-3\\0\end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix}\mathbf{d}_1 & \mathbf{d}_2\end{bmatrix} = \begin{bmatrix}-1 & 3\\2 & -3\\-1 & 0\end{bmatrix}$$

$$\mathbf{S} = \left(\frac{1}{n-1}\right)\mathbf{D}'\mathbf{D} = \left(\frac{1}{3-1}\right)\begin{bmatrix}-1 & 2 & -1\\3 & -3 & 0\end{bmatrix}\begin{bmatrix}-1 & 3\\2 & -3\\-1 & 0\end{bmatrix} =$$

$$\frac{1}{2}\begin{bmatrix}6 & -9\\-9 & 18\end{bmatrix} = \begin{bmatrix}3 & -(9/2)\\-(9/2) & 9\end{bmatrix}$$

$$|\mathbf{S}| = \begin{vmatrix}3 & -(9/2)\\-(9/2) & 9\end{vmatrix} = 27 - (81/4) = \frac{27}{4}$$

3.6

Consider the data matrix

$$\mathbf{X} = \begin{bmatrix} -1 & 3 & -2 \\ 2 & 4 & 2 \\ 5 & 2 & 3 \end{bmatrix}$$

(a) Calculate the matrix of deviations (residuals), $\mathbf{X} - \mathbf{1}\bar{\mathbf{x}}'$. Is this matrix of full rank? Explain.

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

$$\mathbf{X} - \mathbf{1}\bar{\mathbf{x}}' = \begin{bmatrix} -1 & 3 & -2 \\ 2 & 4 & 2 \\ 5 & 2 & 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} -1 & 3 & -2 \\ 2 & 4 & 2 \\ 5 & 2 & 3 \end{bmatrix} - \begin{bmatrix} 2 & 3 & 1 \\ 2 & 3 & 1 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} -3 & 0 & -3 \\ 0 & 1 & 1 \\ 3 & -1 & 2 \end{bmatrix}$$

No, the residual matrix is not full rank. The third column is column one plus column two, so there's a linear dependency. To be full rank the three columns in the square matrix must be linearly independent.

(b) Determine $\mathbf{S}$ and calculate the generalized sample variance $|\mathbf{S}|$. Interpret the latter geometrically.

$$\mathbf{S} = \left(\frac{1}{n-1}\right)\mathbf{D}^{-1}\mathbf{D} = \left(\frac{1}{2}\right) \begin{bmatrix} -3 & 0 & 3 \\ 0 & 1 & -1 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} -3 & 0 & -3 \\ 0 & 1 & 1 \\ 3 & -1 & 2 \end{bmatrix} =$$

$$= \left(\frac{1}{2}\right) \begin{bmatrix} 18 & -3 & 15 \\ -3 & 2 & -1 \\ 15 & -1 & 14 \end{bmatrix} = \begin{bmatrix} 9 & -(3/2) & (15/2) \\ -(3/2) & 1 & -(1/2) \\ (15/2) & -(1/2) & 7 \end{bmatrix}$$

$$|\mathbf{S}| = \begin{vmatrix} 9 & -(3/2) & (15/2) \\ -(3/2) & 1 & -(1/2) \\ (15/2) & -(1/2) & 7 \end{vmatrix} = 0$$

The matrix isn't full rank, so the determinant is 0. From result 3.2 on page 130, when at least one of the deviation vectors lies in the hyperplane formed by the linear combinations from the others, the generalized variance is zero.

(c) Using the results in (b), calculate the total sample variance. [See (3-23).]

The total samle variance is the trace of $\mathbf{S}$

$$tr\,\{\mathbf{S}\} = 9 + 1 + 7 = 17$$

3.7

Sketch the solid ellipsoids $(\mathbf{x}-\bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}-\bar{\mathbf{x}})$ [see (3-16)] for the three matrices

$$\mathbf{S}=\begin{bmatrix}5 & 4\\ 4 & 5\end{bmatrix},\qquad \mathbf{S}=\begin{bmatrix}5 & -4\\ -4 & 5\end{bmatrix},\qquad \mathbf{S}=\begin{bmatrix}3 & 0\\ 0 & 3\end{bmatrix}$$

(Note that these matrices have the *same* generalized variance $|\mathbf{S}|$.)

$$\mathbf{S}=\begin{bmatrix}5 & 4\\ 4 & 5\end{bmatrix}$$

$$0=|\mathbf{S}-\lambda\mathbf{I}|=\begin{vmatrix}5-\lambda & 4\\ 4 & 5-\lambda\end{vmatrix}=(5-\lambda)^2-16=\lambda^2-10\lambda+9=(\lambda-9)(\lambda-1)$$

The eigenvalues are $\{\lambda_1,\lambda_2\}=\{1,9\}$.
$\underline{\lambda_1=1:}$

$$\mathbf{S}\mathbf{x}_1=\lambda_1\mathbf{x}_1\Rightarrow\begin{bmatrix}5 & 4\\ 4 & 5\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}x_1\\ x_2\end{bmatrix}\Rightarrow\begin{bmatrix}4 & 4\\ 4 & 4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\mathbf{x}_1=\begin{bmatrix}1\\ -1\end{bmatrix}\Rightarrow\mathbf{e}_1=\begin{bmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{bmatrix}$$

$\underline{\lambda_2=9:}$

$$\mathbf{S}\mathbf{x}_2=\lambda_2\mathbf{x}_2\Rightarrow\begin{bmatrix}5 & 4\\ 4 & 5\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}9x_1\\ 9x_2\end{bmatrix}\Rightarrow\begin{bmatrix}-4 & 4\\ 4 & -4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\mathbf{x}_2=\begin{bmatrix}1\\ 1\end{bmatrix}\Rightarrow\mathbf{e}_2=\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$$

$$\mathbf{S}=\lambda_1\mathbf{e}_1\mathbf{e}_1'+\lambda_2\mathbf{e}_2\mathbf{e}_2'$$

$$\Rightarrow\mathbf{x}'\mathbf{S}\mathbf{x}=\mathbf{x}'\left(\lambda_1\mathbf{e}_1\mathbf{e}_1'+\lambda_2\mathbf{e}_2\mathbf{e}_2'\right)\mathbf{x}=\lambda_1\mathbf{x}'\mathbf{e}_1\mathbf{e}_1'\mathbf{x}+\lambda_2\mathbf{x}'\mathbf{e}_2\mathbf{e}_2'\mathbf{x}=$$

$$=\left(\frac{\mathbf{x}'\mathbf{e}_1}{(1/\sqrt{\lambda_1})}\right)\left(\frac{\mathbf{x}'\mathbf{e}_1}{(1/\sqrt{\lambda_1})}\right)'+\left(\frac{\mathbf{x}'\mathbf{e}_2}{(1/\sqrt{\lambda_2})}\right)\left(\frac{\mathbf{x}'\mathbf{e}_2}{(1/\sqrt{\lambda_2})}\right)'=$$

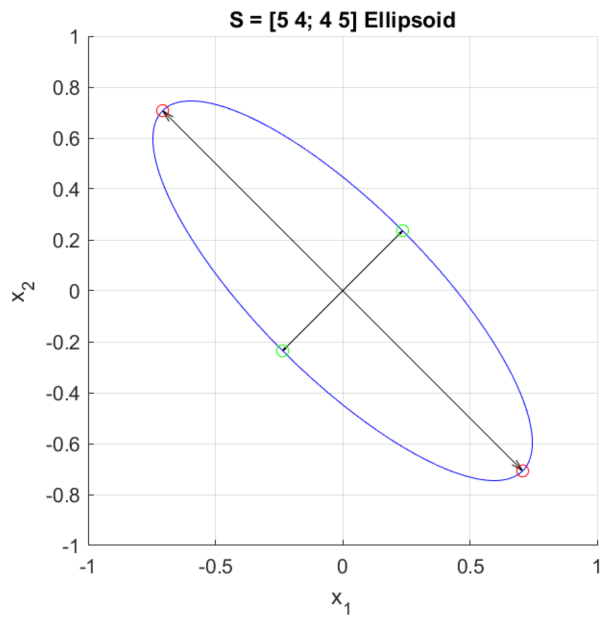
$$=\frac{(\mathbf{x}'\mathbf{e}_1)^2}{(1/\sqrt{\lambda_1})^2}+\frac{(\mathbf{x}'\mathbf{e}_2)^2}{(1/\sqrt{\lambda_2})^2}$$

Set equal to c^2 to have the equation of an ellipse centered at the origin.

$$\Rightarrow\mathbf{x}'\mathbf{S}\mathbf{x}=c^2=\frac{(\mathbf{x}'\mathbf{e}_1)^2}{(c/\sqrt{\lambda_1})^2}+\frac{(\mathbf{x}'\mathbf{e}_2)^2}{(c/\sqrt{\lambda_2})^2}$$

$$=\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}$$

The major axis (related to the smallest eigenvalue) is in the direction of $\mathbf{e}_1 = [(1/\sqrt{2}), -(1/\sqrt{2})]'$ with length $a = c/\sqrt{\lambda_1} = 1/1 = 1$. The minor axis (related to the largest eigenvalue) is in the direction of $\mathbf{e}_2 = [(1/\sqrt{2}), (1/\sqrt{2})]'$, with length $a = c/\sqrt{\lambda_2} = 1/3$. Here, $c = 1$.




```
S = [5 4; 4 5];
[V,D] = eig(S);
MyPlotEllipse(V,D,1,output_path,'sol3.7.1')
```

$$\mathbf{S} = \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}$$

$$0 = |\mathbf{S} - \lambda \mathbf{I}| = \begin{vmatrix} 5-\lambda & -4 \\ -4 & 5-\lambda \end{vmatrix} = (5-\lambda)^2 - 16 = \lambda^2 - 10\lambda + 9 = (\lambda - 9)(\lambda - 1)$$

The eigenvalues are $\{\lambda_1, \lambda_2\} = \{1, 9\}$.
$\underline{\lambda_1 = 1:}$

$$\mathbf{S}\mathbf{x}_1 = \lambda_1 \mathbf{x}_1 \Rightarrow \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & -4 \\ -4 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \mathbf{e}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$\underline{\lambda_2 = 9}$:

$$\mathbf{S}\mathbf{x}_2 = \lambda_2 \mathbf{x}_2 \Rightarrow \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 9x_1 \\ 9x_2 \end{bmatrix} \Rightarrow \begin{bmatrix} -4 & -4 \\ -4 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

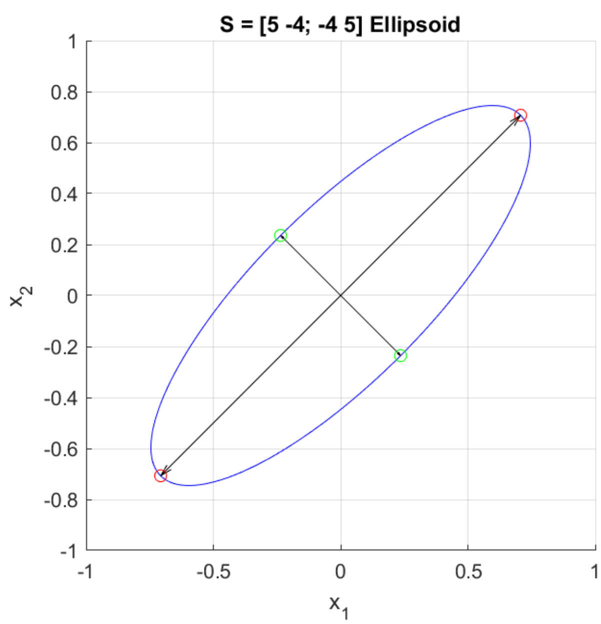
$$\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Rightarrow \mathbf{e}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

Set equal to c^2 to have the equation of an ellipse centered at the origin.

$$\Rightarrow \mathbf{x}'\mathbf{S}\mathbf{x} = c^2 = \frac{(\mathbf{x}'\mathbf{e}_1)^2}{(c/\sqrt{\lambda_1})^2} + \frac{(\mathbf{x}'\mathbf{e}_2)^2}{(c/\sqrt{\lambda_2})^2}$$

$$= \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2}$$

The major axis (related to the smallest eigenvalue) is in the direction of $\mathbf{e}_1 = [(1/\sqrt{2}), (1/\sqrt{2})]'$ with length $a = c/\sqrt{\lambda_1} = 1/1 = 1$. The minor axis (related to the largest eigenvalue) is in the direction of $\mathbf{e}_2 = [(1/\sqrt{2}), -(1/\sqrt{2})]'$, with length $a = c/\sqrt{\lambda_2} = 1/3$. Here, $c = 1$. This sample covariance matrix, $\mathbf{S}$, has the same eigenvalues as the one above, but the eigenvectors here are switched so the major axis is also switched.




```
S = [5 -4; -4 5];
[V,D] = eig(S);
MyPlotEllipse(V,D,1,output_path,'sol3.7.2')
```

$$\mathbf{S} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

$$0 = |\mathbf{S} - \lambda \mathbf{I}| = \begin{vmatrix} 3-\lambda & 0 \\ 0 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 0 = (\lambda - 3)(\lambda - 3)$$

The eigenvalues are both 3, $\{\lambda_1, \lambda_2\} = \{3, 3\}$.
$\underline{\lambda_1 = \lambda_2 = 3:}$

$$\mathbf{S}\mathbf{x}_1 = \lambda_1 \mathbf{x}_1 \Rightarrow \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 3x_1 \\ 3x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 \\ 3x_2 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Both eigenvalues are the same, so need to pick two vectors. They can be anything, so why not the stand basis vectors for $\mathbb{R}^2$.

$$\mathbf{x}_1 = \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

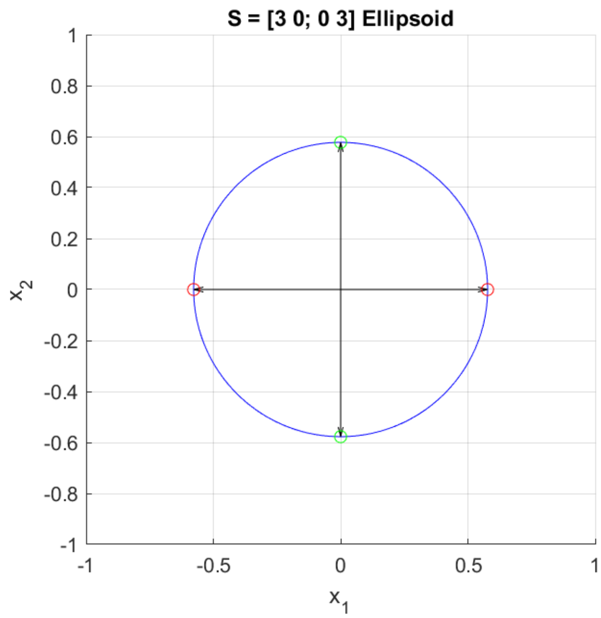
$$\mathbf{x}_2 = \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Set equal to c^2 to have the equation of an ellipse centered at the origin.

$$\Rightarrow \mathbf{x}'\mathbf{S}\mathbf{x} = c^2 = \frac{(\mathbf{x}'\mathbf{e}_1)^2}{(c/\sqrt{\lambda_1})^2} + \frac{(\mathbf{x}'\mathbf{e}_2)^2}{(c/\sqrt{\lambda_2})^2}$$

$$= \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2}$$

Here, we have two eigenvalues that are the same, so the lengths in the major and minor axis are also the same. The eigenvectors are the standard basis, so we have a circle centered at the origin with radius $a = b = c/\sqrt{\lambda_1} = 1/\sqrt{3}$. Here, $c = 1$.




```
S = [3 0; 0 3];
[V,D] = eig(S);
MyPlotEllipse(V,D,1,output_path,'sol3.7.3')
```

3.8

Given

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{S} = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$$

(a) Calculate the total sample variance for each $\mathbf{S}$. Compare the results.

$$tr\{\mathbf{S}\} = tr\left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\} = 3$$

$$tr\{\mathbf{S}\} = tr\left\{\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}\right\} = 3$$

Both sample covariance matrices have the same total sample variance values, since both have the same sample variance values of 1 on the diagonal. The total sample variance metric doesn't account for any the covariance structure for $i \neq j$ (off-diagonal values).

(b) Calculate the generalized sample variance for each $\mathbf{S}$, and compare the results. Comment on the discrepancies, if any, found between Parts a and b.

$$|\mathbf{S}| = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - 0 + 0 = 1(1-0) = 1$$

$$|\mathbf{S}| = \begin{vmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{vmatrix} = 1 \begin{vmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{vmatrix} + \frac{1}{2} \begin{vmatrix} -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{vmatrix} - \frac{1}{2} \begin{vmatrix} -\frac{1}{2} & 1 \\ -\frac{1}{2} & -\frac{1}{2} \end{vmatrix} =$$

$$= 1\left(1 - \frac{1}{4}\right) + \frac{1}{2}\left(-\frac{1}{2} - \frac{1}{4}\right) - \frac{1}{2}\left(\frac{1}{4} + \frac{1}{2}\right) = \frac{3}{4} - \frac{3}{8} - \frac{3}{8} = 0$$

The generalized sample variance of the diagonal matrix is 1. The sample covariance matrix form the standard basis for $\mathbb{R}^3$, so each element is one unit from zero and form a cube of length 1 on all sides. The volumn of this cube determined by the determinant for the generalized sample variance is of course 1. The second covariance matrix has $\text{Cov}(x_i, x_j) = -1/2 \; \forall \; i \neq j$. For this sample covariance matrix there is a linear dependence where the third column is -1 times the first column plus -1 times the second column. By result 3.2 on page 130, if at least one deviation vector lies in the (hyper) plane formed by all linear combos of the others then we have a linear dependence. If there's a linear dependence the parallelepiped will have volumn 0 (using (2) on page 133). The generalized sample variance (GSV) accounts for the off-diagonal covariance values, not simply the diagonal (variance) values, like the total sample variance does, so its result is more representative of the data when we have nonzero covariance. If the vectors are closely related the GSV is small, or zero if some vectors lie in the same (hyper) plane. If vectors are far from each other the GSV will be large.

3.9

The following data matrix contains data on test scores, with x_1 = score on first test, x_2 = score on second test, and x_3 = total score on the two tests:

$$\mathbf{X} = \begin{bmatrix} 12 & 17 & 29 \\ 18 & 20 & 38 \\ 14 & 16 & 30 \\ 20 & 18 & 38 \\ 16 & 19 & 35 \end{bmatrix}$$

(a) Obtain the mean corrected data matrix, and verify that the columns are linearly dependent. Specify an $\mathbf{a}' = [a_1, a_2, a_3]$ vector that establishes the linear dependence.

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} 80/5 \\ 90/5 \\ 170/5 \end{bmatrix} = \begin{bmatrix} 16 \\ 18 \\ 34 \end{bmatrix}$$

$$\mathbf{X} - \mathbf{1}_5\bar{\mathbf{x}}' = \begin{bmatrix} 12 & 17 & 29 \\ 18 & 20 & 38 \\ 14 & 16 & 30 \\ 20 & 18 & 38 \\ 16 & 19 & 35 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 16 & 18 & 34 \end{bmatrix} =$$

$$= \begin{bmatrix} 12 & 17 & 29 \\ 18 & 20 & 38 \\ 14 & 16 & 30 \\ 20 & 18 & 38 \\ 16 & 19 & 35 \end{bmatrix} - \begin{bmatrix} 16 & 18 & 34 \\ 16 & 18 & 34 \\ 16 & 18 & 34 \\ 16 & 18 & 34 \\ 16 & 18 & 34 \end{bmatrix} = \begin{bmatrix} -4 & -1 & -5 \\ 2 & 2 & 4 \\ -2 & -2 & -4 \\ 4 & 0 & 4 \\ 0 & 1 & 1 \end{bmatrix}$$

We have a linear dependence for column 3, whose the sum of the first two columns.

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$\mathbf{Xa} = \begin{bmatrix} 12 & 17 & 29 \\ 18 & 20 & 38 \\ 14 & 16 & 30 \\ 20 & 18 & 38 \\ 16 & 19 & 35 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}$$

or

$$(\mathbf{X} - \mathbf{1}_5\bar{\mathbf{x}}')\,\mathbf{a} = \begin{bmatrix} -4 & -1 & -5 \\ 2 & 2 & 4 \\ -2 & -2 & -4 \\ 4 & 0 & 4 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}$$

(b) Obtain the sample covariance matrix $\mathbf{S}$, and verify that the generalized variance is zero. Also, show that $\mathbf{Sa} = \mathbf{0}$, so a can be rescaled to be an eigenvector corresponding to eigenvalue zero.

$$\mathbf{D} = (\mathbf{X} - \mathbf{1}_5\bar{\mathbf{x}})$$

$$\mathbf{S} = \left(\frac{1}{n-1}\right)\mathbf{D}'\mathbf{D} = \left(\frac{1}{4}\right) \begin{bmatrix} -4 & 2 & -2 & 4 & 0 \\ -1 & 2 & -2 & 0 & 1 \\ -5 & 4 & -4 & 4 & 1 \end{bmatrix} \begin{bmatrix} -4 & -1 & -5 \\ 2 & 2 & 4 \\ -2 & -2 & -4 \\ 4 & 0 & 4 \\ 0 & 1 & 1 \end{bmatrix} =$$

$$= \left(\frac{1}{4}\right) \begin{bmatrix} 40 & 12 & 52 \\ 12 & 10 & 22 \\ 52 & 22 & 74 \end{bmatrix} = \begin{bmatrix} 10 & 3 & 13 \\ 3 & (5/2) & (11/2) \\ 13 & (11/2) & (37/2) \end{bmatrix}$$

Another way, using (3-27) on page 139,

$$\mathbf{S} = \left(\frac{1}{n-1}\right)\mathbf{X}'\left(\mathbf{I} - \frac{1}{n}\mathbf{1}_5\mathbf{1}_5'\right)\mathbf{X} = \left(\frac{1}{n-1}\right)\mathbf{X}'\left(\mathbf{X} - \frac{1}{n}\mathbf{1}_5\mathbf{1}_5'\mathbf{X}\right) =$$

$$= \left(\frac{1}{n-1}\right)\mathbf{X}'\left(\mathbf{X} - \mathbf{1}_5\bar{\mathbf{x}}'\right) = \left(\frac{1}{4}\right)\begin{bmatrix} 12 & 18 & 14 & 20 & 16 \\ 17 & 20 & 16 & 18 & 19 \\ 29 & 38 & 30 & 38 & 35 \end{bmatrix}\begin{bmatrix} -4 & -1 & -5 \\ 2 & 2 & 4 \\ -2 & -2 & -4 \\ 4 & 0 & 4 \\ 0 & 1 & 1 \end{bmatrix} =$$

$$= \left(\frac{1}{4}\right)\begin{bmatrix} 40 & 12 & 52 \\ 12 & 10 & 22 \\ 52 & 22 & 74 \end{bmatrix} = \begin{bmatrix} 10 & 3 & 13 \\ 3 & (5/2) & (11/2) \\ 13 & (11/2) & (37/2) \end{bmatrix}$$

Computing the generalized sample variance

$$|\mathbf{S}| = \begin{vmatrix} 10 & 3 & 13 \\ 3 & (5/2) & (11/2) \\ 13 & (11/2) & (37/2) \end{vmatrix} =$$

$$= 10\begin{vmatrix} (5/2) & (11/2) \\ (11/2) & (37/2) \end{vmatrix} - 3\begin{vmatrix} 3 & (11/2) \\ 13 & (37/2) \end{vmatrix} + 13\begin{vmatrix} 3 & (5/2) \\ 13 & (11/2) \end{vmatrix} =$$

$$= \frac{10}{4}(185 - 121) - \frac{3}{2}(111 - 143) + \frac{13}{2}(33 - 65) =$$

$$= \frac{640}{4} + \frac{96}{2} - \frac{416}{2} =$$

$$= \frac{416}{2} - \frac{416}{2} = 0$$

In part (a) we could see that the third column is the sum of the first two, so we defined a vector $\mathbf{a}$ as

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

and using that same vector to compute $\mathbf{Sa}$,

$$\mathbf{Sa} = \begin{bmatrix} 10 & 3 & 13 \\ 3 & (5/2) & (11/2) \\ 13 & (11/2) & (37/2) \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 13 - 13 \\ 11/2 - 11/2 \\ 37/2 - 37/2 \end{bmatrix} = \mathbf{0}$$

(c) Verify that the third column of the data matrix is the sum of the first two columns. That is, show that there is linear dependence, with $a_1 = 1$, $a_2 = 1$, and $a_3 = -1$.

In part (a) this was shown for the computation of $\mathbf{Xa} = \mathbf{0}$. Here's another way using column vectors in $\mathbf{X}$

$$\mathbf{Xa} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + a_3\mathbf{x}_3 = \mathbf{x}_1 + \mathbf{x}_2 - \mathbf{x}_3 =$$

$$= \begin{bmatrix}12\\18\\14\\20\\16\end{bmatrix} + \begin{bmatrix}17\\20\\16\\18\\19\end{bmatrix} - \begin{bmatrix}29\\38\\30\\38\\35\end{bmatrix} = \begin{bmatrix}29\\38\\30\\38\\35\end{bmatrix} - \begin{bmatrix}29\\38\\30\\38\\35\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix} = \mathbf{0}$$

3.10

When the generalized variance is zero, it is the columns of the mean corrected data matrix $\mathbf{X}_c = \mathbf{X} - \mathbf{1}\bar{\mathbf{x}}'$ that are linearly dependent, not necessarily those of the data matrix itself Given the data

$$\mathbf{X} = \begin{bmatrix}3 & 1 & 0\\6 & 4 & 6\\4 & 2 & 2\\7 & 0 & 3\\5 & 3 & 4\end{bmatrix}$$

(a) Obtain the mean corrected data matrix, and verify that the columns are linearly dependent. Specify an $\mathbf{a}' = [a_1, a_2, a_3]$ vector that establishes the linear dependence.

$$\bar{\mathbf{x}} = \begin{bmatrix}\bar{x}_1\\\bar{x}_2\\\bar{x}_3\end{bmatrix} = \begin{bmatrix}25/5\\10/5\\15/5\end{bmatrix} = \begin{bmatrix}5\\2\\3\end{bmatrix}$$

$$\mathbf{X} - \mathbf{1}_5\bar{\mathbf{x}}' = \begin{bmatrix}3 & 1 & 0\\6 & 4 & 6\\4 & 2 & 2\\7 & 0 & 3\\5 & 3 & 4\end{bmatrix} - \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}\begin{bmatrix}5 & 2 & 3\end{bmatrix} =$$

$$= \begin{bmatrix}3 & 1 & 0\\6 & 4 & 6\\4 & 2 & 2\\7 & 0 & 3\\5 & 3 & 4\end{bmatrix} - \begin{bmatrix}5 & 2 & 3\\5 & 2 & 3\\5 & 2 & 3\\5 & 2 & 3\\5 & 2 & 3\end{bmatrix} = \begin{bmatrix}-2 & -1 & -3\\1 & 2 & 3\\-1 & 0 & -1\\2 & -2 & 0\\0 & 1 & 1\end{bmatrix}$$

For $\mathbf{X}_c$ we have a linear dependence for column 3, whose the sum of the first two columns.

$$\mathbf{a} = \begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix} = \begin{bmatrix}1\\1\\-1\end{bmatrix}$$

In Exercise 3.9. both $\mathbf{X}_c$a $= \mathbf{0}$ and $\mathbf{Xa} = \mathbf{0}$ for the same $\mathbf{a}$, but that isn't the case here. Here, $r(\mathbf{X}) = p = 3$ so the columns of $\mathbf{X}$ are linearly independent, and the columns of $\mathbf{X}_c$ are linearly dependent ($r(\mathbf{X}_c) = 2 <$

$p = 3$).

$$(\mathbf{X} - \mathbf{1}_5\bar{\mathbf{x}}')\,\mathbf{a} = \begin{bmatrix} -2 & -1 & -3 \\ 1 & 2 & 3 \\ -1 & 0 & -1 \\ 2 & -2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}$$

(b) Obtain the sample covariance matrix $\mathbf{S}$, and verify that the generalized variance is zero.

$$\mathbf{D} = (\mathbf{X} - \mathbf{1}_5\bar{\mathbf{x}})$$

$$\mathbf{S} = \left(\frac{1}{n-1}\right)\mathbf{D}'\mathbf{D} = \left(\frac{1}{4}\right) \begin{bmatrix} -2 & 1 & -1 & 2 & 0 \\ -1 & 2 & 0 & -2 & 1 \\ -3 & 3 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & -1 & -3 \\ 1 & 2 & 3 \\ -1 & 0 & -1 \\ 2 & -2 & 0 \\ 0 & 1 & 1 \end{bmatrix} =$$

$$= \left(\frac{1}{4}\right) \begin{bmatrix} 10 & 0 & 10 \\ 0 & 10 & 10 \\ 10 & 10 & 20 \end{bmatrix} = \begin{bmatrix} (5/2) & 0 & (5/2) \\ 0 & (5/2) & (5/2) \\ (5/2) & (5/2) & 5 \end{bmatrix}$$

Another way, using (3-27) on page 139,

$$\mathbf{S} = \left(\frac{1}{n-1}\right)\mathbf{X}'\left(\mathbf{I} - \frac{1}{n}\mathbf{1}_5\mathbf{1}_5'\right)\mathbf{X} = \left(\frac{1}{n-1}\right)\mathbf{X}'\left(\mathbf{X} - \frac{1}{n}\mathbf{1}_5\mathbf{1}_5'\mathbf{X}\right) =$$

$$= \left(\frac{1}{n-1}\right)\mathbf{X}'(\mathbf{X} - \mathbf{1}_5\bar{\mathbf{x}}') = \left(\frac{1}{4}\right) \begin{bmatrix} 3 & 6 & 4 & 7 & 5 \\ 1 & 4 & 2 & 0 & 3 \\ 0 & 6 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} -2 & -1 & -3 \\ 1 & 2 & 3 \\ -1 & 0 & -1 \\ 2 & -2 & 0 \\ 0 & 1 & 1 \end{bmatrix} =$$

$$= \left(\frac{1}{4}\right) \begin{bmatrix} 10 & 0 & 10 \\ 0 & 10 & 10 \\ 10 & 10 & 20 \end{bmatrix} = \begin{bmatrix} (5/2) & 0 & (5/2) \\ 0 & (5/2) & (5/2) \\ (5/2) & (5/2) & 5 \end{bmatrix}$$

Computing the generalized sample variance

$$|\mathbf{S}| = \begin{vmatrix} (5/2) & 0 & (5/2) \\ 0 & (5/2) & (5/2) \\ (5/2) & (5/2) & 5 \end{vmatrix} =$$

$$= \left(\frac{5}{2}\right)\begin{vmatrix} (5/2) & (5/2) \\ (5/2) & 5 \end{vmatrix} - 0 + \left(\frac{5}{2}\right)\begin{vmatrix} 0 & (5/2) \\ (5/2) & (5/2) \end{vmatrix} =$$

$$= \frac{5}{8}(50-25) + \frac{5}{8}(0-25) =$$

$$= \frac{125}{8} - \frac{125}{8} = 0$$

(c) Show that the columns of the data matrix are linearly independent in this case.

To show this, work $\mathbf{X}$ into reduced row echelon form and count the pivot columns.

$$\mathbf{X} = \begin{bmatrix} 3 & 1 & 0 \\ 6 & 4 & 6 \\ 4 & 2 & 2 \\ 7 & 0 & 3 \\ 5 & 3 & 4 \end{bmatrix} \xrightarrow{\text{Simplify rows}} \begin{bmatrix} 3 & 1 & 0 \\ 3 & 2 & 3 \\ 2 & 1 & 1 \\ 7 & 0 & 3 \\ 5 & 3 & 4 \end{bmatrix} \xrightarrow{\text{Row } 5 - \frac{5}{3}\text{Row } 1}$$

$$\begin{bmatrix} 3 & 1 & 0 \\ 3 & 2 & 3 \\ 2 & 1 & 1 \\ 7 & 0 & 3 \\ 0 & (4/3) & 4 \end{bmatrix} \xrightarrow{\text{Row } 4 - \frac{7}{3}\text{Row } 1} \begin{bmatrix} 3 & 1 & 0 \\ 3 & 2 & 3 \\ 2 & 1 & 1 \\ 0 & -(7/3) & 3 \\ 0 & (4/3) & 4 \end{bmatrix} \xrightarrow{\text{Row } 3 - \frac{2}{3}\text{Row } 1}$$

$$\begin{bmatrix} 3 & 1 & 0 \\ 3 & 2 & 3 \\ 0 & (1/3) & 1 \\ 0 & -(7/3) & 3 \\ 0 & (4/3) & 4 \end{bmatrix} \xrightarrow{\text{Row } 3 - \text{Row } 1} \begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 3 \\ 0 & (1/3) & 1 \\ 0 & -(7/3) & 3 \\ 0 & (4/3) & 4 \end{bmatrix} \xrightarrow{\text{Simplify rows}}$$

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 3 \\ 0 & 1 & 3 \\ 0 & -7 & 9 \\ 0 & 1 & 3 \end{bmatrix} \xrightarrow{\text{Row } 5 - \text{Row } 2} \begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 3 \\ 0 & 1 & 3 \\ 0 & -7 & 9 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Row } 4 + 7\text{Row } 2}$$

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 30 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Row } 3 - \text{Row } 2} \begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 30 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Swap rows}}$$

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 30 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Simplify}} \begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Row } 2 - 3\text{Row } 3}$$

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Row } 1 - \text{Row } 2} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Simplify}}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

We can see that in reduced row echelon form that there are 3 pivot columns (nonzero rows), so the rank of $\mathbf{X}$ is 3 and is of full rank ($r(\mathbf{X}) = 3 = p$).

3.11

Use the sample covariance obtained in Example 3.7 to verify (3-29) and (3-30), which state that $\mathbf{R} = \mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2}$ and $\mathbf{D}^{1/2}\mathbf{R}\mathbf{D}^{1/2} = \mathbf{S}$.

$$\mathbf{S} = \begin{bmatrix} 252.04 & -68.43 \\ -68.43 & 123.67 \end{bmatrix}$$

From (3-28) on Page 139

$$\mathbf{D}^{1/2} = \begin{bmatrix} \sqrt{s_{11}} & 0 \\ 0 & \sqrt{s_{22}} \end{bmatrix} = \begin{bmatrix} \sqrt{252.04} & 0 \\ 0 & \sqrt{123.67} \end{bmatrix} = \begin{bmatrix} 15.87577 & 0 \\ 0 & 11.1207 \end{bmatrix}$$

$$\mathbf{D}^{-1/2} = \begin{bmatrix} \frac{1}{\sqrt{s_{11}}} & 0 \\ 0 & \frac{1}{\sqrt{s_{22}}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{252.04}} & 0 \\ 0 & \frac{1}{\sqrt{123.67}} \end{bmatrix} = \begin{bmatrix} 0.06298908 & 0 \\ 0 & 0.08992239 \end{bmatrix}$$

$$\mathbf{R} = \mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2} =$$

$$= \begin{bmatrix} 0.06298908 & 0 \\ 0 & 0.08992239 \end{bmatrix} \begin{bmatrix} 252.04 & -68.43 \\ -68.43 & 123.67 \end{bmatrix} \begin{bmatrix} 0.06298908 & 0 \\ 0 & 0.08992239 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -0.3876 \\ -0.3876 & 1 \end{bmatrix}$$

$$\mathbf{D}^{1/2}\mathbf{R}\mathbf{D}^{1/2} =$$

$$= \begin{bmatrix} 15.87577 & 0 \\ 0 & 11.1207 \end{bmatrix} \begin{bmatrix} 1 & -0.3876 \\ -0.3876 & 1 \end{bmatrix} \begin{bmatrix} 15.87577 & 0 \\ 0 & 11.1207 \end{bmatrix} =$$

$$\begin{bmatrix} 252.04 & -68.43 \\ -68.43 & 123.67 \end{bmatrix}$$

3.12

Show that $|\mathbf{S}| = (s_{11}s_{22}\cdots s_{pp})\,|\mathbf{R}|$.

The output of the determinant is scalar value, so we can rearrange them (commutative property), i.e., $|\mathbf{A}|\,|\mathbf{B}| = |\mathbf{B}|\,|\mathbf{A}|$.

$$|\mathbf{S}| \overset{\text{(3-30)}}{=} \left|\mathbf{D}^{1/2}\mathbf{R}\mathbf{D}^{1/2}\right| \overset{\text{Result 2A.11(e)}}{=} \left|\mathbf{D}^{1/2}\right| \left|\mathbf{R}\right| \left|\mathbf{D}^{1/2}\right| = \left|\mathbf{D}^{1/2}\right| \left|\mathbf{D}^{1/2}\right| \left|\mathbf{R}\right| =$$

$$= \left|\mathbf{D}^{1/2}\mathbf{D}^{1/2}\right| \left|\mathbf{R}\right| = |\mathbf{D}|\,|\mathbf{R}| = \prod_{i=1}^{p} s_{ii}\,|\mathbf{R}|$$

3.13

Given a data matrix $\mathbf{X}$ and the resulting sample correlation matrix $\mathbf{R}$, consider the standardized observations $(x_{ik} - \bar{x}_k)/\sqrt{s_{kk}}$, $k = 1, 2, \ldots, p$, $j = 1, 2, \ldots, n$. Show that these standardized quantities have sample covariance matrix $\mathbf{R}$.

Here are two solutions. One, using matrices and a second using the summations and algebra.

Solution 1:

$$[(x_{ik} - \bar{x}_k)/\sqrt{s_{kk}}]_{ik} = (\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2} = \mathbf{X}_\star$$

Use the formula for $\mathbf{S}$ from (3-27),

$$\mathbf{S} = \frac{1}{n-1}\mathbf{X}_\star'\left(\mathbf{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\mathbf{X}_\star =$$

$$= \frac{1}{n-1}\left((\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2}\right)'\left(\mathbf{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\left((\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2}\right) =$$

$$= \frac{1}{n-1}\left(\mathbf{D}^{-1/2}\right)'(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'\left(\mathbf{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\left((\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2}\right) =$$

$$= \frac{1}{n-1}\mathbf{D}^{-1/2}(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'\left(\mathbf{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2} =$$

$$= \frac{1}{n-1}\left(\mathbf{D}^{-1/2}(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'\,(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2} - \frac{1}{n}\mathbf{D}^{-1/2}(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'\mathbf{1}_n\mathbf{1}_n'\,(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2}\right) =$$

$$= \frac{1}{n-1}\left(\mathbf{D}^{-1/2}((n-1)\mathbf{S})\mathbf{D}^{-1/2} - \frac{1}{n}\mathbf{D}^{-1/2}(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'\,(\mathbf{1}_n\mathbf{1}_n'\mathbf{X} - \mathbf{1}_n\mathbf{1}_n'\mathbf{1}_n\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2}\right) =$$

$$= \frac{1}{n-1}\left((n-1)\mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2} - \frac{1}{n}\mathbf{D}^{-1/2}(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'\,(\mathbf{1}_n\,(\mathbf{1}_n'\mathbf{X}) - \mathbf{1}_n\,(\mathbf{1}_n'\mathbf{1}_n)\,\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2}\right) =$$

$$= \frac{1}{n-1}\left((n-1)\mathbf{R} - \frac{1}{n}\mathbf{D}^{-1/2}(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'\,(\mathbf{1}_n\,(n\bar{\mathbf{x}}') - \mathbf{1}_n\,(n)\,\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2}\right) =$$

$$= \frac{1}{n-1}\left((n-1)\mathbf{R} - \frac{1}{n}\mathbf{D}^{-1/2}(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'\,(n\mathbf{1}_n\bar{\mathbf{x}}' - n\mathbf{1}_n\bar{\mathbf{x}}')\,\mathbf{D}^{-1/2}\right) =$$

$$= \frac{1}{n-1}\left((n-1)\mathbf{R} - \frac{1}{n}\mathbf{D}^{-1/2}(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'0\mathbf{D}^{-1/2}\right) =$$

$$= \frac{1}{n-1}((n-1)\mathbf{R} - 0) =$$

$$= \frac{n-1}{n-1}\mathbf{R} =$$

$$= \mathbf{R}$$

Solution 2:
For an observation j variable k, define the standardized values as

$$y_{jk} = \frac{x_{jk} - \bar{x}_k}{\sqrt{s_{kk}}}$$

A mean value for variable k from the p by 1 mean vector would be

$$\bar{y}_k = \sum_{j=1}^{n} \frac{y_{jk}}{n} = \sum_{j=1}^{n} \frac{(x_{jk} - \bar{x}_k)}{n\sqrt{s_{kk}}}$$

An element ik from the p by p covariance matrix would be

$$s_{ik} = \left(\frac{1}{n-1}\right) \sum_{j=1}^{n} (y_{ji} - \bar{y}_i)(y_{jk} - \bar{y}_k) =$$

$$= \left(\frac{1}{n-1}\right) \sum_{j=1}^{n} \left(\left(\frac{x_{ji} - \bar{x}_i}{\sqrt{s_{ii}}}\right) - \left(\sum_{\ell=1}^{n} \frac{(x_{\ell i} - \bar{x}_i)}{n\sqrt{s_{ii}}}\right)\right) \left(\left(\frac{x_{jk} - \bar{x}_k}{\sqrt{s_{kk}}}\right) - \left(\sum_{\ell=1}^{n} \frac{(x_{\ell k} - \bar{x}_k)}{n\sqrt{s_{kk}}}\right)\right) =$$

$$= \left(\frac{1}{n-1}\right)\left(\frac{1}{\sqrt{s_{ii}}\sqrt{s_{kk}}}\right) \sum_{j=1}^{n} \left(x_{ji} - \bar{x}_i - \left(\sum_{\ell=1}^{n} \frac{x_{\ell i} - \bar{x}_i}{n}\right)\right) \left(x_{jk} - \bar{x}_k - \left(\sum_{\ell=1}^{n} \frac{x_{\ell k} - \bar{x}_k}{n}\right)\right) =$$

$$= \left(\frac{1}{n-1}\right)\left(\frac{1}{\sqrt{s_{ii}}\sqrt{s_{kk}}}\right) \sum_{j=1}^{n} \left(x_{ji} - \bar{x}_i - \left(\sum_{\ell=1}^{n} \frac{x_{\ell i}}{n}\right) + n\frac{\bar{x}_i}{n}\right) \left(x_{jk} - \bar{x}_k - \left(\sum_{\ell=1}^{n} \frac{x_{\ell k}}{n}\right) - n\frac{\bar{x}_k}{n}\right) =$$

$$= \left(\frac{1}{n-1}\right)\left(\frac{1}{\sqrt{s_{ii}}\sqrt{s_{kk}}}\right) \sum_{j=1}^{n} (x_{ji} - \bar{x}_i - \bar{x}_i + \bar{x}_i)(x_{jk} - \bar{x}_k - \bar{x}_k + \bar{x}_k) =$$

$$= \left(\frac{1}{n-1}\right)\left(\frac{1}{\sqrt{s_{ii}}\sqrt{s_{kk}}}\right) \sum_{j=1}^{n} (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k) =$$

$$= \left(\frac{1}{n-1}\right)\left(\frac{1}{\sqrt{s_{ii}}\sqrt{s_{kk}}}\right)(n-1)s_{ik} =$$

$$= \frac{s_{ik}}{\sqrt{s_{ii}}\sqrt{s_{kk}}} = r_{ik}$$

This is element ik from the p by p sample correlation matrix.

3.14

Consider the data matrix $\mathbf{X}$ in Exercise 3.1. We have $n = 3$ observations on $p = 2$ variables X_1 and X_2. Form the linear combinations

$$\mathbf{c}'\mathbf{X} = \begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = -X_1 + 2X_2$$

$$\mathbf{b}'\mathbf{X} = \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 2X_1 + 3X_2$$

(a) Evaluate the sample means, variances, and covariance of $\mathbf{b}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$ from first principles. That is, calculate the observed values of $\mathbf{b}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$ and then use the sample mean, variance, and covariance formulas.

$$\mathbf{X} = \begin{bmatrix} 9 & 1 \\ 5 & 3 \\ 1 & 2 \end{bmatrix}$$

$$\begin{aligned} \mathbf{c}'\mathbf{x}_1 &= -1x_{11} + 2x_{12} = -1(9) + 2(1) = -7 \\ \mathbf{c}'\mathbf{x}_2 &= -1x_{21} + 2x_{22} = -1(5) + 2(3) = 1 \\ \mathbf{c}'\mathbf{x}_3 &= -1x_{31} + 2x_{32} = -1(1) + 2(2) = 3 \end{aligned}$$

$$\text{Sample mean} = \frac{(-7+1+3)}{3} = -1$$

$$\text{Sample variance} = \frac{(-7+1)^2 + (1+1)^2 + (3+1)^2}{3-1} = 28$$

$$\begin{aligned} \mathbf{b}'\mathbf{x}_1 &= 2x_{11} + 3x_{12} = 2(9) + 3(1) = 21 \\ \mathbf{b}'\mathbf{x}_2 &= 2x_{21} + 3x_{22} = 2(5) + 3(3) = 19 \\ \mathbf{b}'\mathbf{x}_3 &= 2x_{31} + 3x_{32} = 2(1) + 3(2) = 8 \end{aligned}$$

$$\text{Sample mean} = \frac{(21+19+8)}{3} = 16$$

$$\text{Sample variance} = \frac{(21-16)^2 + (19-16)^2 + (8-16)^2}{3-1} = 49$$

The covariance between $\mathbf{b}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$

$$\text{Sample covariance} =$$

$$= \frac{(-7+1)(21-16) + (1+1)(19-16) + (3+1)(8-16)}{3-1} =$$
$$= -28$$

(b) Calculate the sample means, variances, covariance of $\mathbf{B}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$ using (3-36). Compare the results in (a) and (b).

First, compute $\bar{\mathbf{x}}$ and $\mathbf{S}_{\mathbf{X}}$

$$\bar{\mathbf{x}} = \frac{1}{n}\mathbf{X}'\mathbf{1}_n = \frac{1}{3}\mathbf{X}'\mathbf{1}_3 = \frac{1}{3}\begin{bmatrix} 9 & 5 & 1 \\ 1 & 3 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 15 \\ 6 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

The deviance is

$$\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}' = \mathbf{X} - \mathbf{1}_3\bar{\mathbf{x}}' = \begin{bmatrix} 9 & 1 \\ 5 & 3 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 5 & 2 \end{bmatrix} =$$

$$= \begin{bmatrix} 9 & 1 \\ 5 & 3 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 5 & 2 \\ 5 & 2 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ 0 & 1 \\ -4 & 0 \end{bmatrix}$$

Using the deviance to get $\mathbf{S_X}$

$$\mathbf{S_X} = \left(\frac{1}{n-1}\right)(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')' (\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}') =$$

$$= \left(\frac{1}{3-1}\right) \begin{bmatrix} 4 & 0 & -4 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ 0 & 1 \\ -4 & 0 \end{bmatrix} =$$

$$= \left(\frac{1}{3-1}\right) \begin{bmatrix} 32 & -4 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 16 & -2 \\ -2 & 1 \end{bmatrix}$$

$$\text{sample mean} = E[\mathbf{c}'\mathbf{X}] = \mathbf{c}'E[\mathbf{X}] = \mathbf{c}'\bar{\mathbf{x}} = \begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = -5 + 4 = -1$$

$$\text{sample variance} = V[\mathbf{c}'\mathbf{X}] = \mathbf{c}'V[\mathbf{X}]\mathbf{c} = \mathbf{c}'\mathbf{S_X}\mathbf{c} =$$

$$= \begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} 16 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -20 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = 28$$

$$\text{sample mean} E[\mathbf{b}'\mathbf{X}] = \mathbf{b}'E[\mathbf{X}] = \mathbf{b}'\bar{\mathbf{x}} = \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = 10 + 6 = 16$$

$$\text{sample variance} = V[\mathbf{b}'\mathbf{X}] = \mathbf{b}'V[\mathbf{X}]\mathbf{c} = \mathbf{b}'\mathbf{S_X}\mathbf{c} =$$

$$= \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 16 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 26 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 49$$

The sample covariance between $\mathbf{b}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$

Sample covariance $= \text{Cov}(\mathbf{c}'\mathbf{X}, \mathbf{b}'\mathbf{X}) = \mathbf{c}'\text{Cov}(\mathbf{X}, \mathbf{X})\mathbf{b} = \mathbf{c}'V[\mathbf{X}]\mathbf{b} = \mathbf{c}'\mathbf{S_X}\mathbf{b} =$

$$= \begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} 16 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -20 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = -40 + 12 = -28$$

Everything computed here in part (b) agrees with the corresponding values computed in part (a) using first principles.

3.15

Repeat Exercise 3.14 using the data matrix

$$\begin{bmatrix} 1 & 4 & 3 \\ 6 & 2 & 6 \\ 8 & 3 & 3 \end{bmatrix}$$

and the linear combinations

$$\mathbf{b}'\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

$$\mathbf{c}'\mathbf{X} = \begin{bmatrix} 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

(a) Evaluate the sample means, variances, and covariance of $\mathbf{b}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$ from first principles. That is, calculate the observed values of $\mathbf{b}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$ and then use the sample mean, variance, and covariance formulas.

$$\begin{aligned}
\mathbf{b}'\mathbf{x}_1 &= 1x_{11} + 1x_{12} + 1x_{13} = 1(1) + 1(4) + 1(3) = 8 \\
\mathbf{b}'\mathbf{x}_2 &= 1x_{21} + 1x_{22} + 1x_{23} = 1(6) + 1(2) + 1(6) = 14 \\
\mathbf{b}'\mathbf{x}_3 &= 1x_{31} + 1x_{32} + 1x_{33} = 1(8) + 1(3) + 1(3) = 14
\end{aligned}$$

$$\text{Sample mean} = \frac{(8 + 14 + 14)}{3} = 12$$

$$\text{Sample variance} = \frac{(8-12)^2 + (14-12)^2 + (14-12)^2}{3-1} = 12$$

$$\begin{aligned}
\mathbf{c}'\mathbf{x}_1 &= 1x_{11} + 2x_{12} - 3x_{13} = 1(1) + 2(4) - 3(3) = 0 \\
\mathbf{c}'\mathbf{x}_2 &= 1x_{21} + 2x_{22} - 3x_{23} = 1(6) + 2(2) - 3(6) = -8 \\
\mathbf{c}'\mathbf{x}_3 &= 1x_{31} + 2x_{32} - 3x_{33} = 1(8) + 2(3) - 3(3) = 5
\end{aligned}$$

$$\text{Sample mean} = \frac{(0 - 8 + 5)}{3} = -1$$

$$\text{Sample variance} = \frac{(0+1)^2 + (-8+1)^2 + (5+1)^2}{3-1} = 43$$

The covariance between $\mathbf{b}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$

$$\text{Sample covariance} =$$

$$= \frac{(8-12)(0+1) + (14-12)(-8+1) + (14-12)(5+1)}{3-1} = $$
$$= -3$$

(b) Calculate the sample means, variances, covariance of $\mathbf{B}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$ using (3-36). Compare the results in (a) and (b).

First, compute $\bar{\mathbf{x}}$ and $\mathbf{S}_\mathbf{X}$

$$\bar{\mathbf{x}} = \frac{1}{n}\mathbf{X}'\mathbf{1}_n = \frac{1}{3}\mathbf{X}'\mathbf{1}_3 = \frac{1}{3}\begin{bmatrix} 1 & 6 & 8 \\ 4 & 2 & 3 \\ 3 & 6 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 15 \\ 9 \\ 12 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix}$$

The deviance is

$$\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}' = \mathbf{X} - \mathbf{1}_3\bar{\mathbf{x}}' = \begin{bmatrix} 1 & 4 & 3 \\ 6 & 2 & 6 \\ 8 & 3 & 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 5 & 3 & 4 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 4 & 3 \\ 6 & 2 & 6 \\ 8 & 3 & 3 \end{bmatrix} - \begin{bmatrix} 5 & 3 & 4 \\ 5 & 3 & 4 \\ 5 & 3 & 4 \end{bmatrix} = \begin{bmatrix} -4 & 1 & -1 \\ 1 & -1 & 2 \\ 3 & 0 & -1 \end{bmatrix}$$

Using the deviance to get $\mathbf{S_X}$

$$\mathbf{S_X} = \left(\frac{1}{n-1}\right) (\mathbf{X} - \mathbf{1}_n \bar{\mathbf{x}}')' (\mathbf{X} - \mathbf{1}_n \bar{\mathbf{x}}') =$$

$$= \left(\frac{1}{3-1}\right) \begin{bmatrix} -4 & 1 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} -4 & 1 & -1 \\ 1 & -1 & 2 \\ 3 & 0 & -1 \end{bmatrix} =$$

$$= \left(\frac{1}{3-1}\right) \begin{bmatrix} 26 & -5 & 3 \\ -5 & 2 & -3 \\ 3 & -3 & 6 \end{bmatrix} = \begin{bmatrix} 13 & -5/2 & 3/2 \\ -5/2 & 1 & -3/2 \\ 3/2 & -3/2 & 3 \end{bmatrix}$$

$$\text{sample mean} = E[\mathbf{b}'\mathbf{X}] = \mathbf{b}'E[\mathbf{X}] = \mathbf{b}'\bar{\mathbf{x}} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix} = 5+3+4 = 12$$

$$\text{sample variance} = V[\mathbf{b}'\mathbf{X}] = \mathbf{b}'V[\mathbf{X}]\mathbf{c} = \mathbf{b}'\mathbf{S_X}\mathbf{c} =$$

$$= \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 13 & -5/2 & 3/2 \\ -5/2 & 1 & -3/2 \\ 3/2 & -3/2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 12 & -3 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 12$$

$$\text{sample mean} = E[\mathbf{c}'\mathbf{X}] = \mathbf{c}'E[\mathbf{X}] = \mathbf{c}'\bar{\mathbf{x}} = \begin{bmatrix} 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix} =$$

$$= 5 + 6 - 12 = -1$$

$$\text{sample variance} = V[\mathbf{c}'\mathbf{X}] = \mathbf{c}'V[\mathbf{X}]\mathbf{c} = \mathbf{c}'\mathbf{S_X}\mathbf{c} =$$

$$= \begin{bmatrix} 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 13 & -5/2 & 3/2 \\ -5/2 & 1 & -3/2 \\ 3/2 & -3/2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 7/2 & 4 & 21/5 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 43$$

The sample covariance between $\mathbf{b}'\mathbf{X}$ and $\mathbf{c}'\mathbf{X}$

$$\text{Sample covariance} = \mathrm{Cov}(\mathbf{c}'\mathbf{X}, \mathbf{b}'\mathbf{X}) = \mathbf{c}'\mathrm{Cov}(\mathbf{X}, \mathbf{X})\mathbf{b} = \mathbf{c}'V[\mathbf{X}]\mathbf{b} =$$

$$= \mathbf{c}'\mathbf{S_X}\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 13 & -5/2 & 3/2 \\ -5/2 & 1 & -3/2 \\ 3/2 & -3/2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} =$$

$$= \begin{bmatrix} 12 & -3 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} = 12 - 6 - 9 = -3$$

Everything computed here in part (b) agrees with the corresponding values computed in part (a) using first principles.

3.16

Let $\mathbf{V}$ be a vector random variable with mean vector $E(\mathbf{V}) = \boldsymbol{\mu}_{\mathbf{V}}$ and covariance matrix $E(\mathbf{V} - \boldsymbol{\mu}_{\mathbf{V}})(\mathbf{V} - \boldsymbol{\mu}_{\mathbf{V}})' = \boldsymbol{\Sigma}_{\mathbf{V}}$. Show that $E(\mathbf{V}\mathbf{V}') = \boldsymbol{\Sigma}_{\mathbf{V}} + \boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}'$.

$$\begin{aligned}
\boldsymbol{\Sigma}_{\mathbf{V}} = E(\mathbf{V} - \boldsymbol{\mu}_{\mathbf{V}})(\mathbf{V} - \boldsymbol{\mu}_{\mathbf{V}})' = E(\mathbf{V} - \boldsymbol{\mu}_{\mathbf{V}})(\mathbf{V}' - \boldsymbol{\mu}_{\mathbf{V}}') = \\
= E\left[\mathbf{V}\mathbf{V}' - \mathbf{V}\boldsymbol{\mu}_{\mathbf{V}}' - \boldsymbol{\mu}_{\mathbf{V}}\mathbf{V}' + \boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}'\right] = \\
= E\left[\mathbf{V}\mathbf{V}'\right] - E\left[\mathbf{V}\boldsymbol{\mu}_{\mathbf{V}}'\right] - E\left[\boldsymbol{\mu}_{\mathbf{V}}\mathbf{V}'\right] + E\left[\boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}'\right] = \\
= E\left[\mathbf{V}\mathbf{V}'\right] - E\left[\mathbf{V}\right]\boldsymbol{\mu}_{\mathbf{V}}' - \boldsymbol{\mu}_{\mathbf{V}}E\left[\mathbf{V}'\right] + \boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}' = \\
= E\left[\mathbf{V}\mathbf{V}'\right] - \boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}' - \boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}' + \boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}' = \\
= E\left[\mathbf{V}\mathbf{V}'\right] - \boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}' = \boldsymbol{\Sigma}_{\mathbf{V}} \\
\Rightarrow E\left[\mathbf{V}\mathbf{V}'\right] = \boldsymbol{\Sigma}_{\mathbf{V}} + \boldsymbol{\mu}_{\mathbf{V}}\boldsymbol{\mu}_{\mathbf{V}}'
\end{aligned}$$

3.17

Show that if $\underset{(p\times 1)}{\mathbf{X}}$ and $\underset{(q\times 1)}{\mathbf{Z}}$ are independent, then each component of $\mathbf{X}$ is independent of each component of $\mathbf{Z}$.
Hint:

$$\begin{aligned}
&P[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_p \leq x_p \text{ and } Z_1 \leq z_1, Z_2 \leq z_2, \ldots, Z_q \leq z_q] \\
&\qquad P[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_p \leq x_p] \cdot P[Z_1 \leq z_1, Z_2 \leq z_2, \ldots, Z_q \leq z_q]
\end{aligned}$$

by independence. Let $x_2, \ldots, x_p$ and $z_2, \ldots, z_q$ tend to infinity, to obtain

$$P[X_1 \leq x_1 \text{ and } Z_1 \leq Z_1] = P[X_1 \leq x_1] \cdot P[Z_1 \leq Z_1]$$

for all x_1, z_1. So X_1 and Z_1 are independent. Repeat for all other pairs.

We know $\mathbf{X}$ and $\mathbf{Z}$ are independent, so first we can break the joint multivariate CDF's into two. If we're looking at some pair of components i and j, where i is in $\mathbf{X}$ and j is in $\mathbf{Z}$, and taking the limit to ∞ for everything other than the i, j pair the limit will go to 1 for the non i, j components and we're left with the marginal CDF.

$$\lim_{\substack{i'\to\infty \\ i'\neq i}} \lim_{\substack{j'\to\infty \\ j'\neq j}} P[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_p \leq x_p \text{ and } Z_1 \leq z_1, Z_2 \leq z_2, \ldots, Z_q \leq z_q] =$$

$$= \lim_{\substack{i'\to\infty \\ i'\neq i}} P[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_p \leq x_p] \cdot \lim_{\substack{j'\to\infty \\ j'\neq j}} P[Z_1 \leq z_1, Z_2 \leq z_2, \ldots, Z_q \leq z_q] =$$

$$= P[X_i \leq x_i \text{ and } Z_j \leq z_j] = P[X_i \leq x_i] \cdot P[Z_j \leq z_j]$$

3.18

Energy consumption in 2001, by state, from the major resources

$$\begin{array}{ll} x_1 = \text{petroleum} & x_2 = \text{natural gas} \\ x_3 = \text{hydroelectric power} & x_4 = \text{nuclear electric power} \end{array}$$

is recorded in quadrillions (10^{15}) of BTUs (Source: *Statistical Abstract of the United States 2006*)
The resulting mean and covariance matrix are

$$\bar{\mathbf{x}} = \begin{bmatrix} 0.766 \\ 0.508 \\ 0.438 \\ 0.161 \end{bmatrix} \qquad \mathbf{S} = \begin{bmatrix} 0.856 & 0.635 & 0.173 & 0.096 \\ 0.635 & 0.568 & 0.128 & 0.067 \\ 0.173 & 0.127 & 0.171 & 0.039 \\ 0.096 & 0.067 & 0.039 & 0.043 \end{bmatrix}$$

(a) Using the summary statistics, determine the sample mean and variance of a state's total energy consumption for these major sources.

$$\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{b}'\mathbf{x} = X_1 + X_2 + X_3 + X_4$$

$$E[\mathbf{b}'\mathbf{x}] = \mathbf{b}'E[\mathbf{x}] = \mathbf{b}'\bar{\mathbf{x}} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0.766 \\ 0.508 \\ 0.438 \\ 0.161 \end{bmatrix} =$$

$$= 0.766 + 0.508 + 0.438 + 0.161 = 1.8730$$

$$\text{Cov}(\mathbf{b}'\mathbf{x}) = \mathbf{b}'\text{Cov}(\mathbf{x})\mathbf{b} = \mathbf{b}'\mathbf{S}\mathbf{b} =$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0.856 & 0.635 & 0.173 & 0.096 \\ 0.635 & 0.568 & 0.128 & 0.067 \\ 0.173 & 0.127 & 0.171 & 0.039 \\ 0.096 & 0.067 & 0.039 & 0.043 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1.76 & 1.3970 & 0.5110 & 0.2450 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 3.9130$$

(b) Determine the sample mean and variance of the excess of petroleum consumption over natural gas consumption. Also find the sample covariance of this variable with the total variable in part a.

$$\mathbf{c} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{c}'\mathbf{x} = X_1 - X_2$$

$$E[\mathbf{c}'\mathbf{x}] = \mathbf{c}'E[\mathbf{x}] = \mathbf{c}'\bar{\mathbf{x}} = \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.766 \\ 0.508 \\ 0.438 \\ 0.161 \end{bmatrix} =$$

$$= 0.766 - 0.508 + 0 + 0 = 0.2580$$

$$\mathrm{Cov}(\mathbf{b}'\mathbf{x}) = \mathbf{c}'\mathrm{Cov}(\mathbf{x})\mathbf{c} = \mathbf{c}'\mathbf{S}\mathbf{c} =$$

$$= \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.856 & 0.635 & 0.173 & 0.096 \\ 0.635 & 0.568 & 0.128 & 0.067 \\ 0.173 & 0.127 & 0.171 & 0.039 \\ 0.096 & 0.067 & 0.039 & 0.043 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} =$$

$$= \begin{bmatrix} 0.2210 & 0.0670 & 0.0450 & 0.0290 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 0.1540$$

The sample covariance between excess petroleum consumption over gas consumption with the total energy consumption is

$$\mathrm{Cov}(\mathbf{c}'\mathbf{x}, \mathbf{b}'\mathbf{x}) = \mathbf{c}'\mathrm{Cov}(\mathbf{x})\mathbf{b} = \mathbf{c}'\mathbf{S}\mathbf{b} =$$

$$= \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.856 & 0.635 & 0.173 & 0.096 \\ 0.635 & 0.568 & 0.128 & 0.067 \\ 0.173 & 0.127 & 0.171 & 0.039 \\ 0.096 & 0.067 & 0.039 & 0.043 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 0.2210 & 0.0670 & 0.0450 & 0.0290 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 0.3620$$

3.19

Using the summary statistics for the first three variables in Exercise 3.18, verify the relation

$$|\mathbf{S}| = (s_{11}s_{22}s_{33})|\mathbf{R}|$$

$$|\mathbf{S}| = \begin{vmatrix} 0.856 & 0.635 & 0.173 \\ 0.635 & 0.568 & 0.128 \\ 0.173 & 0.127 & 0.171 \end{vmatrix} =$$

$$= 0.856 \begin{vmatrix} 0.568 & 0.128 \\ 0.127 & 0.171 \end{vmatrix} - 0.635 \begin{vmatrix} 0.635 & 0.128 \\ 0.173 & 0.171 \end{vmatrix} + 0.173 \begin{vmatrix} 0.635 & 0.568 \\ 0.173 & 0.127 \end{vmatrix} =$$

$$= 0.856(0.568 * 0.171 - 0.128 * 0.127) -$$
$$0.635(0.635 * 0.171 - 0.128 * 0.173) +$$
$$0.173(0.635 * 0.127 - 0.568 * 0.173) =$$

$$= 0.01128831$$

$$\mathbf{D}^{-1/2} = \begin{bmatrix} 1/\sqrt{s_{11}} & 0 & 0 \\ 0 & 1/\sqrt{s_{22}} & 0 \\ 0 & 0 & 1/\sqrt{s_{33}} \end{bmatrix} =$$

$$= \begin{bmatrix} 1/\sqrt{0.856} & 0 & 0 \\ 0 & 1/\sqrt{0.568} & 0 \\ 0 & 0 & 1/\sqrt{0.171} \end{bmatrix} =$$

$$= \begin{bmatrix} 1.0808 & 0 & 0 \\ 0 & 1.3269 & 0 \\ 0 & 0 & 2.4183 \end{bmatrix}$$

$$\mathbf{R} = \mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2} =$$

$$= \begin{bmatrix} 1.0808 & 0 & 0 \\ 0 & 1.3269 & 0 \\ 0 & 0 & 2.4183 \end{bmatrix} \begin{bmatrix} 0.856 & 0.635 & 0.173 \\ 0.635 & 0.568 & 0.128 \\ 0.173 & 0.127 & 0.171 \end{bmatrix} \begin{bmatrix} 1.0808 & 0 & 0 \\ 0 & 1.3269 & 0 \\ 0 & 0 & 2.4183 \end{bmatrix} =$$

$$= \begin{bmatrix} 0.9252 & 0.6863 & 0.1870 \\ 0.8426 & 0.7537 & 0.1698 \\ 0.4184 & 0.3071 & 0.4135 \end{bmatrix} \begin{bmatrix} 1.0808 & 0 & 0 \\ 0 & 1.3269 & 0 \\ 0 & 0 & 2.4183 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0.9107 & 0.4522 \\ 0.9107 & 1 & 0.4107 \\ 0.4522 & 0.4075 & 1 \end{bmatrix}$$

$$|\mathbf{R}| = \begin{vmatrix} 1 & 0.9107 & 0.4522 \\ 0.9107 & 1 & 0.4107 \\ 0.4522 & 0.4075 & 1 \end{vmatrix} =$$

$$= 1 \begin{vmatrix} 1 & 0.4107 \\ 0.4075 & 1 \end{vmatrix} - 0.9107 \begin{vmatrix} 0.9107 & 0.4107 \\ 0.4522 & 1 \end{vmatrix} + 0.4522 \begin{vmatrix} 0.9107 & 1 \\ 0.4522 & 0.4075 \end{vmatrix} =$$

$$= 1(1 * 1 - 0.4107 * 0.4075) -$$
$$0.9107(0.9107 * 1 - 0.4107 * 0.4522) +$$
$$0.4522(0.9107 * 0.4075 - 1 * 0.4522) =$$

$$= 0.13577$$

$(s_{11}s_{22}s_{33})|\mathbf{R}| = (0.856 * 0.568 * 0.171)0.135770.083141568 * 0.13577 = 0.011288$

Okay, so we have that $|\mathbf{S}| = (s_{11}s_{22}s_{33})|\mathbf{R}|$.

3.20

In northern climates, roads must be cleared of snow quickly following a storm. One measure of storm severity is x_1 = its duration in hours, while the effectiveness of snow removal can be quantified by x_2 = the number of hours crews, men, and machine, spend to clear snow. Here are the results for 25 incidents in Wisconsin.

Table 3.2 Snow Data

x_1	x_2	x_3	x_4	x_5	x_6
12.5	13.7	9.0	24.4	3.5	26.1
14.5	16.5	6.5	18.2	8.0	14.5
8.0	17.4	10.5	22.0	17.5	42.3
9.0	11.0	10.0	32.5	10.5	17.5
19.5	23.6	4.5	18.7	12.0	21.8
8.0	13.2	7.0	15.8	6.0	10.4
9.0	32.1	8.5	15.6	13.0	25.6
7.0	12.3	6.5	12.0		
7.0	11.8	8.0	12.8		

(a) Find the sample mean and variance of the difference $x_2 - x_1$ by first obtaining the summary statistics.

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 9.42 \\ 19.272 \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} 14.139167 & 13.472667 \\ 13.472667 & 62.238767 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{b}'\mathbf{X} = X_1 - X_2$$

$$E[\mathbf{b}'\mathbf{X}] = \mathbf{b}'E[\mathbf{X}] = \mathbf{b}'\bar{\mathbf{x}} = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 9.42 \\ 19.272 \end{bmatrix} = -9.852$$

$$\mathrm{Cov}(\mathbf{b}'\mathbf{X}) = \mathbf{b}'\mathrm{Cov}(\mathbf{X})\mathbf{b} =$$

$$= \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 14.139167 & 13.472667 \\ 13.472667 & 62.238767 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} =$$

$$= \begin{bmatrix} 0.6665 & -84.7661 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 49.43260$$

(b) Obtain the mean and variance by first obtaining the individual values $x_{j2} - x_{j1}$ for $j = 1, 2, \ldots, 25$ and then calculating the mean and variance.

Compare these values with those obtained in part a.

$$\mathbf{Y} = \mathbf{X}\mathbf{b} = \begin{bmatrix} 12.5 & 13.7 \\ \vdots & \vdots \\ 13.0 & 25.6 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1.2 \\ \vdots \\ -12.6 \end{bmatrix}$$

$$E[\mathbf{Y}] = \frac{-1.2 + \cdots - 12.6}{25} = -9.852$$

$$\text{Cov}(\mathbf{Y}) = V[\mathbf{Y}] = \frac{(-1.2 + 9.852)^2 + \cdots + (-12.6 + 9.852)^2}{25 - 1} = 49.43260$$

These values are the same as those computed in part (a).

4 Chapter 4

4.1

Consider a bivariate normal distribution with $\mu_1 = 1$, $\mu_2 = 3$, $\sigma_{11} = 2$, $\sigma_{22} = 1$ and $\rho_{12} = -0.8$.

(a) Write out the bivariate normal density. For

$$\mathbf{x} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sqrt{\sigma_{11}}\sqrt{\sigma_{22}}\rho_{12} \\ \sqrt{\sigma_{11}}\sqrt{\sigma_{22}}\rho_{12} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} 2 & -0.80\sqrt{2} \\ -0.80\sqrt{2} & 1 \end{bmatrix}$$

$$f(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} e^{\frac{-1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}(\mathbf{x}-\boldsymbol{\mu})} =$$

$$= \frac{1}{(2\pi)^{2/2}\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \exp\left\{ \frac{-1}{2(1-\rho_{12}^2)} \left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 \right.\right.$$
$$\left.\left. - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right) \right] \right\} =$$

$$= \frac{1}{(2\pi)^{2/2}\sqrt{2(1-0.64)}} \exp\left\{ \frac{-1}{2(1-0.64)} \left[\left(\frac{x_1-1}{\sqrt{2}}\right)^2 + \left(\frac{x_2-3}{\sqrt{1}}\right)^2 \right.\right.$$
$$\left.\left. + 2(0.80)\left(\frac{x_1-1}{\sqrt{2}}\right)\left(\frac{x_2-3}{\sqrt{1}}\right) \right] \right\} =$$

$$= \frac{1}{1.2\pi\sqrt{2}} \exp\left\{ \frac{-1}{0.72} \left[\left(\frac{x_1-1}{\sqrt{2}}\right)^2 + (x_2-3)^2 \right.\right.$$
$$\left.\left. + 1.6\left(\frac{x_1-1}{\sqrt{2}}\right)(x_2-3) \right] \right\}$$

(b) Write out the squared statistical distance expression $(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$ as a quadratic function of x_1 and x_2.

This is most of what's inside the exponent.

$$(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) =$$

$$= \frac{1}{1-\rho_{12}^2}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right) \right] =$$

$$= \frac{1}{0.36}\left[\left(\frac{x_1-1}{\sqrt{2}}\right)^2 + \left(\frac{x_2-3}{1}\right)^2 + 1.6\left(\frac{x_1-1}{\sqrt{2}}\right)(x_2-3)\right] =$$

$$= \frac{1}{0.36}\left[\frac{1}{2}({x_1}^2-2x_1+1)+({x_2}^2-6x_2+9)+\frac{1.6\sqrt{2}}{2}(x_1x_2-3x_1-x_2+3)\right] =$$

$$= \frac{1}{0.36}\left[\frac{1}{2}({x_1}^2-2x_1+1)+({x_2}^2-6x_2+9)+\frac{1.6\sqrt{2}}{2}(x_1x_2-3x_1-x_2+3)\right] =$$

$$= \frac{25}{18}{x_1}^2+\frac{50}{18}{x_2}^2-\frac{5(5+12\sqrt{2})}{9}x_1-x_2\frac{10(15+2\sqrt{2})}{9}+\frac{20\sqrt{2}}{9}x_1x_2+\frac{5(95+24\sqrt{2})}{18} =$$

$$= 1.3889x_1^2 + 2.7778x_2^2 - 12.2059x_1 - 19.8094x_2 + 3.1427x_1x_2 + 35.8170$$

4.2

Consider a bivariate normal distribution with $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_{11} = 2$, $\sigma_{22} = 1$ and $\rho_{12} = 0.5$.

(a) Write out the bivariate normal density. For

$$\mathbf{x} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sqrt{\sigma_{11}}\sqrt{\sigma_{22}}\rho_{12} \\ \sqrt{\sigma_{11}}\sqrt{\sigma_{22}}\rho_{12} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} 2 & 0.50\sqrt{2} \\ 0.50\sqrt{2} & 1 \end{bmatrix}$$

$$f(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} e^{\frac{-1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}(\mathbf{x}-\boldsymbol{\mu})} =$$

$$= \frac{1}{(2\pi)^{2/2}\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \exp\left\{\frac{-1}{2(1-\rho_{12}^2)}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 \right.\right.$$
$$\left.\left. - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]\right\} =$$

$$= \frac{1}{(2\pi)^{2/2}\sqrt{2(1-0.25)}} \exp\left\{\frac{-1}{2(1-0.25)}\left[\left(\frac{x_1-0}{\sqrt{2}}\right)^2 + \left(\frac{x_2-2}{\sqrt{1}}\right)^2 \right.\right.$$
$$\left.\left. - 2(0.50)\left(\frac{x_1-0}{\sqrt{2}}\right)\left(\frac{x_2-2}{\sqrt{1}}\right)\right]\right\} =$$

$$= \frac{\sqrt{6}}{6\pi} \exp\left\{\frac{-2}{3}\left[\frac{x_1^2}{2} + (x_2-2)^2 - \left(\frac{x_1}{\sqrt{2}}\right)(x_2-2)\right]\right\}$$

(b) Write out the squared statistical distance expression $(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$ as a quadratic function of x_1 and x_2.

This is most of what's inside the exponent.

$$(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) =$$

$$= \frac{1}{1-\rho_{12}^2}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right] =$$

$$= \frac{4}{3}\left[\frac{1}{2}x_1^2 + (x_2-2)^2 + \left(\frac{x_1}{\sqrt{2}}\right)(x_2-2)\right] =$$

$$= \frac{4}{3}\left[\frac{1}{2}x_1^2 + \left(x_2^2 - 4x_2 + 4\right) + \frac{\sqrt{2}}{2}(x_1x_2 - 2x_1)\right] =$$

$$= \frac{2}{3}x_1^2 + \frac{4}{3}x_2^2 - \frac{4\sqrt{2}}{3}x_1 - \frac{16}{3}x_2 + \frac{2\sqrt{2}}{3}x_1x_2 - \frac{16}{3} =$$

$$= 0.6667x_1^2 + 1.3333x_2^2 - 1.8856x_1 - 5.3333x_2 + 0.9428x_1x_2 + 5.3333$$

(c) Determine (and sketch) the constant-density contour that contains 50% of the probability.

First, find the eigenvalues.

$$0 = |\boldsymbol{\Sigma} - \lambda\mathbf{I}| = \begin{vmatrix} 2-\lambda & \sqrt{2}/2 \\ \sqrt{2}/2 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda) - \frac{2}{4} =$$

$$= \lambda^2 - 3\lambda + 2 - \frac{1}{2} = \lambda^2 - 3\lambda + \frac{3}{2} = (\lambda - (3-\sqrt{3})/2)(\lambda - (3+\sqrt{3})/2)$$

The eigenvalues are $(3 \pm \sqrt{3})/2$.

$\underline{\lambda_1 = \frac{3-\sqrt{3}}{2}:}$

$$\boldsymbol{\Sigma}\mathbf{x}_1 = \lambda_1\mathbf{x}_1 \Rightarrow \begin{bmatrix} 2 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (3-\sqrt{3})/2x_1 \\ (3-\sqrt{3})/2x_2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} (1+\sqrt{3})/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & (-1+\sqrt{3})/2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} (1+\sqrt{3})/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & (-1+\sqrt{3})/2 \end{bmatrix} \xrightarrow{\text{Row } 2-\left(\frac{2}{1+\sqrt{3}}\right)\frac{\sqrt{2}}{2}\text{Row } 1} \begin{bmatrix} (1+\sqrt{3})/2 & \\ 0 & 0 \end{bmatrix}$$

$$\frac{(1+\sqrt{3})}{2}x_1 = -\frac{\sqrt{2}}{2}x_2 \Rightarrow x_1 = -\frac{\sqrt{2}}{(1+\sqrt{3})}x_2$$

$$\mathbf{x}_1 = \begin{bmatrix} -\frac{\sqrt{2}}{(1+\sqrt{3})} \\ 1 \end{bmatrix}$$

$$\|\mathbf{x}_1\| = \sqrt{\left(-\frac{\sqrt{2}}{(1+\sqrt{3})}\right)^2 + 1^2} = \sqrt{\frac{2+(4+2\sqrt{3})}{(1+\sqrt{3})^2}} = \frac{\sqrt{2(3+\sqrt{3})}}{(1+\sqrt{3})}$$

$$\mathbf{e}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \frac{(1+\sqrt{3})}{\sqrt{2(3+\sqrt{3})}} \begin{bmatrix} -\frac{\sqrt{2}}{(1+\sqrt{3})} \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{3+\sqrt{3}}} \frac{\sqrt{3-\sqrt{3}}}{\sqrt{3-\sqrt{3}}} \\ \frac{1+\sqrt{3}}{\sqrt{2(3+\sqrt{3})}} \frac{\sqrt{3-\sqrt{3}}}{\sqrt{3-\sqrt{3}}} \end{bmatrix} =$$

$$= \begin{bmatrix} -\frac{\sqrt{3-\sqrt{3}}}{\sqrt{6}} \frac{\sqrt{6}}{\sqrt{6}} \\ \frac{\sqrt{2(3+\sqrt{3})}}{\sqrt{12}} \frac{\sqrt{6}}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{6(3-\sqrt{3})}}{6} \\ \frac{\sqrt{6(3+\sqrt{3})}}{6} \end{bmatrix}$$

$\underline{\lambda_2 = \frac{3+\sqrt{3}}{2}:}$

$$\mathbf{\Sigma x}_2 = \lambda_2 \mathbf{x}_2 \Rightarrow \begin{bmatrix} 2 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (3+\sqrt{3})/2x_1 \\ (3+\sqrt{3})/2x_2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} (1-\sqrt{3})/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & (-1-\sqrt{3})/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} (1-\sqrt{3})/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & (-1-\sqrt{3})/2 \end{bmatrix} \xrightarrow{\text{Row } 2 - \left(\frac{2}{1-\sqrt{3}}\right)\frac{\sqrt{2}}{2}\text{Row } 1} \begin{bmatrix} (1-\sqrt{3})/2 & \\ 0 & 0 \end{bmatrix}$$

$$\frac{(1-\sqrt{3})}{2}x_1 = -\frac{\sqrt{2}}{2}x_2 \Rightarrow x_1 = -\frac{\sqrt{2}}{(1-\sqrt{3})}x_2$$

$$\mathbf{x}_2 = \begin{bmatrix} -\frac{\sqrt{2}}{(1-\sqrt{3})} \\ 1 \end{bmatrix}$$

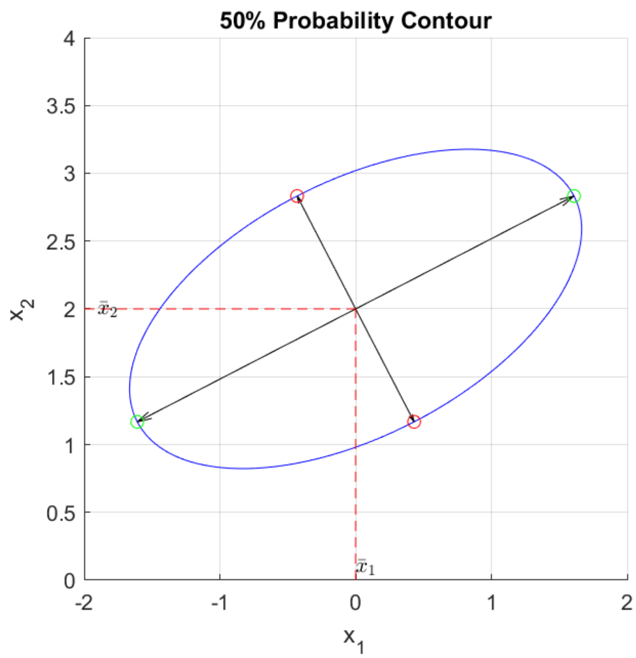
$$\|\mathbf{x}_2\| = \sqrt{\left(-\frac{\sqrt{2}}{(1-\sqrt{3})}\right)^2 + 1^2} = \sqrt{\frac{2+(4-2\sqrt{3})}{(1-\sqrt{3})^2}} = \frac{\sqrt{2(3-\sqrt{3})}}{(1-\sqrt{3})}$$

$$\mathbf{e}_2 = \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} = \frac{(1-\sqrt{3})}{\sqrt{2(3-\sqrt{3})}} \begin{bmatrix} -\frac{\sqrt{2}}{(1-\sqrt{3})} \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{3-\sqrt{3}}} \frac{\sqrt{3+\sqrt{3}}}{\sqrt{3+\sqrt{3}}} \\ \frac{1-\sqrt{3}}{\sqrt{2(3-\sqrt{3})}} \frac{\sqrt{3+\sqrt{3}}}{\sqrt{3-\sqrt{3}}} \end{bmatrix} =$$

$$= \begin{bmatrix} -\frac{\sqrt{3+\sqrt{3}}}{\sqrt{6}} \frac{\sqrt{6}}{\sqrt{6}} \\ -\frac{\sqrt{2(3-\sqrt{3})}}{\sqrt{12}} \frac{\sqrt{6}}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{6(3+\sqrt{3})}}{6} \\ -\frac{\sqrt{6(3-\sqrt{3})}}{6} \end{bmatrix}$$

(d) Determine (and sketch) the constant-density contour that contains 50% of the probability.

The $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution assigns probability 0.5 to the solid ellipsoid $\{\mathbf{x} : (\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \leq \chi^2(0.50)\} = F_2^{-1}(0.50)$, which happens at 1.3863, so $c = \sqrt{1.3863}$. The major axis has length $c\sqrt{\lambda_2} = \sqrt{1.3863 * \frac{3+\sqrt{3}}{2}} = 1.8111$ and the minor axis has length $c\sqrt{\lambda_1} = \sqrt{1.3863 * \frac{3-\sqrt{3}}{2}} = 0.9375$.



4.3

Let $\mathbf{X}$ be $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu}' = [-3, 1, 4]$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Which of the following random variables are independent? Explain.

(a) X_1 and X_2

$\text{Cov}(X_1, X_2) = -2 \neq 0$, so no, X_1 and X_2 are not $\perp\!\!\!\perp$.

(b) X_2 and X_3

$\text{Cov}(X_2, X_3) = 0$, so yes, X_2 and X_3 are $\perp\!\!\!\perp$.

(c) (X_1, X_2) and X_3

If we partition the matrix so that column 3 with rows 1 and 2 make up $\mathbf{\Sigma}_{12}$,

$$\mathbf{\Sigma} = \left[\begin{array}{rr:r} 1 & -2 & 0 \\ -2 & 5 & 0 \\ \hdashline 0 & 0 & 2 \end{array}\right] = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

$$\mathbf{\Sigma}_{12} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Because $\mathbf{\Sigma}_{12} = \text{Cov}((X_1, X_2), X_2) = \mathbf{0}$, yes, (X_1, X_2) and X_2 are $\perp\!\!\!\perp$.

(d) $\frac{X_1+X_2}{2}$ and X_3

If we define $Y_1 = \frac{X_1+X_2}{2}$ and $Y_2 = X_3$ we could then setup $\mathbf{A}$ to be

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{AX} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} \frac{X_1+X_2}{2} \\ X_3 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$\text{Cov}(\mathbf{AX}) = \mathbf{A}\text{Cov}(\mathbf{X})\mathbf{A}' = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} =$$

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

Because $\text{Cov}(Y_1, Y_2) = 0$, yes, $\frac{X_1+X_2}{2}$ and X_3 are $\perp\!\!\!\perp$. Another way would be to partition the matrix, like in chapter 3

$$\mathbf{\Sigma} = \left[\begin{array}{rr:r} 1 & -2 & 0 \\ -2 & 5 & 0 \\ \hdashline 0 & 0 & 2 \end{array}\right] = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

Then apply a linear combination to X_1 and X_2, $Y = \mathbf{b}'\mathbf{X}^\star = \frac{X_1+X_2}{2}$, where $\mathbf{X}^\star = [X_1, X_2]'$ then

$$\mathbf{\Sigma}_{11} = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

$$\mathrm{Cov}(Y, X_3) = \mathrm{Cov}(\mathbf{b}'\mathbf{X}^\star, X_3) = \mathbf{b}'\mathrm{Cov}(\mathbf{X}^\star, X_3) = \mathbf{b}'\mathbf{\Sigma}_{12} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0$$

Because $\mathrm{Cov}(Y, X_3) = 0$, yes, $Y = \frac{X_1+X_2}{2}$ and X_3 are $\perp\!\!\!\perp$.

(e) X_2 and $X_2 - \frac{5}{2}X_1 - X_3$

If we define $Y_1 = X_2 - \frac{5}{2}X_1 - X_3$ and $Y_2 = X_2$ we could then setup $\mathbf{A}$ to be

$$\mathbf{A} = \begin{bmatrix} -\frac{5}{2} & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{AX} = \begin{bmatrix} -\frac{5}{2} & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} X_2 - \frac{5}{2}X_1 - X_3 \\ X_2 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$\mathrm{Cov}(\mathbf{AX}) = \mathbf{A}\mathrm{Cov}(\mathbf{X})\mathbf{A}' = \begin{bmatrix} -\frac{5}{2} & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} -\frac{5}{2} & 0 \\ 1 & 1 \\ -1 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} -\frac{9}{2} & 10 & -2 \\ -2 & 5 & 0 \end{bmatrix} \begin{bmatrix} -\frac{5}{2} & 0 \\ 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 93/4 & 10 \\ 10 & 5 \end{bmatrix}$$

Because $\mathrm{Cov}(Y_1, Y_2) = 10$, no, X_2 and $X_2 - \frac{5}{2}X_1 - X_3$ are not $\perp\!\!\!\perp$.

4.4

Let $\mathbf{X}$ be $N_3(\boldsymbol{\mu}, \mathbf{\Sigma})$ with $\boldsymbol{\mu} = [2, -3, 1]$ and

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

(a) Find the distribution of $3X_1 - 2X_2 + X_3$.

$$\mathbf{b} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}, \qquad \mathbf{b}'\mathbf{X} = \begin{bmatrix} 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = 3X_1 - 2X_2 + X_3$$

$$E[\mathbf{b}'\mathbf{X}] = \mathbf{b}'E[\mathbf{X}] = \mathbf{b}'\boldsymbol{\mu} = \begin{bmatrix} 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} = 13$$

$$\mathrm{Cov}(\mathbf{b}'\mathbf{X}) = \mathbf{b}'\mathrm{Cov}(\mathbf{X})\mathbf{b} = \mathbf{b}'\mathbf{\Sigma}\mathbf{b} = \begin{bmatrix} 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = 9$$

Now we have that

$$\mathbf{b}'\mathbf{X} \sim N(13, 9)$$

(b) Relabel the variables if necessary, and find a 2×1 vector $\mathbf{a}$ such that X_2 and $X_2 - \mathbf{a}'\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ are independent.

I don't see X_3 in $X_2 - \mathbf{a}'\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, so if we first partition $\mathbf{\Sigma}$ and focus on $\mathbf{\Sigma}_{11}$, that way we're working with only X_1 and X_2, also defining $\mathbf{X}^\star = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$.

$$\mathbf{\Sigma} = \left[\begin{array}{cc:c} 1 & 1 & 1 \\ 1 & 3 & 2 \\ \hdashline 1 & 2 & 2 \end{array}\right] = \left[\begin{array}{c:c} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \hdashline \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{array}\right]$$

$$X_2 - \mathbf{a}'\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = X_2 - a_1X_1 - a_2X_2 = -a_1X_1 + (1-a_2)X_2 = \begin{bmatrix} -a_1 & (1-a_2) \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

We're interested in the covariance of X_2 with this, so we can setup $\mathbf{A}$ to be

$$A = \begin{bmatrix} 0 & 1 \\ -a_1 & (1-a_2) \end{bmatrix}$$

$$\text{Cov}(\mathbf{AX}^\star) = \mathbf{A}\text{Cov}(\mathbf{X}^\star)\mathbf{A}' = \mathbf{A\Sigma}_{11}\mathbf{A}' =$$

$$= \begin{bmatrix} 0 & 1 \\ -a_1 & (1-a_2) \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 0 & -a_1 \\ 1 & (1-a_2) \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 3 \\ -a_1 & (3-a_1-3a_2) \end{bmatrix}\begin{bmatrix} 0 & -a_1 \\ 1 & (1-a_2) \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & (3-a_1-3a_2) \\ (3-a_1-3a_2) & (a_1^2 + (1-a_2)(3-a_1-3a_2)) \end{bmatrix}$$

We want to know when X_2 and $X_2 - \mathbf{a}'\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ are independent. This is when the off-diagonal elements of this covariance matrix are zero, that is, when

$$3 - a_1 - 3a_2 = 0 \Rightarrow a_1 + 3a_2 = 3$$

If we pick $a_1 = 1$, then $a_2 = \frac{2}{3}$ and $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}$. Now to check that this is correct.

$$A = \begin{bmatrix} 0 & 1 \\ -1 & (1-\frac{2}{3}) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & \frac{1}{3} \end{bmatrix}$$

$$\text{Cov}(\mathbf{AX}^\star) = \mathbf{A}\text{Cov}(\mathbf{X}^\star)\mathbf{A}' = \mathbf{A\Sigma}_{11}\mathbf{A}' =$$

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

The off-diagonal values are zero, so $\text{Cov}(X_2, X_2 - \mathbf{a}'\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}) = 0$ and are $\perp\!\!\!\perp$.

4.5

Specify the following.

(a) The conditional distribution of X_1, given that $X_2 = x_2$ for the joint distribution in Exercise 4.2.

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} \\ \rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} 2 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1 \end{bmatrix}$$

Using **Result 4.6** on page 160, the conditional distribution is normally distributed with

$$\text{Mean} = \mu_1 + \sigma_{12}\sigma_{22}^{-1}(x_2 - \mu_2) = 0 + \left(\frac{\sqrt{2}}{2}\right)\left(\frac{1}{1}\right)(x_2 - 2) = \frac{\sqrt{2}}{2}(x_2 - 2)$$

$$\text{Covariance} = \sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{21} = 2 - \left(\frac{\sqrt{2}}{2}\right)\left(\frac{1}{1}\right)\left(\frac{\sqrt{2}}{2}\right) = 2 - \frac{2}{4} = \frac{3}{2}$$

$$X_1 \Big| X_2 = x_2 \sim N\left(\frac{\sqrt{2}}{2}(x_2 - 2), \frac{3}{2}\right)$$

(b) The conditional distribution of X_2, given that $X_1 = x_1$ and $X_3 = x_3$ for the joint distribution in Exercise 4.3.

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

This time we need to rearrange the rows and columns so that X_2 is the first row and the first column. Then, from **Result 4.4** on page 158, the subsets of a normal are also normal, and again, using **Result 4.6** to find the mean and variance of the normal for the conditional distribution.

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_2 \\ \hdashline \mu_1 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 1 \\ \hdashline -3 \\ 4 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \hdashline \boldsymbol{\mu}_2 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} \xrightarrow[\text{Swap column 1 and column 2}]{\text{Swap row 1 and row 2}} \left[\begin{array}{c:cc} 5 & -2 & 0 \\ \hdashline -2 & 1 & 0 \\ 0 & 0 & 2 \end{array}\right] = \left[\begin{array}{c:c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hdashline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]$$

The conditional distribution is normally distributed with

$$\text{Mean} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) = 1 + \begin{bmatrix} -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} - \begin{bmatrix} 1 \\ 4 \end{bmatrix} \right) =$$

$$= 1 + \begin{bmatrix} -2 & 0 \end{bmatrix} \begin{bmatrix} (x_1 - 1) \\ (x_3 - 4) \end{bmatrix} = 1 - 2\,(x_1 - 1) = 3 - 2x_1$$

$$\text{Covariance} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} = 5 - \begin{bmatrix} -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix} =$$

$$= 5 - \begin{bmatrix} -2 & 0 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix} = 5 - 4 = 1$$

$$X_2 \Big| X_1 = x_1, X_3 = x_3 \sim N\Big((3 - 2x_1), 1\Big)$$

(c) The conditional distribution of X_3, given that $X_1 = x_1$ and $X_2 = x_2$ for the joint distribution in Exercise 4.4.

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

This time, again, we need to rearrange the rows and columns of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. For this conditional distribution we need to partition so that X_3 is the first row and the first column. From **Result 4.4**, the subsets of the normal are also normal, and using **Result 4.6** to get the conditional distribution.

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_3 \\ \hdashline \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \hdashline 2 \\ -3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \hdashline \boldsymbol{\mu}_2 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \xrightarrow[\text{Move column 3 to column 1}]{\text{Move row 3 and row 1}} \begin{bmatrix} \sigma_{33} & \sigma_{31} & \sigma_{32} \\ \sigma_{13} & \sigma_{11} & \sigma_{12} \\ \sigma_{23} & \sigma_{21} & \sigma_{22} \end{bmatrix} =$$

$$= \left[\begin{array}{c:cc} 2 & 1 & 2 \\ \hdashline 1 & 1 & 1 \\ 2 & 1 & 3 \end{array}\right] = \left[\begin{array}{c:c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hdashline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]$$

The conditional distribution is normally distributed with

$$\text{Mean} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) = 1 + \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 2 \\ -3 \end{bmatrix} \right) =$$

$$= 1 + \begin{bmatrix} 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} (x_1 - 2) \\ (x_2 + 3) \end{bmatrix} = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{3}{2}$$

$$\text{Covariance} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} = 2 - \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} =$$

$$= 2 - \begin{bmatrix} 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2 - \frac{3}{2} = \frac{1}{2}$$

$$X_3 \Big| X_1 = x_1, X_2 = x_2 \sim N\left(\frac{1}{2}(x_1 + x_2 + 3), \frac{1}{2}\right)$$

4.6

Let $\mathbf{X}$ be distributed as $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}' = [1, -1, 2]$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

Which of the following random variables are independent? Explain.

(a) X_1 and X_2

Picking out the value of $\text{Cov}(X_1, X_2) = \sigma_{12} = \sigma_{21}$ from $\boldsymbol{\Sigma}$, we can see that $\text{Cov}(X_1, X_2) = 0$, so by part **(a)** of **Result 4.5** on page 159 X_1 and X_2 are $\perp\!\!\!\perp$.

(b) X_1 and X_3

Picking out the value of $\text{Cov}(X_1, X_3) = \sigma_{13} = \sigma_{31}$ from $\boldsymbol{\Sigma}$, we can see that $\text{Cov}(X_1, X_3) = -1$, so by part **(a)** of **Result 4.5** on page 159 X_1 and X_3 are not $\perp\!\!\!\perp$.

(c) X_2 and X_3

Picking out the value of $\text{Cov}(X_2, X_3) = \sigma_{23} = \sigma_{32}$ from $\boldsymbol{\Sigma}$, we can see that $\text{Cov}(X_2, X_3) = 0$, so by part **(a)** of **Result 4.5** on page 159 X_2 and X_3 are $\perp\!\!\!\perp$.

(d) (X_1, X_3) and X_2

This time partition $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ so that X_2 is the last row of $\boldsymbol{\mu}$ and the last row/column of $\boldsymbol{\Sigma}$

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_3 \\ \hdashline \mu_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ \hdashline -1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \hdashline \boldsymbol{\mu}_2 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \xrightarrow[\text{Swap row 2 and row 3}]{\text{Swap column 2 and column 3}} \left[\begin{array}{cc:c} \sigma_{11} & \sigma_{13} & \sigma_{12} \\ \sigma_{31} & \sigma_{33} & \sigma_{32} \\ \hdashline \sigma_{21} & \sigma_{23} & \sigma_{22} \end{array}\right] =$$

$$= \left[\begin{array}{cc:c} 4 & -1 & 0 \\ -1 & 2 & 0 \\ \hdashline 0 & 0 & 5 \end{array}\right] = \left[\begin{array}{c:c} \boldsymbol{\Sigma_{11}} & \boldsymbol{\Sigma_{12}} \\ \hdashline \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{array}\right]$$

Here, $\mathbf{X}_1 = \begin{bmatrix} X_1 \\ X_3 \end{bmatrix}$ and $\mathbf{X}_2 = \begin{bmatrix} X_2 \end{bmatrix}$. Picking out the value of $\text{Cov}(\mathbf{X}_1, \mathbf{X}_2) = \boldsymbol{\Sigma}_{12} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0}$ from the partitioned $\boldsymbol{\Sigma}$, we can see that $\boldsymbol{\Sigma}_{12} = \mathbf{0}$, so by part **(b)** of **Result 4.5** on page 159 (X_1, X_3) and X_2 are $\perp\!\!\!\perp$.

(e) X_1 and $X_1 + 3X_2 - 2X_3$

For this, we have $q = 2$ linear combinations, so

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 3 & -2 \end{bmatrix}$$

From **Result 4.3** on page 157, this would be distributed as $N_q(, (\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$. Computing $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'$ we have

$$\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}' = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 3 & -2 \end{bmatrix} \begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 3 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 4 & 0 & -1 \\ 6 & 15 & -5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 3 \\ 0 & -2 \end{bmatrix} =$$

$$= \begin{bmatrix} 4 & 6 \\ 6 & 61 \end{bmatrix}$$

This is the variance-covariance matrix of X_1 and $X_1+3X_2-2X_3$. Picking out the off-diagonal term, $\text{Cov}(X_1, X_1 + 3X_2 - 2X_3) = 6$. This value is not zero, so by **Result 4.5** part **(a)**, X_1 and $X_1 + 3X_2 - 2X_3$ are not $\perp\!\!\!\perp$.

4.7

Refer to Exercise 4.6 and specify each of the following.

(a) The conditional distribution of X_1, given that $X_3 = x_3$.

We only care about X_1 and X_3, so to remove X_2 we could first apply the linear combinations

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

so by **Result 4.3**, this would be distributed as $N_2(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$, where

$$\mathbf{A}\boldsymbol{\mu} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_3 \end{bmatrix} \quad \text{and} \quad \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}' = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{13} \\ \sigma_{31} & \sigma_{33} \end{bmatrix}$$

From here we can use **Result 4.6** to get the conditional distribution

$$\text{Mean} = \mu_1 + \sigma_{13}\sigma_{33}^{-1}(x_3 - \mu_3) = 1 + (-1)(1/2)(x_3 - 2) = 2 - \frac{1}{2}x_3$$

$$\text{Covariance} = \sigma_{11} + \sigma_{13}\sigma_{33}^{-1}\sigma_{31} = 4 - (-1)^2(1/2) - 4 - 1/2 = \frac{7}{2}$$

and putting it together we have

$$X_1 \Big| X_3 = x_3 \sim N\left(2 - \frac{1}{2}x_3, \frac{7}{2}\right)$$

(b) The conditional distribution of X_1, given that $X_2 = x_2$ and $X_3 = x_3$.

This time partition things so that one group contains X_1 and another group contains X_2 and X_3.

$$\boldsymbol{\mu} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \left[\begin{array}{r} 1 \\ \hdashline -1 \\ 2 \end{array}\right] = \left[\begin{array}{c} \boldsymbol{\mu}_1 \\ \hdashline \boldsymbol{\mu}_2 \end{array}\right]$$

$$\boldsymbol{\mu} = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix} = \left[\begin{array}{r:rr} 4 & 0 & -1 \\ \hdashline 0 & 5 & 0 \\ -1 & 0 & 2 \end{array}\right] = \left[\begin{array}{c:c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hdashline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]$$

Again, apply **Result 4.6**,

$$\text{Mean} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) = 1 + \begin{bmatrix} 0 & -1 \end{bmatrix} \begin{bmatrix} 2/10 & 0 \\ 0 & 5/10 \end{bmatrix} \left(\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right) =$$

$$= 1 + \begin{bmatrix} 0 & -1/2 \end{bmatrix} \begin{bmatrix} x_2 + 1 \\ x_3 - 2 \end{bmatrix} = 1 - \frac{1}{2}(x_3 - 2) = 2 - \frac{1}{2}x_3$$

$$\text{Covariance} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{21} = 4 - \begin{bmatrix} 0 & -1 \end{bmatrix} \begin{bmatrix} 2/10 & 0 \\ 0 & 5/10 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} =$$

$$= 4 - \begin{bmatrix} 0 & -1/2 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 4 - \frac{1}{2} = \frac{7}{2}$$

Putting this together we have

$$X_1 \Big| X_2, X_3 \sim \left(2 - \frac{1}{2}x_3, \frac{7}{2}\right)$$

4.8

(Example of a nonnormal bivariate distribution with normal marginals.) Let X_1 be $N(0,1)$ and let

$$X_2 = \begin{cases} -X_1 & \text{if } -1 \le X_1 \le 1 \\ X_1 & \text{otherwise} \end{cases}$$

Show each of the following.

(a) X_2 also has an $N(0,1)$ distribution.

The hint is basically the answer. Using symmetry we can start with

$$P(-1 < X_1 \le x) = P(-x \le X_1 < 1)$$

Using the definition of the CDF for X_2

$$\begin{aligned}
F_{X_2}(x_2) &= \\
= P(X_2 \le x_2) &= \\
= P(X_2 \le -1) + P(-1 < X_2 \le x_2) &= \\
= P(X_1 \le -1) + P(-1 < -X_1 \le x_2) &= && \text{Using the definition of } X_2. \\
= P(X_1 \le -1) + P(-x_2 \le X_1 < 1) &= \\
= P(X_1 \le -1) + P(-1 < -X_1 \le x_2) &= && \text{Using symmetry argument.} \\
P(X_1 \le x_2) &= && \text{CDF definition for } X_1. \\
F_{X_1}(x_2)
\end{aligned}$$

The CDF for X_1 and X_2 are the same, so since $X_1 \sim N(0,1)$, then X_2 is also distributed as $N(0,1)$.

(b) X_1 and X_2 do *not* have a bivariate normal distribution.

$$\mathbf{a} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \qquad Y = \mathbf{a}'\mathbf{X} = X_1 - X_2$$

$$\mathbf{a}'\mathbf{X} = X_1 - X_2 = \begin{cases} X_1 - (-X_1) & \text{if } -1 \le X_1 \le 1 \\ X_1 - X_1 & \text{otherwise} \end{cases} =$$

$$= \begin{cases} 2X_1 & \text{if } -1 \le X_1 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

If X_1 is in the interval $-1 \le X_1 \le 1$ then it's distribution as $Y = (X_1 - X_2) \sim N(0,4)$, since $E[\mathbf{a}'\mathbf{X}] = E[2X_1] = 2E[X_1] = 2\mu_1 = 2 * 0 = 0$ and $V[\mathbf{a}'\mathbf{X}] = V[2X_1] = 4V[X_1] = 4\sigma_1^2 = 4 * 1 = 4$. But if X_1 is not in the interval $-1 \le X_1 \le 1$, then $X_1 - X_2 = 0$. This happens with point probability

$$P(\mathbf{a}'\mathbf{X}) = P(X_1 - X_2) = P(0) = P(X_1 < -1 \text{ and } X_1 > 1) =$$
$$= P(X_1 < -1) + P(X_1 > 1) = 2P(X_1 > 1) = P(|X_1| > 1) = 0.3173105$$

So since $P(0) \ne 0$, X_1 and X_2 don't have a bivariate normal. Basically, a Normal distribution is either entirely continuous or entirely discrete, so here we have a discrete point mass at 0 that doesn't have a probability of zero and a continuous $N(0,4)$ distribution from -2 to 2, so this is not consistent with a normal distribution.

The R code for this probability calculation is `2*pnorm(1, lower.tail=FALSE)`. Could have also done the probability calculation as

$$P(|X_1| > 1) = 1 - P(|X_1| \le 1) = 1 - P(-1 \le X_1 \le 1) =$$
$$= 1 - (\Phi(1) - \Phi(-1)) = 1 - 0.6826895 = 0.3173105$$

Where Φ is the CDF of a standard normal distribution. The R code is `1 - (pnorm(1)- pnorm(-1))`.
Hint:

(a) Since X_1 is $N(0,1)$, $P[-1 < X_1 \leq x] = P[-x \leq X_1 < 1]$ for any x. When $-1 < x_2 < 1$, $P[X_2 \leq x_2] = P[X_2 \leq -1]+P[-1 < X_2 \leq x_2] = P[X_1 \leq -1]+P[-1 < -X_1 \leq x_2] = P[X_1 \leq -1]+P[-x_2 \leq X_1 < 1]$. But $P[-x_2 \leq X_1 < 1] = P[-1 < X_1 \leq x_2]$ from the symmetry argument in the first line of this hint. Thus $P[X_2 \leq x_2] = P[X_1 \leq -1] + P[-1 < X_1 \leq x_2] = P[X_1 \leq x_2]$, which is a standard normal probability.

(b) Consider the linear combination $X_1 - X_2$, which equals zero with probability $P[|X_1| > 1] = 0.3174$.

4.9

Refer to Exercise 4.8, but modify the construction by replacing the break point 1 by c so that

$$X_2 = \begin{cases} -X_1 & \text{if } -c \leq X_1 \leq c \\ X_1 & \text{elsewhere} \end{cases}$$

Show that c can be chosen so that $\text{Cov}(X_1, X_2) = 0$, but that the two random variables are not independent.
Hint:
For $c = 0$, evaluate $\text{Cov}(X_1, X_2) = E[X_1(X_1)]$
For c very large, evaluate $\text{Cov}(X_1, X_2) \doteq E[X_1(-X_1)]$

We already know X_1 and X_2 are not $\perp\!\!\!\perp$, since X_2 is a function of X_1, so we just need to show it's possible for c to be chosen so the covariance of X_1 and X_2 will be zero. Conveniently, the hint suggests trying out both $c = 0$ and c very big, these are two extreme values that c could take on. If the covariance for the two extreme c values form an interval that contains 0 we can use the interemediate value theorem to say the covariance of X_1 and X_2 will pass through zero for some c.
First off, the covariance is

$$\text{Cov}(X_1, X_2) = E\left[(X_1 - E[X_1])\,(X_2 - E[X_2])\right] =$$

$$= E\left[X_1X_2 - X_1E[X_2] - E[X_1]X_2 + E[X_1]E[X_2]\right] =$$

$$= E[X_1X_2] - E[X_1]E[X_2] - E[X_1]E[X_2] + E[X_1]E[X_2] =$$

$$= E[X_1X_2] - E[X_1]E[X_2]$$

When $c = 0$, the probability that $X_2 = -X_1$ is

$$P(-c \leq X_1 \leq c) = P(0 \leq X_1 \leq 0) = \Phi(0) - \Phi(0) = 0$$

and so from the definition of X_2, the probability of $X_2 = -X_1$ is zero, and the probability of $X_2 = X_1$ is 1. Now to compute the covariance for $c = 0$,

$$\mathrm{Cov}(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2] = E[X_1(X_1)] - E[X_1]E[-X_1] =$$

$$= E[X_1^2] - (E[X_1])^2 = \mathrm{Var}[X_1] = 1$$

When c is very big, say ∞, the probability that $X_2 = -X_1$ is

$$P(-c \leq X_1 \leq c) = P(-\infty \leq X_1 \leq \infty) = \Phi(\infty) - \Phi(-\infty) = 1 - 0 = 1$$

and so from the definition of X_2, the probability of $X_2 = -X_1$ is 1, and the probability of $X_2 = X_1$ is 0. This is opposite what we got when c was 0. Computing the covariance,

$$\mathrm{Cov}(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2] = E[X_1(-X_1)] - E[X_1]E[-X_1] =$$

$$= -\left(E[X_1^2] - (E[X_1])^2\right) = -\mathrm{Var}[X_1] = -1$$

We now have that the covariance for extreme values of c, 1 for $c = 0$ and -1 for c at ∞. Since the covariance is a smooth function of c, then by the intermediate value theorem, $\mathrm{Cov}(X_1, X_2) = 0$ for some value of c. We're done, we showed that there exists some value of c that causes the covariance to be zero when X_1 and X_2 are not independent.

4.10

Show the following

(a)

$$\begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0}' & \mathbf{B} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{B}|$$

$$\begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0}' & \mathbf{B} \end{vmatrix} = \left| \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}' & \mathbf{B} \end{bmatrix} \right| = \begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0}' & \mathbf{I} \end{vmatrix} \begin{vmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}' & \mathbf{B} \end{vmatrix} =$$

$$= |\mathbf{AI} - \mathbf{00}'|\,|\mathbf{IB} - \mathbf{00}'| = |\mathbf{AI}|\,|\mathbf{IB}| = |\mathbf{A}|\,|\mathbf{B}|$$

(b)

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0}' & \mathbf{B} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{B}| \quad \text{for} \quad |\mathbf{A}| \neq 0$$

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0}' & \mathbf{B} \end{vmatrix} = \left| \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0}' & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{C} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \right| = \begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0}' & \mathbf{B} \end{vmatrix} \begin{vmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{C} \\ \mathbf{0}' & \mathbf{I} \end{vmatrix} =$$

$$= |\mathbf{A}|\,|\mathbf{B}|\,\left|\mathbf{II} - \mathbf{A}^{-1}\mathbf{C}\mathbf{0}'\right| = |\mathbf{A}|\,|\mathbf{B}|\,|\mathbf{I}| = |\mathbf{A}|\,|\mathbf{B}|$$

Hint:
(a) $\begin{vmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0}' & \mathbf{B}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0}' & \mathbf{I}\end{vmatrix}\begin{vmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0}' & \mathbf{B}\end{vmatrix}$. Expanding the determinant $\begin{vmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0}' & \mathbf{B}\end{vmatrix}$ by the first row (see definition 2A.24) gives 1 times a determinant of the same form, with the order of $\mathbf{I}$ reduced by one. This procedure is repeated until $1 \times |\mathbf{B}|$ is obtained. Similarly, expanding the determinant $\begin{vmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0}' & \mathbf{I}\end{vmatrix}$ by the last row gives $\begin{vmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0}' & \mathbf{B}\end{vmatrix} = |\mathbf{A}|$.
(b) $\begin{vmatrix}\mathbf{A} & \mathbf{C}\\ \mathbf{0}' & \mathbf{B}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0}' & \mathbf{B}\end{vmatrix}\begin{vmatrix}\mathbf{I} & \mathbf{A}^{-1}\mathbf{C}\\ \mathbf{0}' & \mathbf{I}\end{vmatrix}$. Expanding the determinant $\begin{vmatrix}\mathbf{I} & \mathbf{A}^{-1}\mathbf{C}\\ \mathbf{0}' & \mathbf{I}\end{vmatrix}$ by the last row gives $\begin{vmatrix}\mathbf{I} & \mathbf{A}^{-1}\mathbf{C}\\ \mathbf{0}' & \mathbf{I}\end{vmatrix} = 1$. Now use the results in Part a.

4.11

Show that, if $\mathbf{A}$ is square,

$$\begin{aligned}|\mathbf{A}| &= |\mathbf{A}_{22}|\left|\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right| \quad \text{for} \quad |\mathbf{A}_{22}| \neq 0\\ &= |\mathbf{A}_{11}|\left|\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right| \quad \text{for} \quad |\mathbf{A}_{11}| \neq 0\end{aligned}$$

First, using the hint

$$\begin{bmatrix}\mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\\ \mathbf{0}' & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I}\end{bmatrix} =$$

$$= \begin{bmatrix}\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{12} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{22}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I}\end{bmatrix} =$$

$$= \begin{bmatrix}\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{12} - \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I}\end{bmatrix} =$$

$$= \begin{bmatrix}\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I}\end{bmatrix} =$$

$$= \begin{bmatrix}\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0}\\ \mathbf{A}_{21} - \mathbf{A}_{22}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix} =$$

$$= \begin{bmatrix}\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0}\\ \mathbf{A}_{21} - \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix} =$$

$$= \begin{bmatrix}\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0}\\ \mathbf{0}' & \mathbf{A}_{22}\end{bmatrix} =$$

Now, for the determinant

$$\left|\begin{bmatrix}\mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\\ \mathbf{0}' & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I}\end{bmatrix}\right| = \left|\begin{bmatrix}\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0}\\ \mathbf{0}' & \mathbf{A}_{22}\end{bmatrix}\right|$$

$$\Rightarrow \begin{vmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{vmatrix} \begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{vmatrix} \begin{vmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{vmatrix} = \begin{vmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} \end{vmatrix}$$

$$\Rightarrow \left|\mathbf{II} - -\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{0}'\right| |\mathbf{A}| \left|\mathbf{II} - -\mathbf{0}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right| = \left|\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right| |\mathbf{A}_{22}|$$

$$\Rightarrow (1)\,|\mathbf{A}|\,(1) = |\mathbf{A}_{22}| \left|\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right|$$

$$\Rightarrow |\mathbf{A}| = |\mathbf{A}_{22}| \left|\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right|$$

Verifying the second part of the hint

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} =$$

$$= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{11} & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} =$$

$$= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} - \mathbf{A}_{21} & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} =$$

$$= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} =$$

$$= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} - \mathbf{A}_{11}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix} =$$

$$= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} - \mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix} =$$

$$= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix} =$$

$$\left| \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \right| = \left| \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix} \right|$$

$$\Rightarrow \begin{vmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I} \end{vmatrix} \begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{vmatrix} \begin{vmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{I} \end{vmatrix} = \left| \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix} \right|$$

$$\Rightarrow \left|\mathbf{II} - -\mathbf{0}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\right| |\mathbf{A}| \left|\mathbf{II} - -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{0}'\right| = |\mathbf{A}_{11}| \left|\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right|$$

$$\Rightarrow (1)\,|\mathbf{A}|\,(1) = |\mathbf{A}_{11}| \left|\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right|$$

$$\Rightarrow |\mathbf{A}| = |\mathbf{A}_{11}| \left|\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right|$$

Finally, we have what we want, that

$$\begin{aligned} |\mathbf{A}| &= |\mathbf{A}_{22}| \left|\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right| \quad \text{for} \quad |\mathbf{A}_{22}| \neq 0 \\ &= |\mathbf{A}_{11}| \left|\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right| \quad \text{for} \quad |\mathbf{A}_{11}| \neq 0 \end{aligned}$$

Hint: Partition $\mathbf{A}$ and verify that

$$\begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} \end{bmatrix}$$

Take the determinants on both sides of this inequality. Use Exercise 4.10 for the first and third determinants on the left and for the determinant on the right. The second inequality for $|\mathbf{A}|$ follows by considering

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix}$$

4.12

Show that, for $\mathbf{A}$ symmetric,

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \left(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix}$$

Thus, $\left(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1}$ is the upper left-hand block of $\mathbf{A}^{-1}$.
From Exercise 4.11 we know

$$\begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} \end{bmatrix}$$

$$\begin{aligned} \Rightarrow & \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \\ & \qquad \times \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix}^{-1} = \\ & \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix}^{-1} \end{aligned}$$

$$\Rightarrow \mathbf{A} = \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix}^{-1}$$

$$\Rightarrow \mathbf{A}^{-1} = \left(\begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix}^{-1} \right)^{-1}$$

$$\Rightarrow \mathbf{A}^{-1} = \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \left(\begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22} \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix}$$

$$\Rightarrow \mathbf{A}^{-1} = \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \begin{bmatrix} \left(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1} & \mathbf{0} \\ \mathbf{0}' & \mathbf{A}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix}$$

Hint: Premultiply the expression in the hint to Exercise 4.11 by $\begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix}$ and postmultiply by $\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix}$. Take inverses of the resulting expression.

4.13

Show the following if $\mathbf{X}$ is $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $|\boldsymbol{\Sigma}| \neq 0$.

(a) Check that $|\boldsymbol{\Sigma}| = |\boldsymbol{\Sigma}_{22}| \left|\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right|$. (Note that $|\boldsymbol{\Sigma}|$ can be factored into the product of contributions from the marginal and conditional distributions.)

This follows directly fom Exercise 4.11 since $\boldsymbol{\Sigma}$ is square.

$$\boldsymbol{\Sigma} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]$$

$$|\boldsymbol{\Sigma}| = |\boldsymbol{\Sigma}_{22}| \left|\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right|$$

(b) Check that

$$\begin{aligned} (\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) = {} & [\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)]' \\ & \times \left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1}[\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{\mathbf{12}}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)] \\ & + (\mathbf{x}_2 - \boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

(Thus the joint density exponent can be written as the sum of two terms corresponding to contributions from the conditional and marginal distributions.)

$$\begin{aligned} & (\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) = \\ & \quad \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix}' \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \\ & \qquad \times \begin{bmatrix} \mathbf{I} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix} = \end{aligned}$$

$$\begin{aligned} = & \left[(\mathbf{x}_1 - \boldsymbol{\mu}_1)' - (\mathbf{x}_2 - \boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}, \quad (\mathbf{x}_2 - \boldsymbol{\mu}_2)'\right] \\ & \times \begin{bmatrix} \left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \\ & \times \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix} = \end{aligned}$$

$$= \left[\left[\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \right]', \quad (\mathbf{x}_2 - \boldsymbol{\mu}_2)' \right]$$
$$\times \begin{bmatrix} \left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix}$$
$$\times \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix} =$$

$$= \left[\left[\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \right]', \quad (\mathbf{x}_2 - \boldsymbol{\mu}_2)' \right]$$
$$\times \begin{bmatrix} \left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} \left[\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \right] \\ \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \end{bmatrix} =$$

$$\begin{aligned} &= \left[\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \right]' \\ &\qquad \times \left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} \\ &\qquad \times \left[\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \right] \\ &\qquad + (\mathbf{x}_2 - \boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

(c) Given the results in Parts a and b, identify the marginal distribution of $\mathbf{X}_2$ and the conditional distribution of $\mathbf{X}_1 \big| \mathbf{X}_2 = \mathbf{x}_2$.

From part (a)

$$|\boldsymbol{\Sigma}| = |\boldsymbol{\Sigma}_{22}| \left| \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} \right|$$

The value $\boldsymbol{\Sigma}_{22}$ is the covariance of the marginal (normal) distribution of $\mathbf{X}_2$. The determinant, $|\boldsymbol{\Sigma}_{22}|$, is the scaling factor found in the first part of the normal density equation for the normal marginal distribution of $\mathbf{X}_2$.

The value of $\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$ is the covariance of the normal conditional distribution of $\mathbf{X}_1 \big| \mathbf{X}_2 = \mathbf{x}_2$. The determinant, $\left| \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} \right|$, is the scaling factor found in the first part of the normal density equation for the normal conditional distribution of $\mathbf{X}_1 \big| \mathbf{X}_2 = \mathbf{x}_2$.

From part (b)

$$\begin{aligned} (\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) &= \left[\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \right]' \\ &\quad \times \left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} [\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{\mathbf{12}}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)] \\ &\quad + (\mathbf{x}_2 - \boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

The left and right parts of the first half of the equation can be written as

$$\mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) = \mathbf{x}_1 - \left(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \right).$$

From there, we can see that,

$$[\mathbf{x}_1 - (\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2))]'(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})^{-1} \times [\mathbf{x}_1 - (\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{\mathbf{12}}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2))]$$

, is the exponent for the normal conditional distribution of $\mathbf{X}_1 \big| \mathbf{X}_2 = x_2$, where $\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$ is the mean from page 160 and $\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$ is the covariance from page 161.
The second part of the addition, $(\mathbf{x}_2 - \boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$, is the exponent of the normal marginal for $\mathbf{X}_2$, where the mean is $\boldsymbol{\mu}_2$ and the covariance is $\boldsymbol{\Sigma}_{22}$.

Hint:
(a) Apply Exercise 4.11
(b) Note from Exercise 4.12 that we can write $(\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})$ as

$$\begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix}' \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \mathbf{I} \end{bmatrix} \begin{bmatrix} (\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \times \begin{bmatrix} \mathbf{I} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix}$$

If we group the product so that

$$\begin{bmatrix} \mathbf{I} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix}$$

the result follows.

4.14

If $\mathbf{X}$ is distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $|\boldsymbol{\Sigma}| \neq 0$, show that the joint density can be written as the product of marginal densities for

$$\underset{(q\times 1)}{\mathbf{X}_1} \quad \text{and} \quad \underset{((p-q)\times 1)}{\mathbf{X}_2} \quad \text{if} \quad \boldsymbol{\Sigma}_{12} = \underset{(q\times(p-q))}{\mathbf{0}}$$

First, parition $\mathbf{X}$ and $\boldsymbol{\Sigma}$,

$$\underset{p\times 1}{\mathbf{X}} = \begin{bmatrix} \underset{(q\times 1)}{\mathbf{X}_1} \\ \hdashline \underset{((p-q)\times 1)}{\mathbf{X}_2} \end{bmatrix}$$

$$\underset{p\times 1}{\boldsymbol{\mu}} = \begin{bmatrix} \underset{(q\times 1)}{\boldsymbol{\mu}_1} \\ \hdashline \underset{((p-q)\times 1)}{\boldsymbol{\mu}_2} \end{bmatrix}$$

$$\underset{p\times p}{\boldsymbol{\Sigma}} = \left[\begin{array}{c:c} \underset{(q\times q)}{\boldsymbol{\Sigma}_{11}} & \underset{(q\times(p-q))}{\boldsymbol{\Sigma}_{12}} \\ \hdashline \underset{((p-q)\times q)}{\boldsymbol{\Sigma}_{21}} & \underset{((p-q)\times(p-q))}{\boldsymbol{\Sigma}_{22}} \end{array}\right] = \left[\begin{array}{c:c} \underset{(q\times q)}{\boldsymbol{\Sigma}_{11}} & \underset{(q\times(p-q))}{\mathbf{0}} \\ \hdashline \underset{((p-q)\times q)}{\mathbf{0}'} & \underset{((p-q)\times(p-q))}{\boldsymbol{\Sigma}_{22}} \end{array}\right]$$

We can pick $\boldsymbol{\Sigma}^{-1}$ and use the definition of an inverse on page 58 to show it is the inverse of $\boldsymbol{\Sigma}$, (if $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, then $\mathbf{B}$ is an inverse of $\mathbf{A}$.) So using the hint to pick $\boldsymbol{\Sigma}^{-1}$

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix}$$

$$\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} = \mathbf{I}$$

$$\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}\boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}' & \mathbf{I} \end{bmatrix} = \mathbf{I}$$

Because $\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1} = \mathbf{I}$, our pick of $\boldsymbol{\Sigma}^{-1}$ is the inverse of $\boldsymbol{\Sigma}$.
Using the second part of the hint

$$\begin{aligned} (\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) &= [(\mathbf{x}_1-\boldsymbol{\mu}_1)', (\mathbf{x}_2-\boldsymbol{\mu}_2)'] \begin{bmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1-\boldsymbol{\mu}_1 \\ \mathbf{x}_2-\boldsymbol{\mu}_2 \end{bmatrix} \\ &= (\mathbf{x}_1-\boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1) + (\mathbf{x}_2-\boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2) \end{aligned}$$

Using the third part of the hint

$$|\boldsymbol{\Sigma}| = \begin{vmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22} \end{vmatrix} = |\boldsymbol{\Sigma}_{11}|\,|\boldsymbol{\Sigma}_{22}| \text{ (from Exercise 4.10 (a))}$$

Now, using the formula for the normal density (4–4)

$$\begin{aligned} f(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\Sigma}) &= \frac{1}{(1/2)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\} \\ &= \frac{1}{(1/2)^{q/2}(1/2)^{(p-q)/2}(|\boldsymbol{\Sigma}_{11}|\,|\boldsymbol{\Sigma}_{22}|)^{1/2}} \\ &\times \exp\left\{-\frac{1}{2}(\mathbf{x}_1-\boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1) + (\mathbf{x}_2-\boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)\right\} \\ &= \left(\frac{1}{(1/2)^{q/2}|\boldsymbol{\Sigma}_{11}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_1-\boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1)\right\}\right) \\ &\times \left(\frac{1}{(1/2)^{(p-q)/2}|\boldsymbol{\Sigma}_{22}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_2-\boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)\right\}\right) \\ &= f(\mathbf{x}_1|\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1) f(\mathbf{x}_2|\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2) \end{aligned}$$

so we've shown that the joint density can be written as the product of marginal densities

$$f(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\Sigma}) = f(\mathbf{x}_1|\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1) f(\mathbf{x}_2|\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)$$

Hint: Show by block multiplication that

$$\begin{bmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \text{ is the inverse of } \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

Then write

$$\begin{aligned}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) &= [(\mathbf{x}_1-\boldsymbol{\mu}_1)', (\mathbf{x}_2-\boldsymbol{\mu}_2)'] \begin{bmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1-\boldsymbol{\mu}_1 \\ \mathbf{x}_2-\boldsymbol{\mu}_2 \end{bmatrix} \\ &= (\mathbf{x}_1-\boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1) + (\mathbf{x}_2-\boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)\end{aligned}$$

Note that $|\boldsymbol{\Sigma}| = |\boldsymbol{\Sigma}_{11}|\,|\boldsymbol{\Sigma}_{22}|$ from Exercise 4.10(a). Now factor the joint density.

4.15

Show that $\sum_{j=1}^{n}(\mathbf{x}_j-\bar{\mathbf{x}})(\bar{\mathbf{x}}-\boldsymbol{\mu})'$ and $\sum_{j=1}^{n}(\bar{\mathbf{x}}-\boldsymbol{\mu})(\mathbf{x}_j-\bar{\mathbf{x}})'$ are both $p\times p$ matrices of zeros. Here $\mathbf{x}_j' = [x_{j1}, x_{j2}, \ldots, x_{jp}]$, $j = 1, 2, \ldots, n$ and

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{x}_j$$

Here are both a short answer (using distribution and vector-vector multiplication) and long ugly answer (multiply matrix out).
<u>Short answer:</u>

$$\sum_{j=1}^{n}(\mathbf{x}_j-\bar{\mathbf{x}})(\bar{\mathbf{x}}-\boldsymbol{\mu})' =$$

$$= \sum_{j=1}^{n}(\mathbf{x}_j-\bar{\mathbf{x}})(\bar{\mathbf{x}}'-\boldsymbol{\mu}') =$$

$$= \sum_{j=1}^{n}\left(\underset{(p\times1)}{\mathbf{x}_j}\underset{(1\times p)}{\bar{\mathbf{x}}'} - \underset{(p\times1)}{\bar{\mathbf{x}}}\underset{(1\times p)}{\bar{\mathbf{x}}'} - \underset{(p\times1)}{\mathbf{x}_j}\underset{(1\times p)}{\boldsymbol{\mu}'} + \underset{(p\times1)}{\bar{\mathbf{x}}}\underset{(1\times p)}{\boldsymbol{\mu}'}\right) =$$

$$= \sum_{j=1}^{n}\left(\underset{(p\times1)}{\mathbf{x}_j}\underset{(1\times p)}{\bar{\mathbf{x}}'}\right) - n\underset{(p\times1)}{\bar{\mathbf{x}}}\underset{(1\times p)}{\bar{\mathbf{x}}'} - \sum_{j=1}^{n}\left(\underset{(p\times1)}{\mathbf{x}_j}\underset{(1\times p)}{\boldsymbol{\mu}'}\right) + n\underset{(p\times1)}{\bar{\mathbf{x}}}\underset{(1\times p)}{\boldsymbol{\mu}'} =$$

$$= n\underset{(p\times1)}{\bar{\mathbf{x}}}\underset{(1\times p)}{\bar{\mathbf{x}}'} - n\underset{(p\times1)}{\bar{\mathbf{x}}}\underset{(1\times p)}{\bar{\mathbf{x}}'} - n\underset{(p\times1)}{\bar{\mathbf{x}}}\underset{(1\times p)}{\boldsymbol{\mu}'} + n\underset{(p\times1)}{\bar{\mathbf{x}}}\underset{(1\times p)}{\boldsymbol{\mu}'} =$$

$$= \underset{p\times p}{\mathbf{0}}$$

and

$$\sum_{j=1}^{n}(\bar{\mathbf{x}}-\boldsymbol{\mu})(\mathbf{x}_j-\bar{\mathbf{x}})' =$$

$$= \sum_{j=1}^{n} (\bar{\mathbf{x}} - \boldsymbol{\mu})(\mathbf{x}_j' - \bar{\mathbf{x}}') =$$

$$= \sum_{i=1}^{n} \left(\underset{(p\times 1)}{\bar{\mathbf{x}}} \underset{(1\times p)}{\mathbf{x}_j'} - \underset{(p\times 1)}{\boldsymbol{\mu}} \underset{(1\times p)}{\mathbf{x}_j'} - \underset{(p\times 1)}{\bar{\mathbf{x}}} \underset{(1\times p)}{\bar{\mathbf{x}}'} + \underset{(p\times 1)}{\boldsymbol{\mu}} \underset{(1\times p)}{\bar{\mathbf{x}}'} \right) =$$

$$= \sum_{i=1}^{n} \left(\underset{(p\times 1)}{\bar{\mathbf{x}}} \underset{(1\times p)}{\mathbf{x}_j'} \right) - \sum_{i=1}^{n} \left(\underset{(p\times 1)}{\boldsymbol{\mu}} \underset{(1\times p)}{\mathbf{x}_j'} \right) - n \underset{(p\times 1)}{\bar{\mathbf{x}}} \underset{(1\times p)}{\bar{\mathbf{x}}'} + n \underset{(p\times 1)}{\boldsymbol{\mu}} \underset{(1\times p)}{\bar{\mathbf{x}}'} =$$

$$= n \underset{(p\times 1)}{\bar{\mathbf{x}}} \underset{(1\times p)}{\bar{\mathbf{x}}'} - n \underset{(p\times 1)}{\boldsymbol{\mu}} \underset{(1\times p)}{\bar{\mathbf{x}}'} - n \underset{(p\times 1)}{\bar{\mathbf{x}}} \underset{(1\times p)}{\bar{\mathbf{x}}'} + n \underset{(p\times 1)}{\boldsymbol{\mu}} \underset{(1\times p)}{\bar{\mathbf{x}}'} =$$

$$= \underset{p\times p}{\mathbf{0}}$$

Long answer:

$$\sum_{j=1}^{n} (\mathbf{x}_j - \bar{\mathbf{x}})(\bar{\mathbf{x}} - \boldsymbol{\mu})' =$$

$$= \sum_{j=1}^{n} \left(\begin{bmatrix} x_{j1} \\ \vdots \\ x_{jp} \end{bmatrix} - \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{bmatrix} \right) \left(\begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix} \right)' =$$

$$= \sum_{j=1}^{n} \left(\begin{bmatrix} x_{j1} - \bar{x}_1 \\ \vdots \\ x_{jp} - \bar{x}_p \end{bmatrix} \right) \left(\begin{bmatrix} \bar{x}_1 - \mu_1 \\ \vdots \\ \bar{x}_p - \mu_p \end{bmatrix} \right)' =$$

$$= \sum_{j=1}^{n} \left(\begin{bmatrix} (x_{j1} - \bar{x}_1)(\bar{x}_1 - \mu_1) & \cdots & (x_{j1} - \bar{x}_1)(\bar{x}_p - \mu_p) \\ \vdots & \ddots & \vdots \\ (x_{jp} - \bar{x}_p)(\bar{x}_1 - \mu_1) & \cdots & (x_{jp} - \bar{x}_p)(\bar{x}_p - \mu_p) \end{bmatrix} \right) =$$

$$= \sum_{j=1}^{n} \left(\begin{bmatrix} (x_{j1}\bar{x}_1 - \bar{x}_1\bar{x}_1 - x_{j1}\mu_1 + \bar{x}_1\mu_1) & \cdots & (x_{j1}\bar{x}_p - \bar{x}_1\bar{x}_p - x_{j1}\mu_p + \bar{x}_1\mu_p) \\ \vdots & \ddots & \vdots \\ (x_{jp}\bar{x}_1 - \bar{x}_p\bar{x}_1 - x_{jp}\mu_1 + \bar{x}_p\mu_1) & \cdots & (x_{jp}\bar{x}_p - \bar{x}_p\bar{x}_p - x_{jp}\mu_p + \bar{x}_p\mu_p) \end{bmatrix} \right) =$$

$$= \begin{bmatrix} \sum_{j=1}^{n} (x_{j1}\bar{x}_1 - \bar{x}_1\bar{x}_1 - x_{j1}\mu_1 + \bar{x}_1\mu_1) & \cdots & \sum_{j=1}^{n} (x_{j1}\bar{x}_p - \bar{x}_1\bar{x}_p - x_{j1}\mu_p + \bar{x}_1\mu_p) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} (x_{jp}\bar{x}_1 - \bar{x}_p\bar{x}_1 - x_{jp}\mu_1 + \bar{x}_p\mu_1) & \cdots & \sum_{j=1}^{n} (x_{jp}\bar{x}_p - \bar{x}_p\bar{x}_p - x_{jp}\mu_p + \bar{x}_p\mu_p) \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{j=1}^{n} (x_{j1}\bar{x}_1) - n\bar{x}_1\bar{x}_1 - \sum_{j=1}^{n} (x_{j1}\mu_1) + n\bar{x}_1\mu_1 & \cdots & \sum_{j=1}^{n} (x_{j1}\bar{x}_p) - n\bar{x}_1\bar{x}_p - \sum_{j=1}^{n} (x_{j1}\mu_p) + n\bar{x}_1\mu_p \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} (x_{jp}\bar{x}_1) - n\bar{x}_p\bar{x}_1 - \sum_{j=1}^{n} (x_{jp}\mu_1) + n\bar{x}_p\mu_1 & \cdots & \sum_{j=1}^{n} (x_{jp}\bar{x}_p) - n\bar{x}_p\bar{x}_p - \sum_{j=1}^{n} (x_{jp}\mu_p) + n\bar{x}_p\mu_p \end{bmatrix} =$$

$$= \begin{bmatrix} n\bar{x}_1\bar{x}_1 - n\bar{x}_1\bar{x}_1 - n\bar{x}_1\mu_1 + n\bar{x}_1\mu_1 & \cdots & n\bar{x}_1\bar{x}_p - n\bar{x}_1\bar{x}_p - n\bar{x}_1\mu_p + n\bar{x}_1\mu_p \\ \vdots & \ddots & \vdots \\ n\bar{x}_p\bar{x}_1 - n\bar{x}_p\bar{x}_1 - n\bar{x}_p\mu_1 + n\bar{x}_p\mu_1 & \cdots & n\bar{x}_p\bar{x}_p - n\bar{x}_p\bar{x}_p - n\bar{x}_p\mu_p + n\bar{x}_p\mu_p \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} = \underset{(p\times p)}{\mathbf{0}}$$

and

$$\sum_{j=1}^{n} (\bar{\mathbf{x}} - \boldsymbol{\mu})(\mathbf{x}_j - \bar{\mathbf{x}})' =$$

$$= \sum_{j=1}^{n} \left(\begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix} \right) \left(\begin{bmatrix} x_{j1} \\ \vdots \\ x_{jp} \end{bmatrix} - \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{bmatrix} \right)' =$$

$$= \sum_{j=1}^{n} \left(\begin{bmatrix} \bar{x}_1 - \mu_1 \\ \vdots \\ \bar{x}_p - \mu_p \end{bmatrix} \right) \left(\begin{bmatrix} x_{j1} - \bar{x}_1 \\ \vdots \\ x_{jp} - \bar{x}_p \end{bmatrix} \right)' =$$

$$= \sum_{j=1}^{n} \left(\begin{bmatrix} (\bar{x}_1 - \mu_1)(x_{j1} - \bar{x}_1) & \cdots & (\bar{x}_1 - \mu_1)(x_{jp} - \bar{x}_p) \\ \vdots & \ddots & \vdots \\ (\bar{x}_p - \mu_p)(x_{j1} - \bar{x}_1) & \cdots & (\bar{x}_p - \mu_p)(x_{jp} - \bar{x}_p) \end{bmatrix} \right) =$$

$$= \sum_{j=1}^{n} \left(\begin{bmatrix} (\bar{x}_1 x_{j1} - \mu_1 x_{j1} - \bar{x}_1\bar{x}_1 + \mu_1\bar{x}_1) & \cdots & (\bar{x}_1 x_{jp} - \mu_1 x_{jp} - \bar{x}_1\bar{x}_p + \mu_1\bar{x}_p) \\ \vdots & \ddots & \vdots \\ (\bar{x}_p x_{j1} - \mu_p x_{j1} - \bar{x}_p\bar{x}_1 - \mu_p\bar{x}_1) & \cdots & (\bar{x}_p x_{jp} - \mu_p x_{jp} - \bar{x}_p\bar{x}_p + \mu_p\bar{x}_p) \end{bmatrix} \right) =$$

$$= \begin{bmatrix} \sum_{j=1}^{n} (\bar{x}_1 x_{j1} - \mu_1 x_{j1} - \bar{x}_1\bar{x}_1 + \mu_1\bar{x}_1) & \cdots & \sum_{j=1}^{n} (\bar{x}_1 x_{jp} - \mu_1 x_{jp} - \bar{x}_1\bar{x}_p + \mu_1\bar{x}_p) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} (\bar{x}_p x_{j1} - \mu_p x_{j1} - \bar{x}_p\bar{x}_1 - \mu_p\bar{x}_1) & \cdots & \sum_{j=1}^{n} (\bar{x}_p x_{jp} - \mu_p x_{jp} - \bar{x}_p\bar{x}_p + \mu_p\bar{x}_p) \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{j=1}^{n} (\bar{x}_1 x_{j1}) - \sum_{j=1}^{n} (\mu_1 x_{j1}) - n\bar{x}_1\bar{x}_1 + n\mu_1\bar{x}_1 & \cdots & \sum_{j=1}^{n} (\bar{x}_1 x_{jp}) - \sum_{j=1}^{n} (\mu_1 x_{jp}) - n\bar{x}_1\bar{x}_p + n\mu_1\bar{x}_p \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} (\bar{x}_p x_{j1}) - \sum_{j=1}^{n} (\mu_p x_{j1}) - n\bar{x}_p\bar{x}_1 + n\mu_p\bar{x}_1 & \cdots & \sum_{j=1}^{n} (\bar{x}_p x_{jp}) - \sum_{j=1}^{n} (\mu_p x_{jp}) - n\bar{x}_p\bar{x}_p + n\mu_p\bar{x}_p \end{bmatrix} =$$

$$= \begin{bmatrix} n\bar{x}_1\bar{x}_1 - n\mu_1\bar{x}_1 - n\bar{x}_1\bar{x}_1 + n\mu_1\bar{x}_1 & \cdots & n\bar{x}_1\bar{x}_p - n\mu_1\bar{x}_p - n\bar{x}_1\bar{x}_p + n\mu_1\bar{x}_p \\ \vdots & \ddots & \vdots \\ n\bar{x}_p\bar{x}_1 - n\mu_p\bar{x}_1 - n\bar{x}_p\bar{x}_1 + n\mu_p\bar{x}_1 & \cdots & n\bar{x}_p\bar{x}_p - n\mu_p\bar{x}_p - n\bar{x}_p\bar{x}_p + n\mu_p\bar{x}_p \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} = \underset{(p\times p)}{\mathbf{0}}$$

4.16

Let $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$, and $\mathbf{X}_4$ be independent $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ random vectors.

(a) Find the marginal distributions for each of the random vectors

$$\mathbf{V}_1 = \frac{1}{4}\mathbf{X}_1 - \frac{1}{4}\mathbf{X}_2 + \frac{1}{4}\mathbf{X}_3 - \frac{1}{4}\mathbf{X}_4$$

and

$$\mathbf{V}_2 = \frac{1}{4}\mathbf{X}_1 + \frac{1}{4}\mathbf{X}_2 - \frac{1}{4}\mathbf{X}_3 - \frac{1}{4}\mathbf{X}_4$$

First, note that we have $n = 4$ random vectors. We're looking at the linear combination of random vectors, not a linear combination of the components of a random vector, so we use Result 4.8 on page 158, $\mathbf{V}_1 = c_1\mathbf{X}_1 + c_2\mathbf{X}_2 + \cdots + c_n\mathbf{X}_n$ is distributed as $N_p\left(\sum_{j=1}^n c_j \boldsymbol{\mu}_j, \left(\sum_{j=1}^n c_j^2\right)\boldsymbol{\Sigma}\right)$.
For $\mathbf{V}_1$ we have

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \end{bmatrix}$$

$$\text{mean} = \sum_{j=1}^n c_j \underset{(p\times1)}{\boldsymbol{\mu}_j} = \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}_1} - \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}_2} + \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}_3} - \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}_4} =$$

$$= \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}} - \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}} + \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}} - \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}} = (\frac{1}{4} - \frac{1}{4} + \frac{1}{4} - \frac{1}{4})\underset{(p\times1)}{\boldsymbol{\mu}} = \underset{(p\times1)}{\mathbf{0}}$$

$$\text{Covariance} = \left(\sum_{j=1}^n c_j^2\right)\boldsymbol{\Sigma} = \left(\left(\frac{1}{4}\right)^2 + \left(-\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2 + \left(-\frac{1}{4}\right)^2\right)\boldsymbol{\Sigma} =$$

$$= \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}\right)\boldsymbol{\Sigma} = \frac{4}{16}\boldsymbol{\Sigma} = \frac{1}{4}\boldsymbol{\Sigma}$$

so that $\mathbf{V}_1 \sim N_p\left(\underset{(p\times1)}{\mathbf{0}}, \underset{(p\times p)}{\left(\frac{1}{4}\right)\boldsymbol{\Sigma}}\right)$.

For $\mathbf{V}_2$ we have

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \end{bmatrix}$$

$$\text{mean} = \sum_{j=1}^n d_j \underset{(p\times1)}{\boldsymbol{\mu}_j} = \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}_1} + \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}_2} - \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}_3} - \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}_4} =$$

$$= \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}} + \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}} - \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}} - \frac{1}{4}\underset{(p\times1)}{\boldsymbol{\mu}} = (\frac{1}{4} + \frac{1}{4} - \frac{1}{4} - \frac{1}{4})\underset{(p\times1)}{\boldsymbol{\mu}} = \underset{(p\times1)}{\mathbf{0}}$$

$$\text{Covariance} = \left(\sum_{j=1}^{n} d_j^2\right)\boldsymbol{\Sigma} = \left(\left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2 + \left(-\frac{1}{4}\right)^2 + \left(-\frac{1}{4}\right)^2\right)\boldsymbol{\Sigma} =$$

$$= \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}\right)\boldsymbol{\Sigma} = \frac{4}{16}\boldsymbol{\Sigma} = \frac{1}{4}\boldsymbol{\Sigma}$$

so that $\mathbf{V}_2 \sim N_p\left(\underset{(p\times 1)}{\mathbf{0}}, \underset{(p\times p)}{\left(\tfrac{1}{4}\right)\boldsymbol{\Sigma}}\right)$.

(b) Find the joint density of the random vectors $\mathbf{V}_1$ and $\mathbf{V}_2$ defined in (a).

Again, from Result 4.8 on page 165, the joint distribution of $\mathbf{V}_1$ and $\mathbf{V}_2$ is multivariate normal. To compute the mean vector using what's on page 166, we know $n = 4$, but p can be anything, for $\underset{(p\times p)}{\mathbf{I}}$,

$$\underset{(2p\times np)}{\mathbf{A}} = \begin{bmatrix} c_1\mathbf{I} & c_2\mathbf{I} & c_3\mathbf{I} & c_4\mathbf{I} \\ d_1\mathbf{I} & d_2\mathbf{I} & d_3\mathbf{I} & d_4\mathbf{I} \end{bmatrix} = \begin{bmatrix} \frac{1}{4}\mathbf{I} & -\frac{1}{4}\mathbf{I} & \frac{1}{4}\mathbf{I} & -\frac{1}{4}\mathbf{I} \\ \frac{1}{4}\mathbf{I} & \frac{1}{4}\mathbf{I} & -\frac{1}{4}\mathbf{I} & -\frac{1}{4}\mathbf{I} \end{bmatrix}$$

$$\text{Vec}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{X}_3 \\ \mathbf{X}_4 \end{bmatrix} = \underset{(np\times 1)}{\begin{bmatrix} X_{11} \\ \vdots \\ X_{1p} \\ X_{21} \\ \vdots \\ X_{2p} \\ X_{31} \\ \vdots \\ X_{3p} \\ X_{41} \\ \vdots \\ X_{4p} \end{bmatrix}}$$

$$\mathbf{A}\text{Vec}(\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3,\mathbf{X}_4) = \begin{bmatrix} c_1\mathbf{I} & c_2\mathbf{I} & c_3\mathbf{I} & c_4\mathbf{I} \\ d_1\mathbf{I} & d_2\mathbf{I} & d_3\mathbf{I} & d_4\mathbf{I} \end{bmatrix} \begin{bmatrix} X_{11} \\ \vdots \\ X_{1p} \\ X_{21} \\ \vdots \\ X_{2p} \\ X_{31} \\ \vdots \\ X_{3p} \\ X_{41} \\ \vdots \\ X_{4p} \end{bmatrix} =$$

$$\begin{bmatrix} c_1X_{11}+c_2X_{21}+c_3X_{31}+c_4X_{41} \\ c_1X_{12}+c_2X_{22}+c_3X_{32}+c_4X_{42} \\ \vdots \\ c_1X_{1p}+c_2X_{2p}+c_3X_{3p}+c_4X_{4p} \\ d_1X_{11}+d_2X_{21}+d_3X_{31}+d_4X_{41} \\ d_1X_{12}+d_2X_{22}+d_3X_{32}+d_4X_{42} \\ \vdots \\ d_1X_{1p}+d_2X_{2p}+d_3X_{3p}+d_4X_{4p} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^4 c_jX_{j1} \\ \sum_{j=1}^4 c_jX_{j2} \\ \vdots \\ \sum_{j=1}^4 c_jX_{jp} \\ \sum_{j=1}^4 d_jX_{j1} \\ \sum_{j=1}^4 d_jX_{j2} \\ \vdots \\ \sum_{j=1}^4 d_jX_{jp} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^4 c_j\mathbf{X}_j \\ \sum_{j=1}^4 d_j\mathbf{X}_j \end{bmatrix}$$

The expected value of this is (using part (a))

$$\text{E}\left\{\begin{bmatrix} \sum_{j=1}^4 c_j\mathbf{X}_j \\ \sum_{j=1}^4 d_j\mathbf{X}_j \end{bmatrix}\right\} = \begin{bmatrix} \sum_{j=1}^4 c_j\text{E}[\mathbf{X}_j] \\ \sum_{j=1}^4 d_j\text{E}[\mathbf{X}_j] \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^4 c_j\boldsymbol{\mu}_j \\ \sum_{j=1}^4 d_j\boldsymbol{\mu}_j \end{bmatrix} = \begin{bmatrix} \underset{p\times 1}{\mathbf{0}} \\ \underset{p\times 1}{\mathbf{0}} \end{bmatrix} = \underset{2p\times 1}{\mathbf{0}}$$

The covariance matrix from Result 4.8 is on page 165,

$$\mathbf{d}'\mathbf{c} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \end{bmatrix} = \frac{1}{16} - \frac{1}{16} - \frac{1}{16} + \frac{1}{16} = 0$$

$$\begin{bmatrix} \left(\sum_{i=1}^n c_j^2\right)\boldsymbol{\Sigma} & (\mathbf{d}'\mathbf{c})\,\boldsymbol{\Sigma} \\ (\mathbf{d}'\mathbf{c})\,\boldsymbol{\Sigma} & \left(\sum_{i=1}^n d_j^2\right)\boldsymbol{\Sigma} \end{bmatrix} = \begin{bmatrix} \left(\frac{1}{4}\right)\boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \left(\frac{1}{4}\right)\boldsymbol{\Sigma} \end{bmatrix} = \left(\frac{1}{4}\right)\begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} \end{bmatrix}$$

and mean vector

$$\text{Vec}(\mathbf{V}_1,\mathbf{V}_2) = \begin{bmatrix} \underset{p\times 1}{\mathbf{0}} \\ \underset{p\times 1}{\mathbf{0}} \end{bmatrix} = \underset{2p\times 1}{\mathbf{0}}$$

Also, $\mathbf{V}_1$ and $\mathbf{V}_2$ are independent, since $\mathbf{d}'\mathbf{c} = 0$.

4.17

Let $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$, and $\mathbf{X}_4$ be independent and identically distributed random vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Find the mean vector and covariance matrices for each of the two linear combinations of random vectors

$$\frac{1}{5}\mathbf{X}_1 + \frac{1}{5}\mathbf{X}_2 + \frac{1}{5}\mathbf{X}_3 + \frac{1}{5}\mathbf{X}_4 + \frac{1}{5}\mathbf{X}_5$$

and

$$\mathbf{X}_1 - \mathbf{X}_2 + \mathbf{X}_3 - \mathbf{X}_4 + \mathbf{X}_5$$

in terms of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Also, obtain the covariance between the two linear combinations of random vectors.

First some detail for my own education. The vector $\mathbf{c} = [1/5, 1/5, 1/5, 1/5, 1/5]'$ is (5×1), but the vector $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4, \mathbf{X}_5]'$ is $(np \times 1)$, so computing $\mathbf{c}'\mathbf{X}$ is not conformable (can't multiply $(1 \times p)$ by a $(np \times 1)$). We can get the job done by multiplying $\mathbf{c}'$ by the kronecker product of the identity matrix, $\mathbf{c}' \otimes \mathbf{I}_p$ first. Each of the $\mathbf{X}_j$ is $p \times 1$ and $\mathbf{I}_p$ is $p \times p$.

$$(\mathbf{c}' \otimes \mathbf{I}_p)\,\mathbf{X} = \left(\begin{bmatrix} c_1 & c_2 & c_3 & c_4 & c_5 \end{bmatrix} \otimes \mathbf{I}_p\right) \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{X}_3 \\ \mathbf{X}_4 \\ \mathbf{X}_5 \end{bmatrix} =$$

$$= \begin{bmatrix} c_1\mathbf{I}_p & c_2\mathbf{I}_p & c_3\mathbf{I}_p & c_4\mathbf{I}_p & c_5\mathbf{I}_p \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{X}_3 \\ \mathbf{X}_4 \\ \mathbf{X}_5 \end{bmatrix} =$$

$$= c_1\mathbf{I}_p\mathbf{X}_1 + c_2\mathbf{I}_p\mathbf{X}_2 + c_3\mathbf{I}_p\mathbf{X}_3 + c_4\mathbf{I}_p\mathbf{X}_4 + c_5\mathbf{I}_p\mathbf{X}_5 =$$

$$= c_1\mathbf{X}_1 + c_2\mathbf{X}_2 + c_3\mathbf{X}_3 + c_4\mathbf{X}_4 + c_5\mathbf{X}_5 =$$

$$= \frac{1}{5}\mathbf{X}_1 + \frac{1}{5}\mathbf{X}_2 + \frac{1}{5}\mathbf{X}_3 + \frac{1}{5}\mathbf{X}_4 + \frac{1}{5}\mathbf{X}_5$$

<u>Mean:</u>

$$\mathrm{E}\left[(\mathbf{c}' \otimes \mathbf{I}_p)\,\mathbf{X}\right] =$$

$$= \mathrm{E}\left[\frac{1}{5}\mathbf{X}_1 + \frac{1}{5}\mathbf{X}_2 + \frac{1}{5}\mathbf{X}_3 + \frac{1}{5}\mathbf{X}_4 + \frac{1}{5}\mathbf{X}_5\right] =$$

$$= \mathrm{E}\left[\frac{1}{5}\mathbf{X}_1\right] + \mathrm{E}\left[\frac{1}{5}\mathbf{X}_2\right] + \mathrm{E}\left[\frac{1}{5}\mathbf{X}_3\right] + \mathrm{E}\left[\frac{1}{5}\mathbf{X}_4\right] + \mathrm{E}\left[\frac{1}{5}\mathbf{X}_5\right] =$$

$$= \frac{1}{5}\mathrm{E}\left[\mathbf{X}_1\right] + \frac{1}{5}\mathrm{E}\left[\mathbf{X}_2\right] + \frac{1}{5}\mathrm{E}\left[\mathbf{X}_3\right] + \frac{1}{5}\mathrm{E}\left[\mathbf{X}_4\right] + \frac{1}{5}\mathrm{E}\left[\mathbf{X}_5\right] =$$

$$= \frac{1}{5}\boldsymbol{\mu} + \frac{1}{5}\boldsymbol{\mu} + \frac{1}{5}\boldsymbol{\mu} + \frac{1}{5}\boldsymbol{\mu} + \frac{1}{5}\boldsymbol{\mu}$$
$$= \left(\frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5}\right)\boldsymbol{\mu} = \boldsymbol{\mu}$$

Covariance:

$$\text{Covar}\left[(\mathbf{c}' \otimes \mathbf{I}_p)\,\mathbf{X}\right] =$$
$$= \text{Covar}\left[\frac{1}{5}\mathbf{X}_1 + \frac{1}{5}\mathbf{X}_2 + \frac{1}{5}\mathbf{X}_3 + \frac{1}{5}\mathbf{X}_4 + \frac{1}{5}\mathbf{X}_5\right] =$$
$$= \text{Covar}\left[\frac{1}{5}\mathbf{X}_1\right] + \text{Covar}\left[\frac{1}{5}\mathbf{X}_2\right] + \text{Covar}\left[\frac{1}{5}\mathbf{X}_3\right] + \text{Covar}\left[\frac{1}{5}\mathbf{X}_4\right] + \text{Covar}\left[\frac{1}{5}\mathbf{X}_5\right] =$$
$$= \left(\frac{1}{5}\right)^2 \text{Covar}\left[\mathbf{X}_1\right] + \left(\frac{1}{5}\right)^2 \text{Covar}\left[\mathbf{X}_2\right] + \left(\frac{1}{5}\right)^2 \text{Covar}\left[\mathbf{X}_3\right] + \left(\frac{1}{5}\right)^2 \text{Covar}\left[\mathbf{X}_4\right] + \left(\frac{1}{5}\right)^2 \text{Covar}\left[\mathbf{X}_5\right] =$$
$$= \left(\frac{1}{5}\right)^2 \boldsymbol{\Sigma} + \left(\frac{1}{5}\right)^2 \boldsymbol{\Sigma} + \left(\frac{1}{5}\right)^2 \boldsymbol{\Sigma} + \left(\frac{1}{5}\right)^2 \boldsymbol{\Sigma} + \left(\frac{1}{5}\right)^2 \boldsymbol{\Sigma} =$$
$$= 5\left(\frac{1}{5}\right)^2 \boldsymbol{\Sigma} = \frac{1}{5}\boldsymbol{\Sigma}$$

and for

$$\mathbf{X}_1 - \mathbf{X}_2 + \mathbf{X}_3 - \mathbf{X}_4 + \mathbf{X}_5$$

$$\text{Here, we have} \quad \mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

Mean:

$$\text{E}\left[\left(\mathbf{d}' \otimes \mathbf{I}_p\right)\mathbf{X}\right] = \sum_{j=1}^{5} d_j \boldsymbol{\mu}_j = \sum_{j=1}^{5} d_j \boldsymbol{\mu} = (1 - 1 + 1 - 1 + 1)\,\boldsymbol{\mu} = \boldsymbol{\mu}$$

Covariance:

$$\text{Covar}\left[\left(\mathbf{d}' \otimes \mathbf{I}_p\right)\mathbf{X}\right] = \left(\sum_{j=1}^{5} d_j^2\right)\boldsymbol{\Sigma} = \left(1^2 - 1^2 + 1^2 - 1^2 + 1^2\right)\boldsymbol{\Sigma} = 5\boldsymbol{\Sigma}$$

Before finding the covariance betweeen the two linear combinations we need to compute

$$\mathbf{d}'\mathbf{c} = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix} = \frac{1}{5}$$

Now, the covariance betweeen the two linear combinations is

$$\begin{bmatrix} \left(\sum_{j=1}^{5} c_j^2\right)\boldsymbol{\Sigma} & (\mathbf{d}'\mathbf{c})\,\boldsymbol{\Sigma} \\ (\mathbf{d}'\mathbf{c})\,\boldsymbol{\Sigma} & \left(\sum_{j=1}^{5} d_j^2\right)\boldsymbol{\Sigma} \end{bmatrix} = \begin{bmatrix} \frac{1}{5}\boldsymbol{\Sigma} & \frac{1}{5}\boldsymbol{\Sigma} \\ \frac{1}{5}\boldsymbol{\Sigma} & 5\boldsymbol{\Sigma} \end{bmatrix}$$

4.18

Find the maximum likelihood estimates of the 2×1 mean vector $\boldsymbol{\mu}$ and the 2×2 covariance matrix $\boldsymbol{\Sigma}$ based on the random sample

$$\begin{bmatrix} 3 & 6 \\ 4 & 4 \\ 5 & 7 \\ 4 & 7 \end{bmatrix}$$

from a bivariate normal population.

Using Result 4.11 on page 171

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} = \begin{bmatrix} \bar{\mathbf{x}}_1 \\ \bar{\mathbf{x}}_2 \end{bmatrix} = \begin{bmatrix} \frac{(3+4+5+4)}{4} \\ \frac{(6+4+7+7)}{4} \end{bmatrix} = \begin{bmatrix} \frac{16}{4} \\ \frac{24}{4} \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

$$\hat{\Sigma} = \left(\frac{(n-1)}{n}\right)\mathbf{S} = \left(\frac{1}{4}\right)\sum_{j=1}^{4}\left(\mathbf{X}_j - \bar{\mathbf{X}}\right)\left(\mathbf{X}_j - \bar{\mathbf{X}}\right)' =$$

$$= \left(\frac{1}{4}\right)\left\{\left(\begin{bmatrix} 3 \\ 6 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \end{bmatrix}\right)\left(\begin{bmatrix} 3 \\ 6 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \end{bmatrix}\right)' + \left(\begin{bmatrix} 4 \\ 4 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \end{bmatrix}\right)\left(\begin{bmatrix} 4 \\ 4 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \end{bmatrix}\right)' + \right.$$

$$\left.\left(\begin{bmatrix} 5 \\ 7 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \end{bmatrix}\right)\left(\begin{bmatrix} 5 \\ 7 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \end{bmatrix}\right)' + \left(\begin{bmatrix} 4 \\ 7 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \end{bmatrix}\right)\left(\begin{bmatrix} 4 \\ 7 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \end{bmatrix}\right)'\right\} =$$

$$= \left(\frac{1}{4}\right)\left\{\begin{bmatrix} -1 \\ 0 \end{bmatrix}\begin{bmatrix} -1 \\ 0 \end{bmatrix}' + \begin{bmatrix} 0 \\ -2 \end{bmatrix}\begin{bmatrix} 0 \\ -2 \end{bmatrix}' + \begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}' + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}'\right\} =$$

$$= \left(\frac{1}{4}\right)\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\} =$$

$$= \left(\frac{1}{4}\right)\begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix} = \begin{bmatrix} (1/2) & -(1/4) \\ -(1/4) & (3/2) \end{bmatrix}$$

4.19

Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{20}$ be a random sample of size $n = 20$ from a $N_6(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ population. Specify each of the following completely.

(a) The distribution of $(\mathbf{X}_1 - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}_1 - \boldsymbol{\mu})$
Can use Result 4.7 (a) on page 163 for this

$$(\mathbf{X}_1 - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}_1 - \boldsymbol{\mu}) \sim \chi^2_6$$

(b) The distribution of $\bar{\mathbf{X}}$ and $\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})$

$$\mathbf{Y}_1 = \bar{\mathbf{X}} = \begin{bmatrix} \bar{\mathbf{x}}_1 \\ \vdots \\ \bar{\mathbf{x}}_6 \end{bmatrix} = (\mathbf{X}_1 + \cdots + \mathbf{X}_{20})/20 = \frac{\sum_{j=1}^{20} \mathbf{X}_j}{20}$$

$$\mathrm{E}[\mathbf{Y}_1] = \mathrm{E}\left[\frac{\sum_{j=1}^{20} \mathbf{X}_j}{20}\right] = \frac{\sum_{j=1}^{20} \mathrm{E}[\mathbf{X}_j]}{20} = \frac{\sum_{j=1}^{20} \boldsymbol{\mu}}{20} = \frac{20\boldsymbol{\mu}}{20} = \boldsymbol{\mu}$$

$$\mathrm{Covar}[\mathbf{Y}_1] = \mathrm{Covar}\left[\frac{\sum_{j=1}^{20} \mathbf{X}_j}{20}\right] = \left(\frac{1}{20}\right)^2 \sum_{j=1}^{20} \mathrm{Covar}[\mathbf{X}_j] =$$

$$\left(\frac{1}{20}\right)^2 \left(\sum_{j=1}^{20} \boldsymbol{\Sigma}\right) = \left(\frac{1}{20}\right)^2 (20\boldsymbol{\Sigma}) = \frac{1}{20}\boldsymbol{\Sigma}$$

$$\mathbf{Y}_1 = \bar{\mathbf{X}} \sim N_6\left(\boldsymbol{\mu}, \frac{1}{20}\boldsymbol{\Sigma}\right)$$

This is the same as (4–23) in section 4.4 on page 174.
If $\mathbf{Y}_2 = \sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})$, then

$$\mathrm{E}[\mathbf{Y}] = \mathrm{E}\left[\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})\right] = \mathrm{E}\left[\sqrt{n}\bar{\mathbf{X}}\right] - \mathrm{E}\left[\sqrt{n}\boldsymbol{\mu}\right] = \mathrm{E}\left[\sqrt{n}\bar{\mathbf{X}}\right] - \mathrm{E}\left[\sqrt{n}\boldsymbol{\mu}\right] =$$

$$= \sqrt{n}\mathrm{E}\left[\bar{\mathbf{X}}\right] - \sqrt{n}\boldsymbol{\mu} = \sqrt{n}\boldsymbol{\mu} - \sqrt{n}\boldsymbol{\mu} = \mathbf{0}$$

$$\mathrm{Covar}[\mathbf{Y}_2] = \mathrm{Covar}\left[\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})\right] = \mathrm{Covar}\left[\sqrt{n}\bar{\mathbf{X}}\right] - \mathrm{Covar}\left[\sqrt{n}\boldsymbol{\mu}\right] =$$

$$= \left(\sqrt{n}\right)^2 \mathrm{Covar}\left[\bar{\mathbf{X}}\right] - \mathbf{0} = n\left(\frac{1}{n}\boldsymbol{\Sigma}\right) = \boldsymbol{\Sigma}$$

$$\mathbf{Y}_2 = \sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim N_6(\mathbf{0}, \boldsymbol{\Sigma})$$

This is the same as the approximation in (4–28) in section 4.5 on page 176.

(c) The distribution of (n-1)**S**

$$(n-1)\mathbf{S} = \frac{n-1}{n-1}\sum_{j=1}^{n}\left(\mathbf{X}_j - \bar{\mathbf{X}}\right)\left(\mathbf{X}_j - \bar{\mathbf{X}}\right)' =$$

$$= \sum_{j=1}^{n}\left(\mathbf{X}_j + \boldsymbol{\mu} - \boldsymbol{\mu} - \bar{\mathbf{X}}\right)\left(\mathbf{X}_j + \boldsymbol{\mu} - \boldsymbol{\mu} - \bar{\mathbf{X}}\right)' =$$

$$= \sum_{j=1}^{n}\left((\mathbf{X}_j - \boldsymbol{\mu}) - (\bar{\mathbf{X}} - \boldsymbol{\mu})\right)\left((\mathbf{X}_j - \boldsymbol{\mu}) - (\bar{\mathbf{X}} - \boldsymbol{\mu})\right)' =$$

$$= \sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - (\bar{\mathbf{X}} - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - (\mathbf{X}_j - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' + (\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' =$$

$$= \sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2\sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' + \sum_{j=1}^{n} (\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' =$$

$$= \sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2(\bar{\mathbf{X}} - \boldsymbol{\mu})\sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' =$$

$$= \sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2(\bar{\mathbf{X}} - \boldsymbol{\mu})\left(\left(\sum_{j=1}^{n} \mathbf{X}_j\right) - n\boldsymbol{\mu}\right)' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' =$$

$$= \sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2(\bar{\mathbf{X}} - \boldsymbol{\mu})(n\bar{\mathbf{X}} - n\boldsymbol{\mu})' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' =$$

$$= \sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - 2n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' =$$

$$= \sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' =$$

$$= \sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' - \left[\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})\right]\left[\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})'\right] =$$

so looking at the first and second part of this equation

$$\mathbf{X}_j \sim N(\underset{(p\times 1)}{\boldsymbol{\mu}}, \underset{(p\times p)}{\boldsymbol{\Sigma}}) \;\Rightarrow\; (\mathbf{X}_j - \boldsymbol{\mu}) \sim N(\mathbf{0}, \Sigma) \Rightarrow$$

$$\Rightarrow (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' \sim \chi^2_1 \Rightarrow$$

$$\Rightarrow \sum_{j=1}^{n} (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' \sim W_n(\Sigma)$$

That is, the first term in the sum is distributed as a Wishart distribution with n degrees of freedom. For the second term,

$$\bar{\mathbf{X}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n) \Rightarrow (\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Sigma}/n) \Rightarrow$$

$$\Rightarrow \sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \Rightarrow$$

$$\left[\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})\right]\left[\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})'\right] \sim W_1(\boldsymbol{\Sigma})$$

That is, the second term in the sum is distributed as a Wishart distribution with 1 degree of freedom.

Combining the two, we now has a Wishart distribution with n degrees of freedom minus another Wishart distribution with 1 degree of freedom to create a Wishart distribution with $n-1$ distribution, so we finally have

$$(n-1)\mathbf{S} \sim W_{n-1}\left(\underset{(p\times p)}{\boldsymbol{\Sigma}}\right)$$

This is the answers found in (4–23) on page 174. For this data we have

$$(n-1)\mathbf{S} \sim W_{20-1}\left(\underset{(6\times 6)}{\boldsymbol{\Sigma}}\right)$$

A Wishart distribution with 19 degrees of freedom.

4.20

For the random variables $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{20}$ in Exercise 4.19, specify the distribution of $\mathbf{B}(19\mathbf{S})\mathbf{B}'$ in each case

(a) $\mathbf{B} = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$

Use the properties of the Wishart distribution from page 174 along with (4–23)

$$\mathbf{B}(19\mathbf{S})\mathbf{B}' \sim W_{19}\left(\mathbf{B}(19\mathbf{S})\mathbf{B}' \middle| \mathbf{B}(19\boldsymbol{\Sigma})\mathbf{B}'\right)$$

where

$$\mathbf{B}(19\mathbf{S})\mathbf{B}' = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix} \times \begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{21} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} & s_{36} \\ s_{41} & s_{42} & s_{43} & s_{44} & s_{45} & s_{46} \\ s_{51} & s_{52} & s_{53} & s_{54} & s_{55} & s_{56} \\ s_{61} & s_{62} & s_{63} & s_{64} & s_{65} & s_{66} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} =$$

$$\begin{bmatrix} (s_{11} - \frac{1}{2}s_{21} - \frac{1}{2}s_{31}) & (s_{12} - \frac{1}{2}s_{22} - \frac{1}{2}s_{32}) & \cdots & (s_{16} - \frac{1}{2}s_{26} - \frac{1}{2}s_{36}) \\ (-\frac{1}{2}s_{41} - \frac{1}{2}s_{51} + s_{61}) & (-\frac{1}{2}s_{42} - \frac{1}{2}s_{52} + s_{62}) & \cdots & (-\frac{1}{2}s_{46} - \frac{1}{2}s_{56} + s_{66}) \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \text{element}_{11} & \text{element}_{12} \\ \text{element}_{21} & \text{element}_{22} \end{bmatrix}$$

where
$\text{element}_{11} = (s_{11} - \frac{1}{2}s_{21} - \frac{1}{2}s_{31}) - \frac{1}{2}(s_{12} - \frac{1}{2}s_{22} - \frac{1}{2}s_{32}) - \frac{1}{2}(s_{13} - \frac{1}{2}s_{23} - \frac{1}{2}s_{33})$
$\text{element}_{12} = -\frac{1}{2}(s_{14} - \frac{1}{2}s_{24} - \frac{1}{2}s_{34}) - \frac{1}{2}(s_{15} - \frac{1}{2}s_{25} - \frac{1}{2}s_{35}) + (s_{16} - \frac{1}{2}s_{26} - \frac{1}{2}s_{36})$
$\text{element}_{21} = (-\frac{1}{2}s_{41} - \frac{1}{2}s_{51} + s_{61}) - \frac{1}{2}(-\frac{1}{2}s_{42} - \frac{1}{2}s_{52} + s_{62}) - \frac{1}{2}(-\frac{1}{2}s_{43} - \frac{1}{2}s_{53} + s_{63})$
$\text{element}_{22} = -\frac{1}{2}(-\frac{1}{2}s_{44} - \frac{1}{2}s_{54} + s_{64}) - \frac{1}{2}(-\frac{1}{2}s_{45} - \frac{1}{2}s_{55} + s_{65}) + (-\frac{1}{2}s_{46} - \frac{1}{2}s_{56} + s_{66})$ and

$$\mathbf{B}(19\mathbf{\Sigma})\mathbf{B}' = \begin{bmatrix} \text{element}_{11} & \text{element}_{12} \\ \text{element}_{21} & \text{element}_{22} \end{bmatrix}$$

where
$\text{element}_{11} = (\sigma_{11} - \frac{1}{2}\sigma_{21} - \frac{1}{2}\sigma_{31}) - \frac{1}{2}(\sigma_{12} - \frac{1}{2}\sigma_{22} - \frac{1}{2}\sigma_{32}) - \frac{1}{2}(\sigma_{13} - \frac{1}{2}\sigma_{23} - \frac{1}{2}\sigma_{33})$
$\text{element}_{12} = -\frac{1}{2}(\sigma_{14} - \frac{1}{2}\sigma_{24} - \frac{1}{2}\sigma_{34}) - \frac{1}{2}(\sigma_{15} - \frac{1}{2}\sigma_{25} - \frac{1}{2}\sigma_{35}) + (\sigma_{16} - \frac{1}{2}\sigma_{26} - \frac{1}{2}\sigma_{36})$
$\text{element}_{21} = (-\frac{1}{2}\sigma_{41} - \frac{1}{2}\sigma_{51} + \sigma_{61}) - \frac{1}{2}(-\frac{1}{2}\sigma_{42} - \frac{1}{2}\sigma_{52} + \sigma_{62}) - \frac{1}{2}(-\frac{1}{2}\sigma_{43} - \frac{1}{2}\sigma_{53} + \sigma_{63})$
$\text{element}_{22} = -\frac{1}{2}(-\frac{1}{2}\sigma_{44} - \frac{1}{2}\sigma_{54} + \sigma_{64}) - \frac{1}{2}(-\frac{1}{2}\sigma_{45} - \frac{1}{2}\sigma_{55} + \sigma_{65}) + (-\frac{1}{2}\sigma_{46} - \frac{1}{2}\sigma_{56} + \sigma_{66})$

(b) $\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

Again, use the properties of the Wishart distribution from page 174 along with (4–23)

$$\mathbf{B}(19\mathbf{S})\mathbf{B}' \sim W_{19}\left(\mathbf{B}(19\mathbf{S})\mathbf{B}' \middle| \mathbf{B}(19\mathbf{\Sigma})\mathbf{B}'\right)$$

where

$$\mathbf{B}(19\mathbf{S})\mathbf{B}' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{21} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} & s_{36} \\ s_{41} & s_{42} & s_{43} & s_{44} & s_{45} & s_{46} \\ s_{51} & s_{52} & s_{53} & s_{54} & s_{55} & s_{56} \\ s_{61} & s_{62} & s_{63} & s_{64} & s_{65} & s_{66} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} & s_{36} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{13} \\ s_{31} & s_{33} \end{bmatrix}$$

and

$$\mathbf{B}(19\boldsymbol{\Sigma})\mathbf{B}' = \begin{bmatrix} \sigma_{11} & \sigma_{13} \\ \sigma_{31} & \sigma_{33} \end{bmatrix}$$

4.21

Let $\mathbf{X}_1, \ldots, \mathbf{X}_{60}$ be a random sample of size 60 from a four-variate normal distribution having mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. Specify each of the following completely.

(a) The distribution of $\bar{\mathbf{X}}$

Using 1. from (4–23) on page 174

$$\bar{\mathbf{X}} \sim N_4 \left(\underset{(4\times 1)}{\boldsymbol{\mu}}, \frac{1}{60} \underset{(4\times 4)}{\boldsymbol{\Sigma}} \right)$$

(b) The distribution of $(\mathbf{X}_1 - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_1 - \boldsymbol{\mu})$

Using Result 4.7 (a) in section 4.2 on page 163.

$$(\mathbf{X}_1 - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_1 - \boldsymbol{\mu}) \sim \chi_4^2$$

(c) The distribution of $n\big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)' \boldsymbol{\Sigma}^{-1} \big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)$

Similar to part (b), but I'll do the work.

$$n\big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)' \boldsymbol{\Sigma}^{-1} \big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big) =$$

$$= \sqrt{n}\sqrt{n}\big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)' \boldsymbol{\Sigma}^{-1/2} \boldsymbol{\Sigma}^{-1/2} \big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big) =$$

$$= \left[\big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)' \boldsymbol{\Sigma}^{-1/2} \sqrt{n}\right] \left[\sqrt{n} \boldsymbol{\Sigma}^{-1/2} \big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)\right] =$$

$$= \left[\sqrt{n} \boldsymbol{\Sigma}^{-1/2} \big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)\right]' \left[\sqrt{n} \boldsymbol{\Sigma}^{-1/2} \big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)\right] =$$

$$= \mathbf{Z}'\mathbf{Z}$$

$$\mathrm{E}[\mathbf{Z}] = \mathrm{E}\left[\sqrt{n} \boldsymbol{\Sigma}^{-1/2} \big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)\right] = \sqrt{n} \boldsymbol{\Sigma}^{-1/2} \mathrm{E}\left[\big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)\right] =$$

$$= \sqrt{n} \boldsymbol{\Sigma}^{-1/2} \big(\mathrm{E}\big[\bar{\mathbf{X}}\big] - \mathrm{E}[\boldsymbol{\mu}]\big) = \sqrt{n} \boldsymbol{\Sigma}^{-1/2} (\boldsymbol{\mu} - \boldsymbol{\mu}) = \sqrt{n} \boldsymbol{\Sigma}^{-1/2} \mathbf{0} = \mathbf{0}$$

$$\mathrm{Covar}[\mathbf{Z}] = \mathrm{Covar}\left[\sqrt{n} \boldsymbol{\Sigma}^{-1/2} \big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)\right] =$$

$$= \big(\sqrt{n}\big)^2 \boldsymbol{\Sigma}^{-1/2} \mathrm{Covar}\left[\big(\bar{\mathbf{X}} - \boldsymbol{\mu}\big)\right] \left(\boldsymbol{\Sigma}^{-1/2}\right)' =$$

$$= n \boldsymbol{\Sigma}^{-1/2} \big(\mathrm{Covar}\big[\bar{\mathbf{X}}\big] + \mathrm{Covar}[\boldsymbol{\mu}]\big) \boldsymbol{\Sigma}^{-1/2} =$$

$$= n\mathbf{\Sigma}^{-1/2}\left(\frac{1}{n}\mathbf{\Sigma}+0\right)\mathbf{\Sigma}^{-1/2} =$$

$$= n\mathbf{\Sigma}^{-1/2}\left(\frac{1}{n}\mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{1/2}\right)\mathbf{\Sigma}^{-1/2} =$$

$$= n\frac{1}{n}\left(\mathbf{\Sigma}^{-1/2}\mathbf{\Sigma}^{1/2}\right)\left(\mathbf{\Sigma}^{-1/2}\mathbf{\Sigma}^{1/2}\right) =$$

$$= (\mathbf{I}_p)(\mathbf{I}_p) = \mathbf{I}_p$$

We now have that

$$\mathbf{Z} \sim N_p(\mathbf{0}, \mathbf{I}_p)$$

Now $\mathbf{Z}'\mathbf{Z} = \sum_{j=1}^{p} Z_j^2$, the sum of squared standard normal random variables is a chi-squared distribution with p degrees of freedom, so

$$n(\bar{\mathbf{X}}-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\bar{\mathbf{X}}-\boldsymbol{\mu}) \sim \chi^2_4$$

(d) The approximate distribution of $n(\bar{\mathbf{X}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}}-\boldsymbol{\mu})$

Well, if we know that $\bar{\mathbf{X}} \sim N(\boldsymbol{\mu}, \frac{1}{n}\mathbf{\Sigma})$ and that $(n-1)\mathbf{S}$ is a $p=4$-variate Wishart distribution with $(60-1)$ degrees of freedom, then $n(\bar{\mathbf{X}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}}-\boldsymbol{\mu})$ is a Hotelling's $T^2_{p=4,n-1=60-1}$ distribution. The T^2 is related to an F-distribution as $T^2_{4,59} = \frac{4(60-1)}{60-4}F_{p=4,n-p=60-4}$, covered in Chapter 5. Here, they're asking for an approximation though, so instead we can use the chi-squared approximation in (4–28) on page 176, where they aren't making assumtions about the distribution of the random sample (same as the answer to Exercise 4.22 (b)), where for large n relative to p, we can use the CLT and

$$n(\bar{\mathbf{X}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}}-\boldsymbol{\mu}) \doteq$$

$$\left[\sqrt{n}(\bar{\mathbf{X}}-\boldsymbol{\mu})'\right]\mathbf{\Sigma}^{-1}\left[\sqrt{n}(\bar{\mathbf{X}}-\boldsymbol{\mu})\right] =$$

$$= \left[\sqrt{n}(\bar{\mathbf{X}}-\boldsymbol{\mu})\right]'\mathbf{\Sigma}^{-1/2}\mathbf{\Sigma}^{-1/2}\left[\sqrt{n}(\bar{\mathbf{X}}-\boldsymbol{\mu})\right] =$$

$$= \left[\sqrt{n}\mathbf{\Sigma}^{-1/2}(\bar{\mathbf{X}}-\boldsymbol{\mu})\right]'\left[\sqrt{n}\mathbf{\Sigma}^{-1/2}(\bar{\mathbf{X}}-\boldsymbol{\mu})\right] = \text{same as in part (c)}$$

$$n(\bar{\mathbf{X}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}}-\boldsymbol{\mu}) \mathrel{\dot\sim} \chi^2_4$$

4.22

Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{75}$ be a random sample from a population distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$. What is the approximate distribution of each of the following?

(a) $\bar{\mathbf{X}}$

We have 75 observations, which is considered fairly large, so the asymptotic results from the section 4.5 apply. As n increases, $\bar{\mathbf{X}}$ converges in probability to $\boldsymbol{\mu}$ and $\mathbf{S}$ converges in probability to $\boldsymbol{\Sigma}$. Result 4.13 basically gives our answer. In the result, they have $\sqrt{n}\left(\bar{\mathbf{X}}-\boldsymbol{\mu}\right) \dot{\sim} N_p\left(\mathbf{0},\boldsymbol{\Sigma}\right)$. If we leave off the $\sqrt{n}$ and the subtraction of $\boldsymbol{\mu}$ we have

$$\bar{\mathbf{X}} \dot{\sim} N_p\left(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}\right)$$

(b) $n\left(\bar{\mathbf{X}}-\boldsymbol{\mu}\right)'\mathbf{S}^{-1}\left(\bar{\mathbf{X}}-\boldsymbol{\mu}\right)$

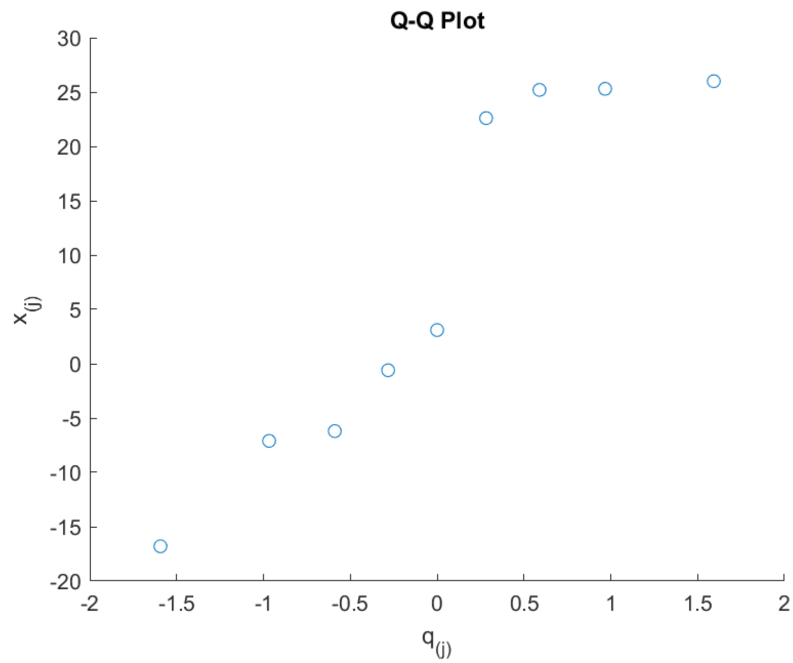
Again, we only know that we have a population mean vector $\boldsymbol{\mu}$ and covaraince matrix $\boldsymbol{\Sigma}$. We're not making any statements about what distribution they belong to. Using the central limit theorem from Result 4.13 from page 176, where $\sqrt{n}\left(\bar{\mathbf{X}}-\boldsymbol{\mu}\right) \dot{\sim} N_p\left(\mathbf{0},\boldsymbol{\Sigma}\right)$ and then $n\left(\bar{\mathbf{X}}-\boldsymbol{\mu}\right)'\boldsymbol{\Sigma}^{-1}\left(\bar{\mathbf{X}}-\boldsymbol{\mu}\right) \dot{\sim} \chi^2_p$ so replacing the true population covaraince, $\boldsymbol{\Sigma}$, with the sample covariance, $\mathbf{S}$, won't cause massive errors as long as $n >> p$, and we have

$$n\left(\bar{\mathbf{X}}-\boldsymbol{\mu}\right)'\mathbf{S}^{-1}\left(\bar{\mathbf{X}}-\boldsymbol{\mu}\right) \dot{\sim} \chi^2_p$$

4.23

Consider the annual rates of return (including dividends) on the Dow-Jones industrial average for the years 1996–2005. These data, multiplied by 100, are -0.6 3.1 25.3 -16.8 -7.1 -6.2 25.2 22.6 26.0.
Use these 10 observations to complete the following.

(a) Construct a Q-Q plot. Do the data seem to be normally distributed? Explain.

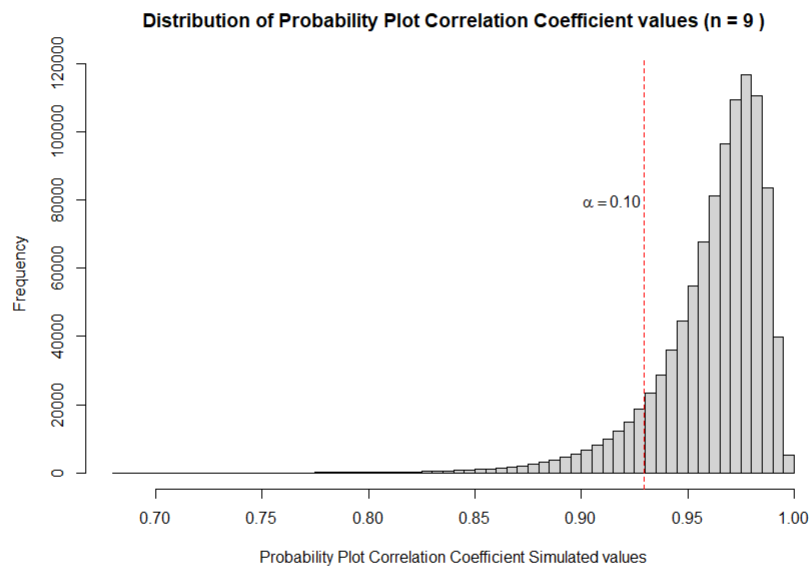


(b) Carry out a test of normality based on the correlation coefficient r_Q. [See (4–31).] Let the significance level be $\alpha = .10$.

$$H_0 : \text{The data is normally distributed.}$$
$$H_\alpha : \text{The data is NOT normally distributed.}$$

The r_Q value is 0.9351. From table 4.2, using $\alpha = 0.10$ and sample size $n = 10$, the critical point is 0.9351. We only have 9 observations, so the citical value is slightly less, so $r_Q \geq 0.9351$, so we'd conclude the data is normal. It's right on the border though.
Going the extra mile, using the Filliben paper from Reference 5 in Chaper 4, *The Probability Plot Test for Correlation coefficient Test for Normality* we can simulate values just like those in Table 4.2, but for a sample size of 9. The simulation invlolves generating 9 points from a standard normal and sorting the 9 values. The probability levels are computed and standard normal quantiles are computed (just like in Example 4.9). The correlation coefficient is computed using the data and the quantiles. The correlation coefficient is then saved. This was performed $N = 10^6 = 1\text{M}$ times (1M samples). In the Filliben paper, there were $N = 10^5 = 100\text{K}$ samples genereated. The 1M values were used to create quantile at the 0.10-level. The simulated value was 0.9295314. This is smaller than the r_Q value of 0.9351. Based on the simulated value we would conclude H_0, that the data is normally distributed. The histogram of the simulated values is

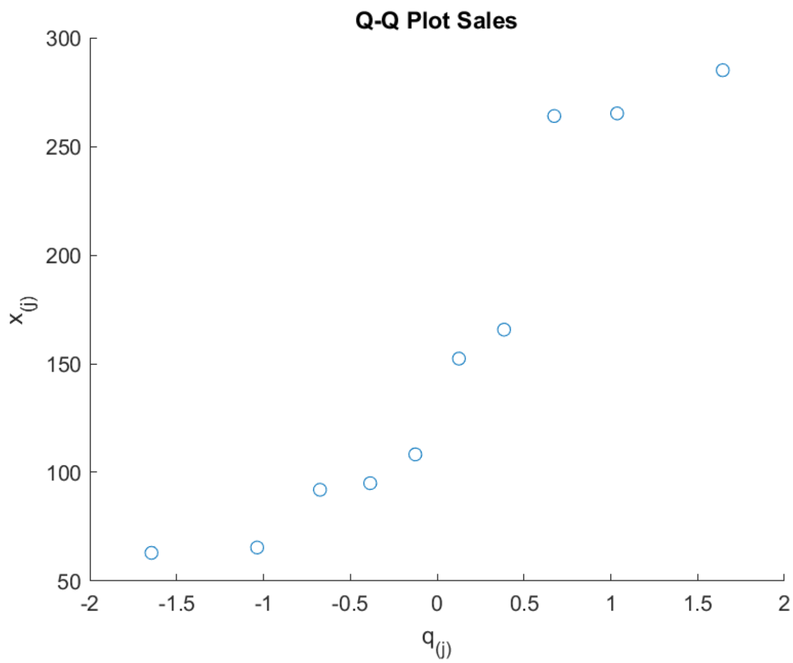
below. In the plot 10% of the data lies below the red dashed vertical line at the value 9295314.

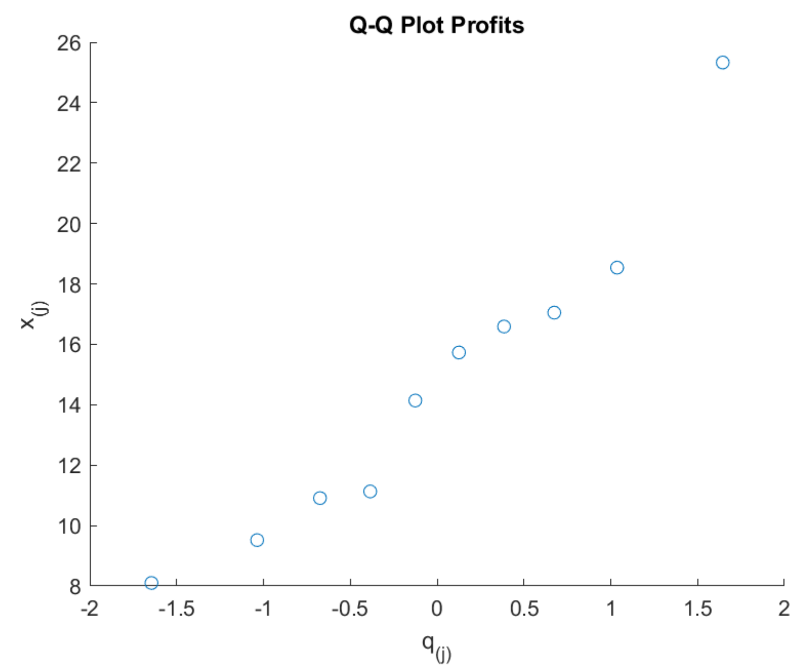


4.24

Exercise 1.4 contains data on three variables for the world's 10 largest companies as of April 2005. For the sales (x_1) and profits (x_2) data:

(a) Construct Q-Q plots. Do these data appear to be normally distributed? Explain.


Q-Q Plot Sales
300
250
200
150
100
50
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
$x_{(j)}$
$q_{(j)}$


Q-Q Plot Profits
26
24
22
20
18
16
14
12
10
8
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
$x_{(j)}$
$q_{(j)}$

The Q-Q plot for sales does not really look linear, but the plot for profits does.

(b) Carry out a test of normality based on the correlation coefficient r_Q. [See (4–31).] Set the significance level at $\alpha = .10$. Do the results of these tests corroborate the results in Part a?

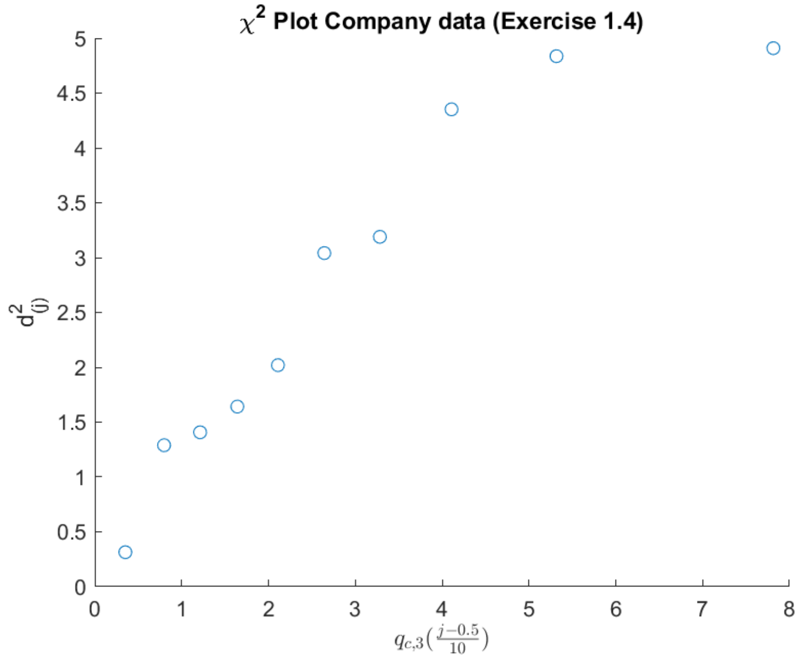
$$r_{Q,\text{sales}} \frac{\sum_{j=1}^{n} \left(x_{(j)} - \bar{x}\right) \left(q_{(j)} - \bar{q}\right)}{\sqrt{\sum_{j=1}^{n} \left(x_{(j)} - \bar{x}\right)^2} \sqrt{\sum_{j=1}^{n} \left(q_{(j)} - \bar{q}\right)^2}} = \frac{721.0797}{\sqrt{67288}\sqrt{8.7979}} = 0.9372$$

$$r_{Q,\text{profits}} \frac{\sum_{j=1}^{n} \left(x_{(j)} - \bar{x}\right) \left(q_{(j)} - \bar{q}\right)}{\sqrt{\sum_{j=1}^{n} \left(x_{(j)} - \bar{x}\right)^2} \sqrt{\sum_{j=1}^{n} \left(q_{(j)} - \bar{q}\right)^2}} = \frac{44.1345}{\sqrt{235.7128}\sqrt{8.7979}} = 0.9692$$

In the book, from Table 4.2, with a sample size $n = 10$, the significance level is 0.9351. Performing the simulation using MATLAB, the significance level for 10M simulations was 0.9345. For sales, $r_Q = 0.9372$, which is slightly larger than the 0.1 significance level values of 0.9351 and 0.9345, so the sales column would be considered normally disributed. For profits, $r_Q = 0.9692$, which is larger than the 0.1 significance level values of 0.9351 and 0.9345, so the profits column would also be considered normally disributed.

4.25

Refer to the data for the world's 10 largest companies in Exercise 1.4. Construct a chi-square plot using all *three* variables. The chi-square quantiles are
0.3518 0.7978 1.2125 1.6416 2.1095 2.6430 3.2831 4.1083 5.3170 7.8147



j	$d^2_{(j)}$	$q_{c,3}\left(\frac{j-\frac{1}{2}}{10}\right)$
1	0.3142	0.3518
2	1.2894	0.7978
3	1.4073	1.2125
4	1.6418	1.6416
5	2.0195	2.1095
6	3.0411	2.6430
7	3.1891	3.2831
8	4.3520	4.1083
9	4.8364	5.3170
10	4.9090	7.8147

4.26

Exercise 1.2 gives the age x_1, measured in years, as well as the selling price x_2, measured in thousands of dollars, for $n = 10$ used cars. These data are reproduced as follows:

x_1	1	2	3	3	4	5	6	8	9	11
x_2	18.95	19.00	17.95	15.54	14.00	12.95	8.94	7.49	6.00	3.99

(a) Use the results of Exercise 1.2 to calculate the squared statistical distances

$(\mathbf{x}_j - \bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}_j - \bar{\mathbf{x}})$, $j = 1, 2, \ldots, 10$, where $\mathbf{x}_j' = [x_{j1}, x_{j2}]$.

$$\bar{\mathbf{x}} = \begin{bmatrix} 5.2 \\ 12.4810 \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} 10.6222 & -17.7102 \\ -17.71102 & 30.8544 \end{bmatrix} \Rightarrow \mathbf{S}^{-1} = \begin{bmatrix} 2.1898 & 1.2569 \\ 1.2569 & 0.7539 \end{bmatrix}$$

$$\mathbf{d}^2 = (\mathbf{x}_j - \bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}_j - \bar{\mathbf{x}}) =$$

$$= \begin{bmatrix} x_{j1} - 5.2000 \\ x_{j2} - 12.4810 \end{bmatrix}' \begin{bmatrix} 2.1898 & 1.2569 \\ 1.2569 & 0.7539 \end{bmatrix} \begin{bmatrix} x_{j1} - 5.2000 \\ x_{j2} - 12.4810 \end{bmatrix}$$

We could compute all $j = 1, 2, \ldots, 10$ values of $\mathbf{d}^2$ at once. Here, $\mathbf{x}_j$ is $p \times 1 = 2 \times 1$ vector, putting them all into a matrix, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{10}]$ is $p \times n = 2 \times 10$. If we compute

$$\Big(\underset{(2\times10)}{\mathbf{X}} - \underset{(2\times1)}{\bar{\mathbf{x}}}\underset{(1\times10)}{\mathbf{1}}\Big)' \underset{(2\times2)}{\mathbf{S}^{-1}} \Big(\underset{(2\times10)}{\mathbf{X}} - \underset{(2\times1)}{\bar{\mathbf{x}}}\underset{(1\times10)}{\mathbf{1}}\Big)$$

directly, the output is a 10×10 matrix where the squared distances are on the diagonal. We can save ourselves from the unnecessary multiplication of the off-diagonal elements by using element-wise multiplication with $\odot$:

$$\mathbf{d}^2 = (\mathbf{x}_j - \bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}_j - \bar{\mathbf{x}}) =$$

$$= \operatorname{diag}\left\{(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1})'\mathbf{S}^{-1}(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1})\right\} = ((\mathbf{X} - \bar{\mathbf{x}}\mathbf{1})'\mathbf{S}^{-1}) \odot (\mathbf{X} - \bar{\mathbf{x}}\mathbf{1})' =$$

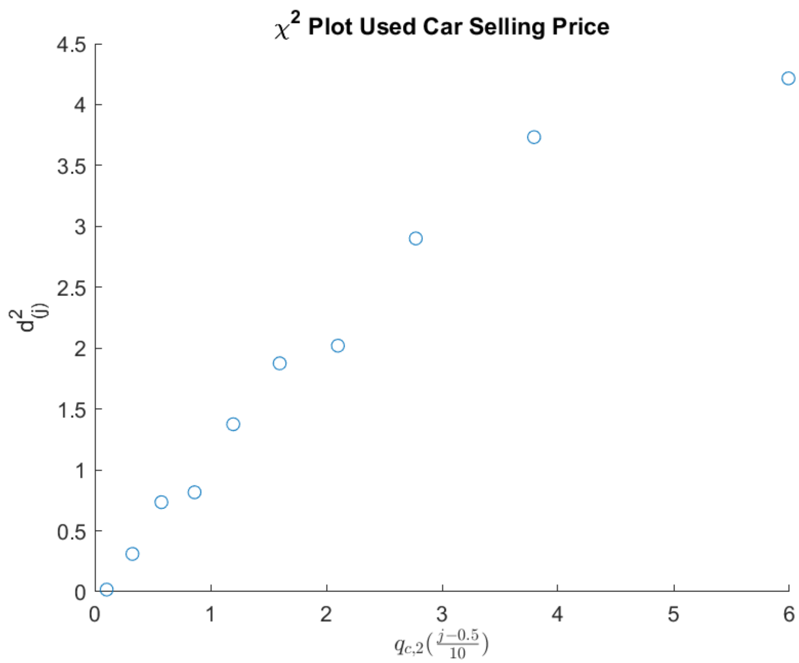
$$= \left(\begin{bmatrix} -4.2 & 6.4690 \\ -3.2 & 6.5190 \\ -2.2 & 5.4690 \\ -2.2 & 3.0590 \\ -1.2 & 1.5190 \\ -0.2 & 0.4690 \\ 0.8 & -3.5410 \\ 2.8 & -4.9910 \\ 3.8 & -6.4810 \\ 5.8 & -8.4910 \end{bmatrix} \begin{bmatrix} 2.1898 & 1.2569 \\ 1.2569 & 0.7539 \end{bmatrix} \right) \odot \begin{bmatrix} -4.2 & 6.4690 \\ -3.2 & 6.5190 \\ -2.2 & 5.4690 \\ -2.2 & 3.0590 \\ -1.2 & 1.5190 \\ -0.2 & 0.4690 \\ 0.8 & -3.5410 \\ 2.8 & -4.9910 \\ 3.8 & -6.4810 \\ 5.8 & -8.4910 \end{bmatrix} =$$

$$= \begin{bmatrix} -1.0661 & -0.4023 \\ 1.1866 & 0.8924 \\ 2.0566 & 1.3577 \\ -0.9726 & -0.4591 \\ -0.7185 & -0.3632 \\ 0.1515 & 0.1022 \\ -2.6990 & -1.6640 \\ -0.1419 & -0.2432 \\ 0.1751 & -0.1096 \\ 2.0282 & 0.8890 \end{bmatrix} \odot \begin{bmatrix} -4.2 & 6.4690 \\ -3.2 & 6.5190 \\ -2.2 & 5.4690 \\ -2.2 & 3.0590 \\ -1.2 & 1.5190 \\ -0.2 & 0.4690 \\ 0.8 & -3.5410 \\ 2.8 & -4.9910 \\ 3.8 & -6.4810 \\ 5.8 & -8.4910 \end{bmatrix} = \begin{bmatrix} 1.8753 \\ 2.0203 \\ 2.9009 \\ 0.7353 \\ 0.3105 \\ 0.0176 \\ 3.7329 \\ 0.8165 \\ 1.3753 \\ 4.2153 \end{bmatrix}$$

(b) Using the distances in Part a, determine the proportion of the observations falling within the estimated 50% probability contour of a bivariate normal distribution.

The Chi-square value where 50% of the data is inside is found using the inverse CDF of a Chi-squared with $p = 2$ degrees of freedom, $F^{-1}(\text{prob} = 0.50|p = 2) = 1.3863$. Using the output from (a), there are 5 values less than 1.3863. Out of 10 total observations that makes 0.50 of the data. This is exactly what we'd expect under normality.

(c) Order the distances in Part a and construct a chi-square plot.



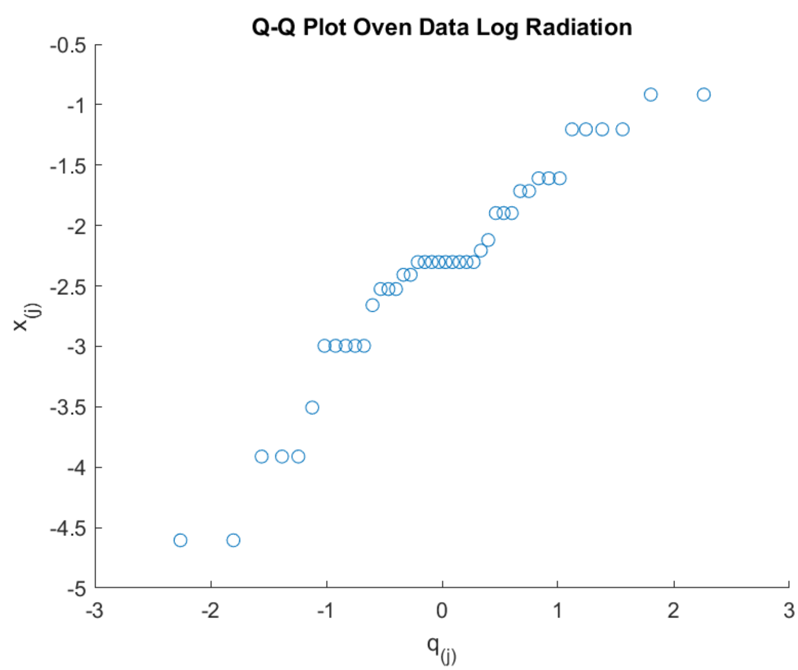
j	$d^2_{(j)}$	$q_{c,2}\left(\frac{j-\frac{1}{2}}{10}\right)$
1	0.0176	0.1026
2	0.3105	0.3250
3	0.7353	0.5754
4	0.8165	0.8616
5	1.3753	1.1957
6	1.8753	1.5970
7	2.0203	2.0996
8	2.9009	2.7726
9	3.7329	3.7942
10	4.2153	5.9915

(d) Given the results in Parts b and c, are these data approximately bivariate normal? Explain.

In part (b) we found that 50% of our data is within the contour of what we would expect if the data was normally distributed. In part (c), the χ^2-plot shows our data to be roughly linear. For these reasons we could conclude that there is evidence that used car sale data is normally distributed.

4.27

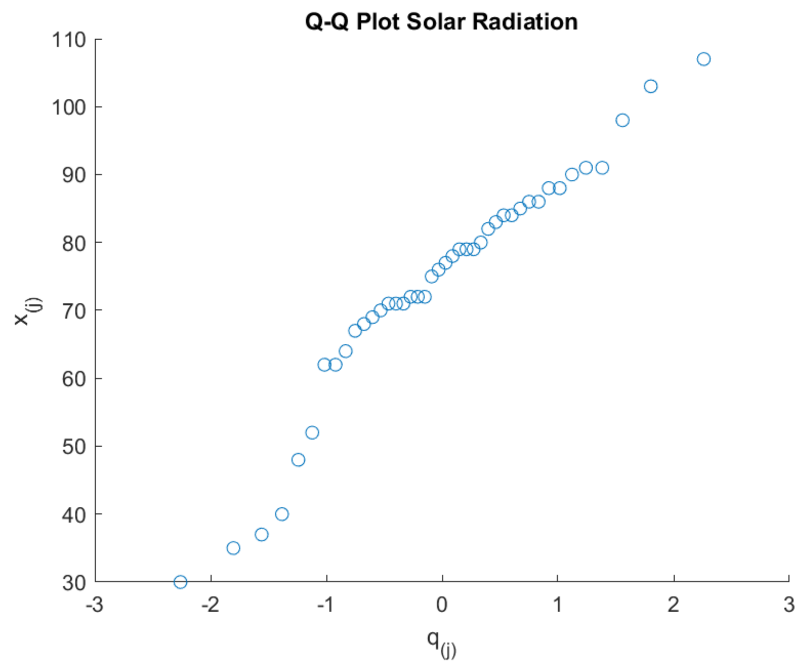
Consider the radiation data (with door closed) in Example 4.10. Construct a *Q-Q* plot for the natural logarithms of these data. [Note that the natural logarithm transformation corresponds to the value $\lambda = 0$ in (4–34).] Do the natural logarithms appear to be normally distributed? Compare your results with Figure 4.13. Does the choice $\lambda = \frac{1}{4}$ or $\lambda = 0$ make much difference in this case?



The Q-Q plot of the natural log transformed data does appear to be linear, so the data would be considered normally distributed after transformation. The plot here, when $\lambda = 0$, looks very similar to the one in Figure 4.13 when $\lambda = \frac{1}{4}$. The plot when $\lambda = \frac{1}{4}$ stretches things verically slightly compared to the log-transformation. I'd guess that either λ value is a fine choice.

4.28

Consider the air-pollution data given in Table 1.5. Construct a Q-Q plot for the solar radiation measurements and carry out a test for normality based on the correlation coefficient r_Q [see (4–31)]. Let $\alpha = .05$ and use the entry corresponding to $n = 40$ in Table 4.2.



The correlation coefficient for the data is $r_Q = \mathrm{Corr}(q_{(j)}, x_{(j)}) = 0.9693$. The value in Table 4.2 when n=40 and $\alpha = 0.05$ is 0.9726. Our value of R_Q is less than the value in the table, so we'd conclude that the data is normally distributed.

4.29

Given the air-pollution data in Table 1.5, examine the pairs $X_5 = NO_2$ and $X_6 = O_3$ for bivariate normality.

(a) Calculate statistical distances $(\mathbf{x}_j - \bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}_j - \bar{\mathbf{x}})$, $j = 1, 2, \ldots, 42$, where $\mathbf{x}_j' = [x_{j5}, x_{j6}]$.

$$\bar{\mathbf{x}} = \begin{bmatrix} 10.0476 \\ 9.4048 \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} 11.3635 & 3.1266 \\ 3.1266 & 30.9785 \end{bmatrix} \Rightarrow \mathbf{S}^{-1} = \begin{bmatrix} 0.0905 & -0.0091 \\ -0.0091 & 0.0332 \end{bmatrix}$$

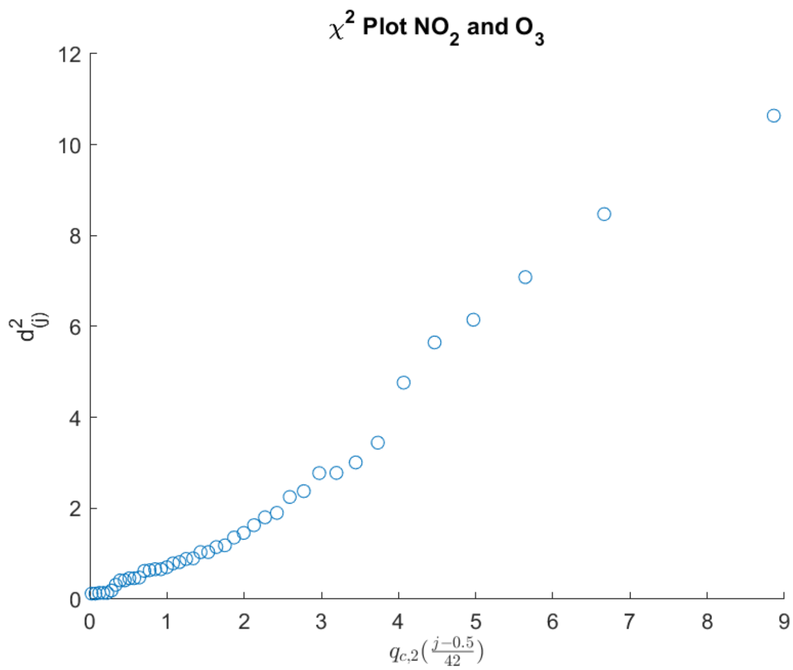
$$(\mathbf{x}_j - \bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}_j - \bar{\mathbf{x}}) =$$

$$= \begin{bmatrix} x_{j5} - 10.0476 \\ x_{j6} - 9.4048 \end{bmatrix} \begin{bmatrix} 0.0905 & -0.0091 \\ -0.0091 & 0.0332 \end{bmatrix} \begin{bmatrix} x_{j5} - 10.0476 \\ x_{j6} - 9.4048 \end{bmatrix}$$

(b) Determine the proportion of observations $\mathbf{x}'_j = [x_{j5}, x_{j6}]$, $j = 1, 2, \ldots, 42$, falling within the approximate 50% probability contour of a bivariate normal distribution.

The Chi-square value where 50% of the data is inside is found using the inverse CDF of a Chi-squared with $p = 2$ degrees of freedom, $F^{-1}(\text{prob} = 0.50|p = 2) = 1.3863$. For this data, 26 of the 42 $\mathbf{d}^2$ observations are equal to, or less than, 1.3863, so 61.90%.

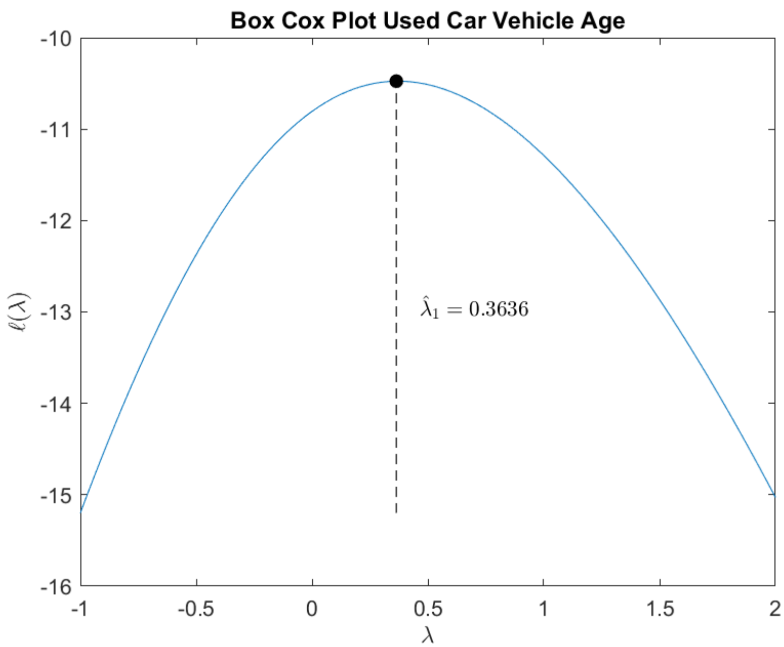
(c) Construct a chi-square plot of the ordered distances in Part a.



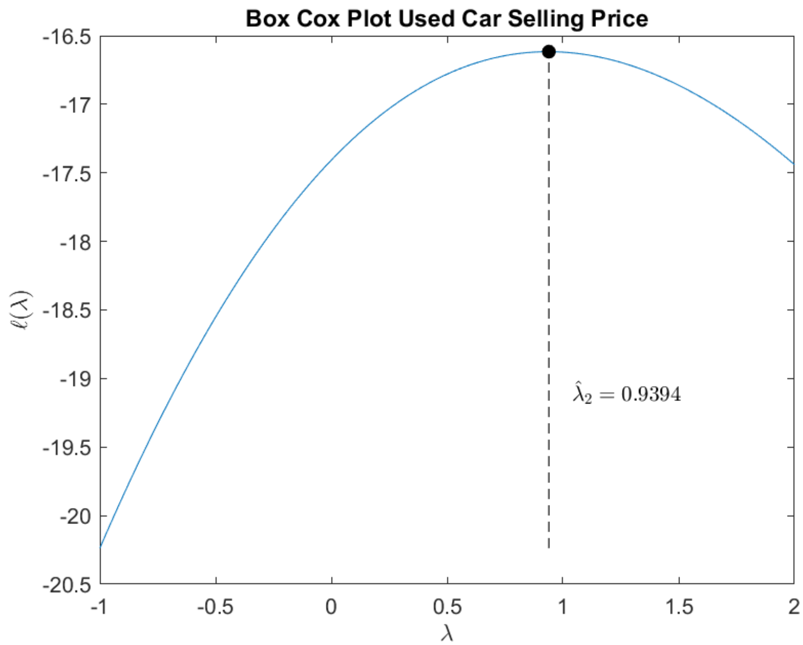
4.30

Consider the used-car data in Exercise 4.26.

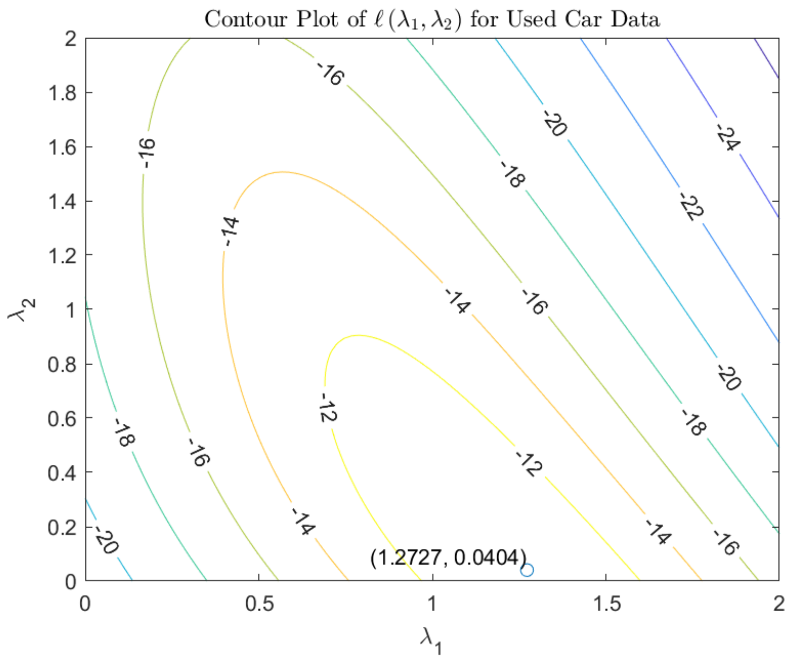
(a) Determine the power transformation $\hat{\lambda}_1$ that makes the x_1 values approximately normal. Construct a Q-Q plot for the transformed data.

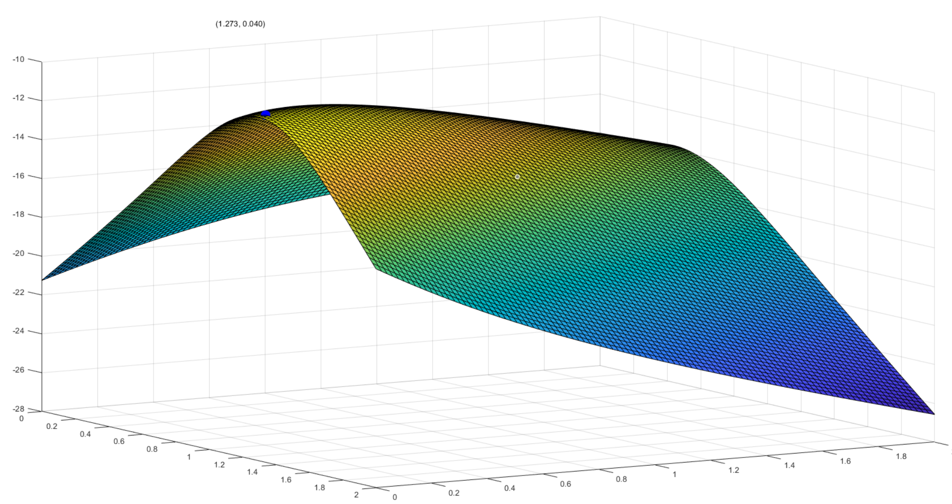


(b) Determine the power transformation $\hat{\lambda}_2$ that makes the x_2 values approximately normal. Construct a Q-Q plot for the transformed data.



(c) Determine the power transformations $\hat{\boldsymbol{\lambda}}' = [\hat{\lambda}_1, \hat{\lambda}_2]$ that make the $[x_1, x_2]$ values jointly normal using (4–40). Compare the results with those obtained in Parts a and b.





Using the joint distribution of x_1 and x_2, the simultaneous maximum for the bestp̈ower transformation was found at $\hat{\boldsymbol{\lambda}}' = [\hat{\lambda}_1, \hat{\lambda}_2] = [1.2727, 0.0404]$. That is, for vehicle age was $\hat{\lambda}_1 = 1.2727$, and for sale price was $\hat{\lambda}_2 = 0.0404$.

These are a bit different from parts a and b where we were maximizing the univariate likelihood. In those cases the maximum for vehicle age was $\hat{\lambda}_1 = 0.3636$ and for vehicle price was $\hat{\lambda}_2 = 0.9394$.

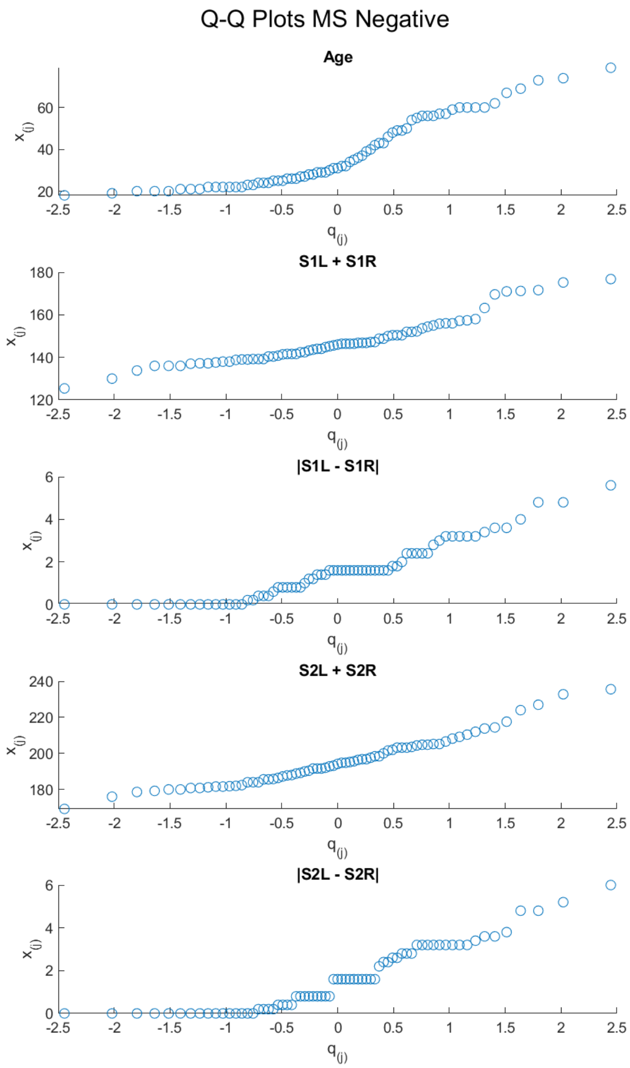
4.31

Examine the marginal normality of the observations on variables $X_1, X_2, \ldots, X_5$ for the multiple-sclerosis data in Table 1.6. Treat the non-multiple-sclerosis and multiple-sclerosis groups separately. Use whatever methodology, including transformations, you feel is appropriate.
There are 10 vectors that need to be checked and I'll pick an α-level of 0.01 for our tests.

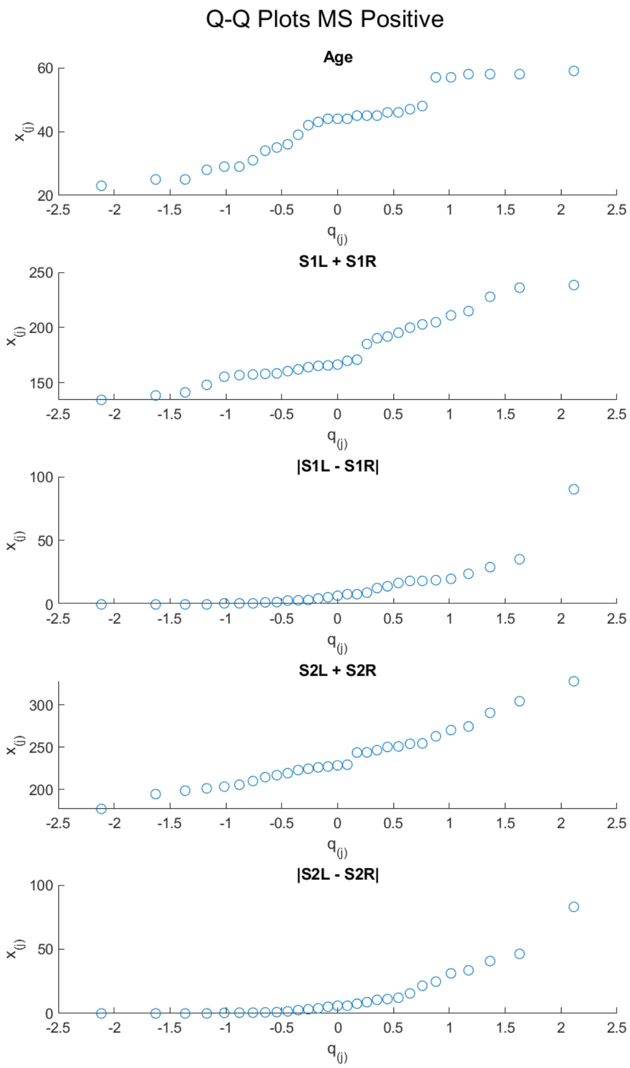
x_i	MS Negative r_Q	MS Positive r_Q
Age	0.9448	0.9714
$S1L + S1R$	0.9613	0.9703
$\lvert S1L - S1R \rvert$	0.9558	0.7952
$S2L + S2R$	0.9757	0.9787
$\lvert S2L - S2R \rvert$	0.9445	0.8413

The correlation coefficients for the Q-Q plots for all the variables are in the table above. The MS positive group has low r_Q values for variables $|S1L - S1R|$ and $|S2L - S2R|$. For MS negative, variable $|S2L - S2R|$ has the lowest r_Q value within group.

From Table 4.2 for the negative MS group the sample size of 69 and significance level $\alpha = 0.01$, the critical point would be somewhere within 0.9720 and 9771. Instead of relying on the table, performing 1M simulations, the value found is 0.9752. Based on this value, only the variable $S2L + S2R$ for the negative MS group would be considered normally distributed. Variables x_1, x_2, x_3, and x_5 are in need of a transformation. The Q-Q plots for each of the five variables is below.



For the positive MS group the sample size is 29, so from Table 4.2 where the significance level is 0.01, the critical point is slightly less than 0.9479, but larger than 0.9410. Performing the simulation for sample size 29 and 1M replications, the critical point found was 0.9475. In this case, only x_3 and x_5, that is, $|S1L - S1R|$ and $|S2L - S2R|$, would be considered not to be normally distributed and need a transformation. The Q-Q plots for the variables in the MS positive group are below.



Univariate Power Transformation

x_i	MS Negative $\max\{\ell(\lambda)\}$	MS Positive $\max\{\ell(\lambda)\}$
Age	-0.4910	—
$S1L + S1R$	-3.5371	—
$\lvert S1L - S1R \rvert$	0.2305	0.2104
$S2L + S2R$	—	—
$\lvert S2L - S2R \rvert$	0.1904	0.1904

The results of performing the univariate power transformation are in the table above. I left the power transformation plots out, but copies can be found in the `matlab` folder. For the MS Negative group, rounding things off, it looks like Age could use a negative square-root transformation. The variable $S1L + S1R$ could use a transformation of -3.5. Variable, $|S1L - S1R|$, rounds to 1/4 and 1/5 for $|S2L - S2R|$. For the MS positive data, again, $|S1L - S1R|$ is transformed using a power of 1/4 and $|S2L - S2R|$ a power of 1/5.

Q-Q Correlation Transformed Data

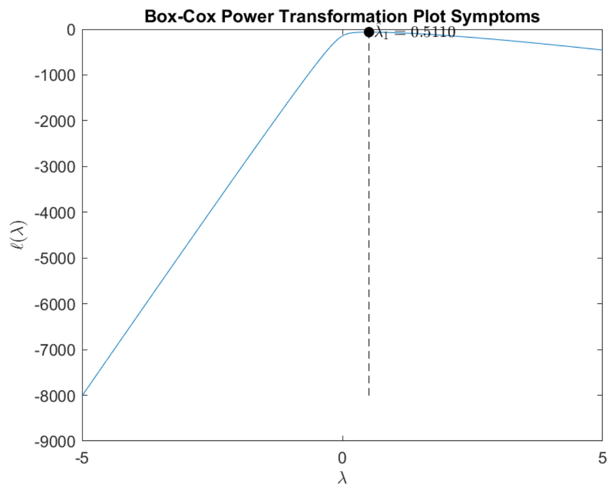
x_i	MS Negative Transformed r_Q	MS Positive Transformed r_Q
Age	0.9694	—
$S1L + S1R$	0.9878	—
$\|S1L - S1R\|$	0.8888	0.9679
$S2L + S2R$	—	—
$\|S2L - S2R\|$	0.8852	0.9651

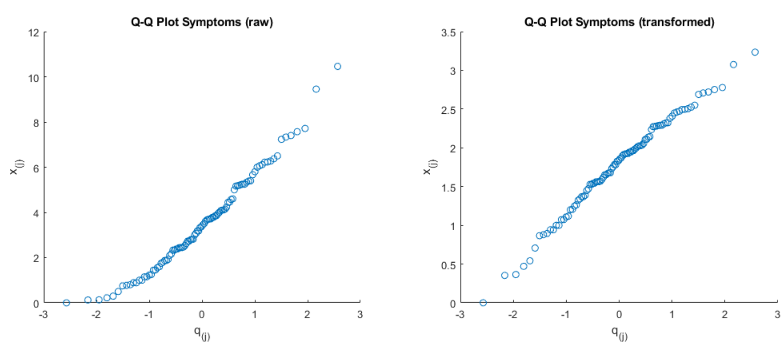
Computing the Q-Q values on the transformed data seems to have normalized the MS positive group columns $|S1L - S1R|$ and $|S2L - S2R|$. Both values are greater than the critical point value of 0.9475. For the MS negative group, the variable $S1L + S1R$ looks okay now, and has a Q-Q correlation value (0.9878) larger than the 0.01-level critical value of 0.9752. For the MS negative group, transforming variable Age was helpful, but the Q-Q correlation value (0.9694) is still not quite larger than the critical value. Lastly, for the MS negative group, transforming variables $|S1L - S1R|$ and $|S2L - S2R|$, the results are worse than the raw data. This might be due to the number of 0 values. For these two variables the MS negative group has 20% to 25% zero values, but the MS positive group it's around 13%, so people without MS have less of an absolute difference between left and right eye than people with MS. Maybe instead we could use an factor variable to indicate if the absolute difference is 0 or not.

4.32

Examine the marginal normality of the observations on variables $X_1, X_2, \ldots, X_6$ for the radiotherapy data in Table 1.7. Use whatever methodology, including transformations, you feel is appropriate.

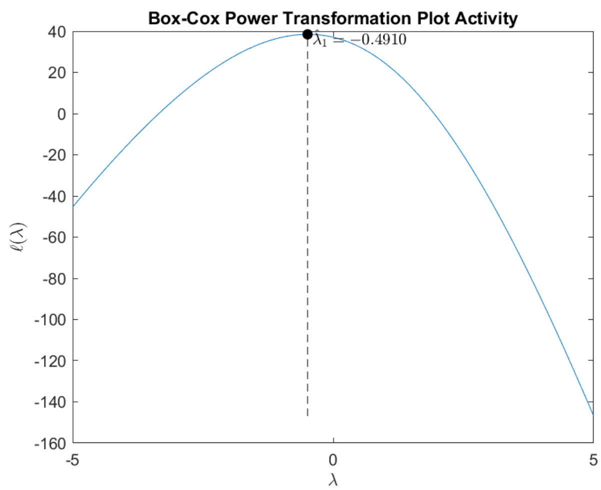
For x_1, the variable symptoms. The Q-Q correlation coefficient using the raw data was 0.9871, but the critical value found at the 0.01, 0.05, and 0.10 levels by performing 1M simulation were, respectively, 0.9819, 0.9871, and 0.9893, so the data would not be considered normally distributed at the 0.05 and 0.10 levels. The transformation suggested by the power transformation was 0.5110, but was rounded to 0.5, so $x_1' = \sqrt{x_1}$. The Q-Q correlation coefficient on the transformed data was 0.9942, which is larger than all the critical values, so the data is now normally distributed at the 0.01, 0.05, and 0.10 levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original plot wasn't too bad, but the transformed data has brought in the larger values and straightened out lower values to make the plot much more linear.

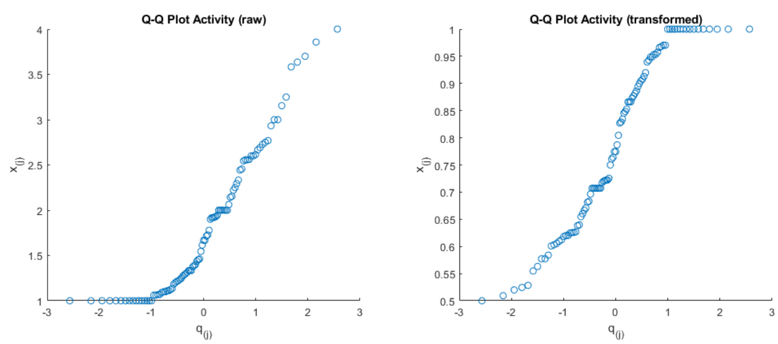




For activity (x_2), we're measuring patient activity on a continuous scale from 1 to 5, and have 97 valid observations. The Q-Q correlation on the raw data was 0.9455. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 97 are, 0.9818, 0.9870, and 0.9892, respectively. Our Q-Q correlation value (0.9455) is smaller than all of these, so out data is not considered normally distributed at any of the 3 levels.

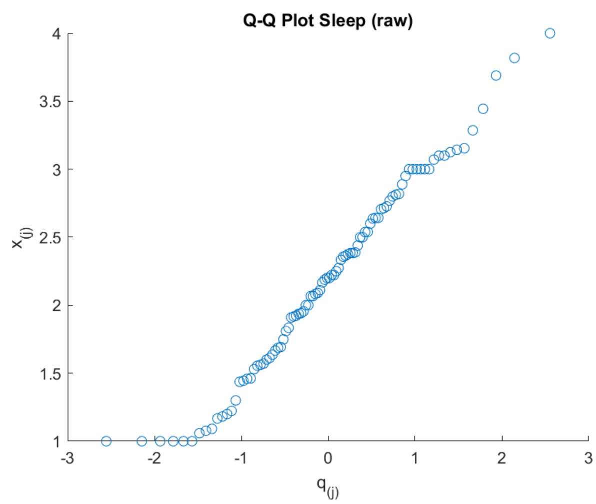
The Box-Cox power transformation is maximized at -0.4910, but I chose to round it to -0.50, so $x_2' = 1/\sqrt{x_2}$. The Q-Q correlation coefficient on the transformed data was increased to 0.9630. This value is still smaller than the critical values at all three levels, but hey, it is an improvement. I'd say the problem with this variable is repeated values. We're measuring patient activity on a continous scale from 1 to 5. There are 16 of the 97 patiens (16.5%) with an activity level of 1. These 16 values are creating a flat area in the Q-Q plot that cannot be fixed by any common transformation.





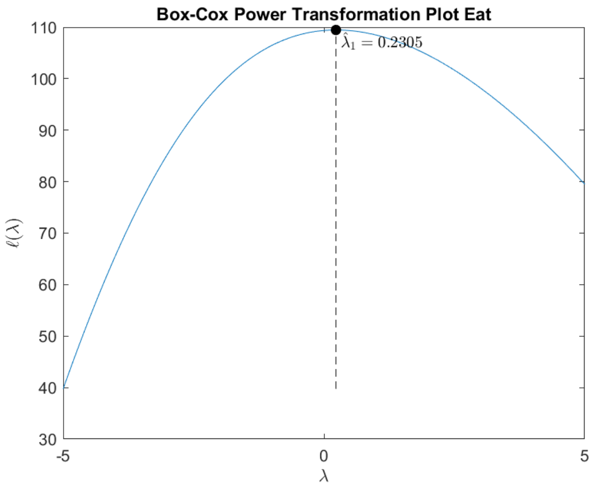
For sleep (x_3), we're measuring patient sleep on a continuous scale from 1 to 5, and have 94 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 94 are, 0.9812, 0.9867, and 0.9889, respectively. Our Q-Q correlation value of 0.9895 is larger than all of these, so out data is actually considered normally distributed at all 3 levels.

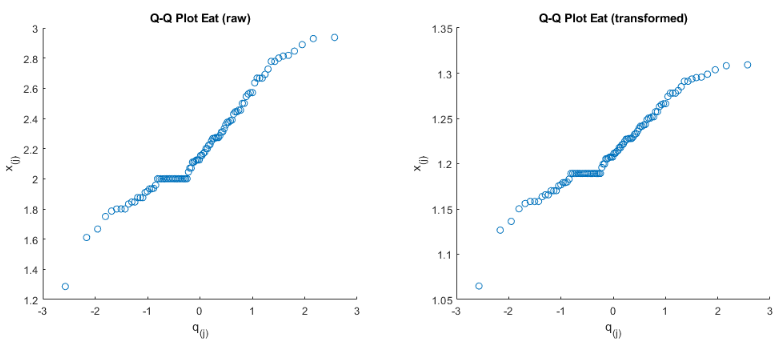
Because of this I won't really bolther displaying the Box-Cox transformation results, but the power transformation of 0.7114 (close to 1) bumps the Q-Q correlation coefficient up slightly to 0.9893. The raw data Q-Q plot is below.



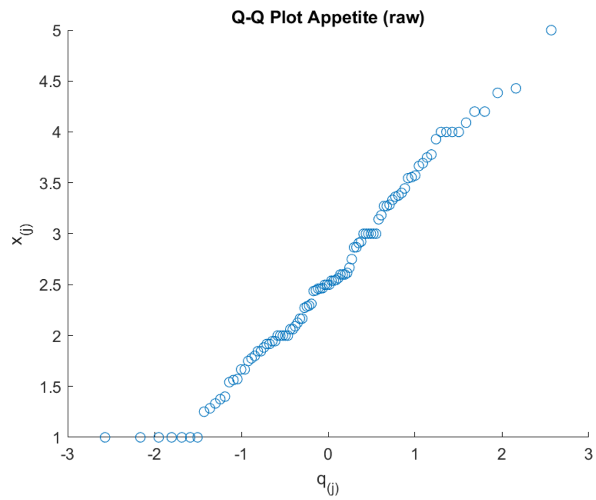
For eat (x_4), we're measuring amount of food consumed on a continuous scale from 1 to 3, and have 98 valid observations. The Q-Q correlation on the raw data was 0.9810. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 98 are, 0.9819, 0.9871, and 0.9893, respectively. Our Q-Q correlation value (0.9810) is smaller than all of these, so out data is not considered normally distributed at any of the 3 levels.

The Box-Cox power transformation is maximized at 0.2305, but I chose to round it to 0.25, so $x_4' = \sqrt[4]{x_4}$. The Q-Q correlation coefficient on the transformed data was increased slightly to 0.9834. This value is larger than the critical value at the 0.01-level (0.9819), but still smaller than the critical values at 0.05 and 0.01 levels. In this case our transformation has improved the normality of our data, if only slightly. I should also note that most of the data is tightly clustered, with the exception of observation 13, who has an eat value lower than the others with value 1.2826 and standardized residual of -2.8.





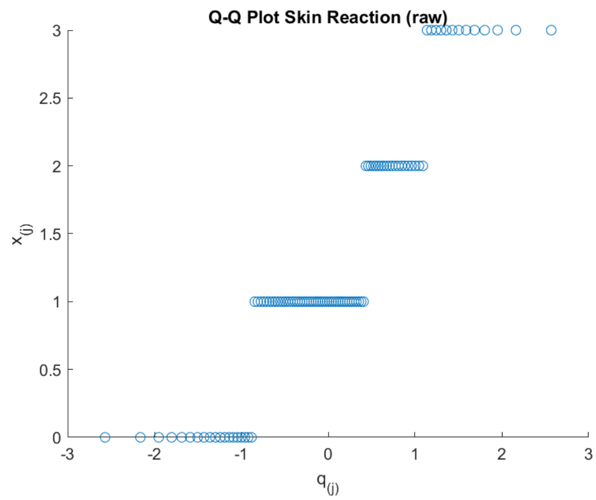
For appetite (x_5), we're measuring patient appetite on a continuous scale from 1 to 5, and have 98 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 94 are, 0.9819, 0.9871, and 0.9893, respectively. Our Q-Q correlation value of 0.9906 is larger than all of these, so out data is actually considered normally distributed at all 3 levels. Because of this I won't really bolther displaying the Box-Cox transformation results, but the power transformation of 0.6112 bumps the Q-Q correlation coefficient up slightly to 0.9924. The raw data Q-Q plot is below.



For skin reaction (x_6), we're measuring amount of skin reaction on a discrete scale from 0 to 3, and have 98 valid observations. The problem here is that the daata is discrete, so I would think a transformation might not be appropriate since this would be considered a factor varaible. I'll do the work anyway just to see what happens.

The Q-Q correlation on the raw data was 0.9278. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 98 are, 0.9819, 0.9871, and 0.9893, respectively. Our Q-Q correlation value (0.9278) is smaller than all of these, so out data is not considered normally distributed at any of the 3 levels.

The Box-Cox power transformation is maximized at 0.2705, so $x_6' = x_6^{0.2705}$. The Q-Q correlation coefficient on the transformed data was decreased to 0.8371. This value is worse than the one for the raw data and so still not larger than any of the 3 critical points. The Q-Q plot on the raw data is below. Just like for the activity variable we have lots of flat spots, so it's unlikely a common transformation will increase the Q-Q correlation value high enough to be larger than any of the critical points. It probably best to use the raw data for analysis or use it as a factor variable.

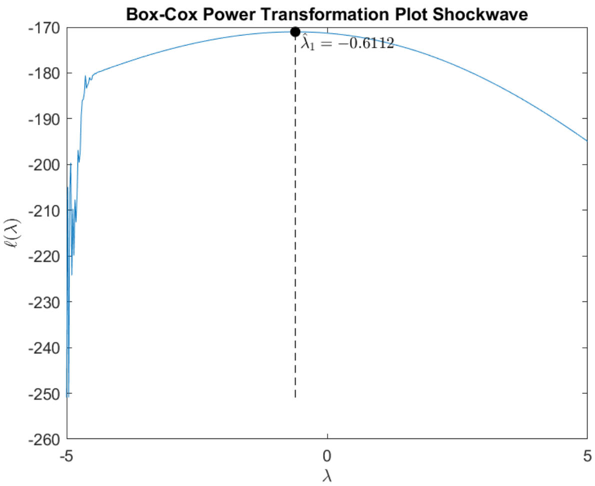


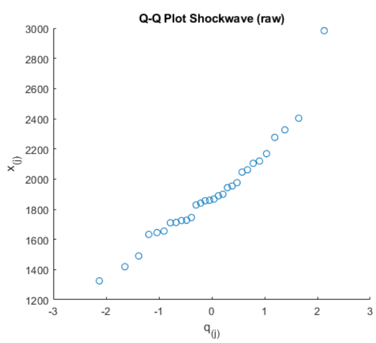
4.33

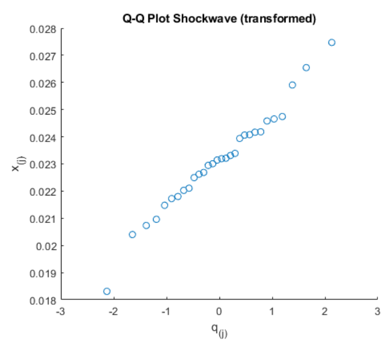
Examine the marginal and bivariate normality of the observations on variables X_1, X_2, X_3, and X_4 for the data in Table 4.3.

For (x_1), we're sending a shockwave down a board and measuring the board stiffness, and have 30 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 30 are, 0.9488, 0.9647, and 0.9711, respectively. The Q-Q correlation coefficient using the raw data was 0.9599, so the data would not be considered normally distributed at the 0.05 and 0.10 levels.

The transformation suggested by the power transformation was -0.6112, but was rounded to -0.5, so $x_1' = 1/\sqrt{x_1}$. The Q-Q correlation coefficient on the transformed data was 0.9873, which is larger than all the critical values, so the data is now normally distributed at the 0.01, 0.05, and 0.10 levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original plot wasn't too bad, with one value with a higher stiffness value than the rest, but the transformed data has brought in the larger value to make the plot much more linear.

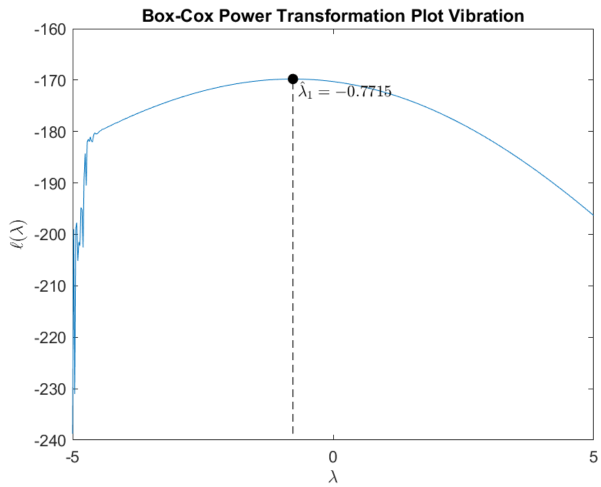


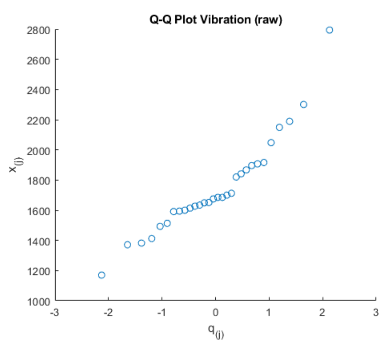


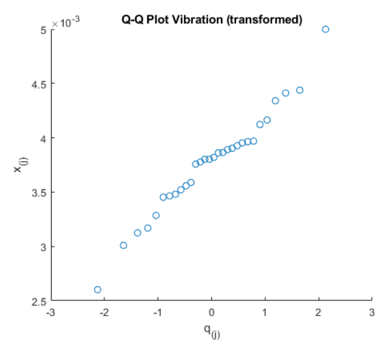


For (x_2), we're determining the stiffness of a bord while it's being vibrated, and have 30 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 30 are, 0.9488, 0.9647, and 0.9711, respectively. The Q-Q correlation coefficient using the raw data was 0.9504, so the data would not be considered normally distributed at the 0.05 and 0.10 levels.

The transformation suggested by the Box-Cox power transformation was -0.7715, but was rounded to -0.75, so $x_2' = 1/\sqrt[4]{x_2^3}$. The Q-Q correlation coefficient on the transformed data was 0.9828, which is larger than all the critical values, so the data is now normally distributed at the 0.01, 0.05, and 0.10 levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original plot wasn't too bad, with one value with a higher stiffness value than the rest, just like for x_1, but the transformed data has brought in the larger value to make the plot much more linear.

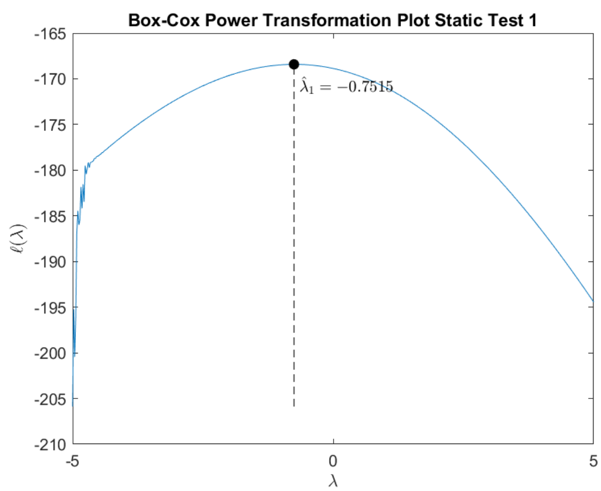


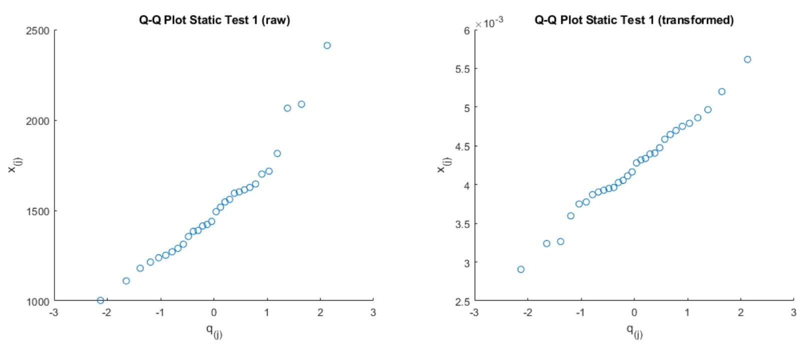




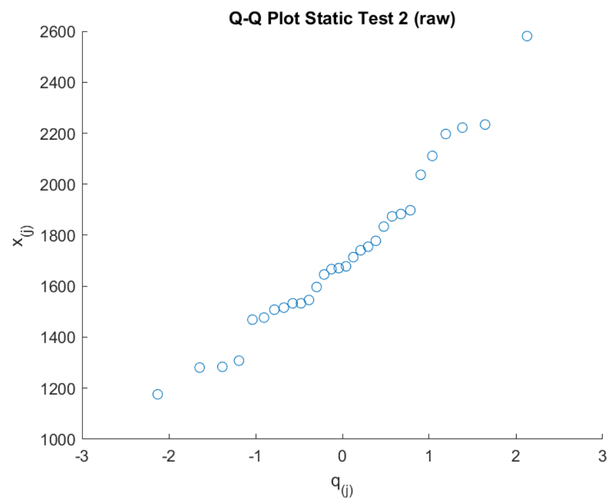
For (x_3), we're determining the stiffness of a bord in a static test, and have 30 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 30 are, 0.9488, 0.9647, and 0.9711, respectively. The Q-Q correlation coefficient using the raw data was 0.9634, so the data would not be considered normally distributed at the 0.05 and 0.10 levels.

The transformation suggested by the Box-Cox power transformation was -0.7515, but was rounded to -0.75, so $x_3' = 1/\sqrt[4]{x_3^3}$. The Q-Q correlation coefficient on the transformed data was 0.9828, which is larger than all the critical values, so the data is now normally distributed at the 0.01, 0.05, and 0.10 levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original plot wasn't too bad, with a few values with higher stiffness value than the rest, similar to x_1 and x_2, but the transformed data has brought in the larger value to make the plot much more linear.





For (x_4), we're determining the stiffness of a bord in a static test, again, and have 30 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 30 are, 0.9488, 0.9647, and 0.9711, respectively. Our Q-Q correlation value of 0.9803 is larger than all of these, so out data is actually considered normally distributed at all 3 levels. Because of this I won't really bolther displaying the Box-Cox transformation results, but the power transformation of -0.3507 bumps the Q-Q correlation coefficient up to 0.9928. The raw data Q-Q plot is below.



Checking out the bivariate, we have $\binom{4}{2} = 6$ pairs of variables to evaluate for

$$\begin{Bmatrix} (x_1, x_2), & (x_1, x_3), & (x_1, x_4), \\ (x_2, x_3), & (x_2, x_4), & \\ (x_3, x_4) & & \end{Bmatrix}$$

Proportion less than $\chi^2_2(0.50) = 1.3863$

Variables	Raw Data	Transformed Data
(x_1, x_2)	0.5333	0.5000
(x_1, x_3)	0.6333	0.6000
(x_1, x_4)	0.6000	0.6000
(x_2, x_3)	0.7000	0.6333
(x_2, x_4)	0.7000	0.6667
(x_3, x_4)	0.5667	0.6000

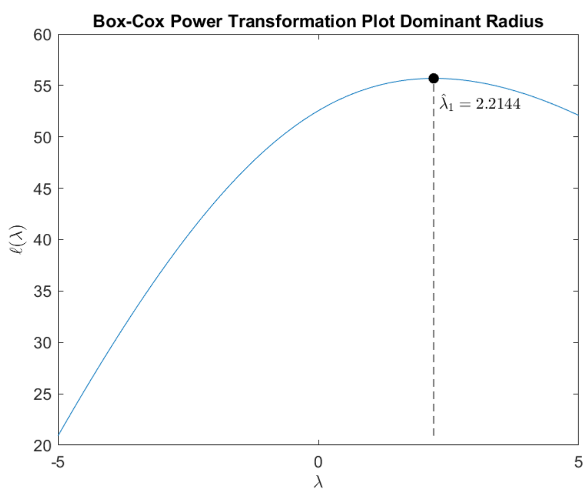
Looking at the table above, for the raw data, some of the pairs of variables look okay, like x_1 and x_2 who has a proportion of 0.53, which is close to the 0.50 we'd expect for bivariate normal data. Some raw data proportions, like x_2 and x_3 or x_2 and x_4 are fairly high proportions of 0.70 when we'd expect to see 0.5. After transforming the data, most of the proportions of pairs of board stiffness variables became closer to 0.5. One exception is x_3 and x_4 rose slightly to 0.6 from 0.57. Overall, transforming the data did improve bivariate normality. I probably should also note that we don't have tons of data, we only 30 observations.

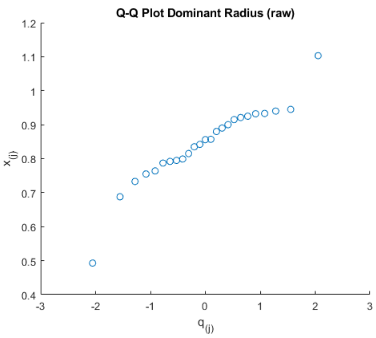
4.34

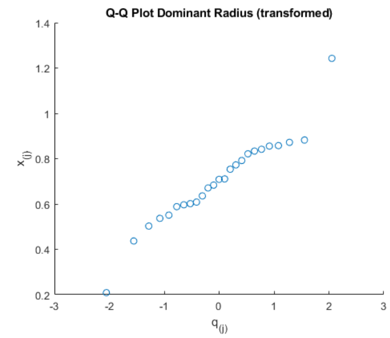
Examine the data on bone mineral content in Table 1.8 for marginal and bivariate normality.

For (x_1), we're using the dominant side of the radius bone found in the forearm to measure the mineral content via photon absorptiometry, and have 25 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 25 are, 0.9405, 0.9591, and 0.9665, respectively. The Q-Q correlation coefficient using the raw data was 0.9516, so the data would not be considered normally distributed at the 0.05 and 0.10 levels.

The transformation suggested by Box-Cox power transformation was 2.2144, so $x_1' = x_1^{2.2144}$. The Q-Q correlation coefficient on the transformed data was 0.9654, which is larger than the 0.01 and 0.05 critical values, but not the 0.10 value, so the data is now normally distributed at two of the three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original plot shows a few outlier values on both the low and high end to make things more curvilinear. The transformed data wasn't able to make the Q-Q plot absolutely linear, but has helped some.

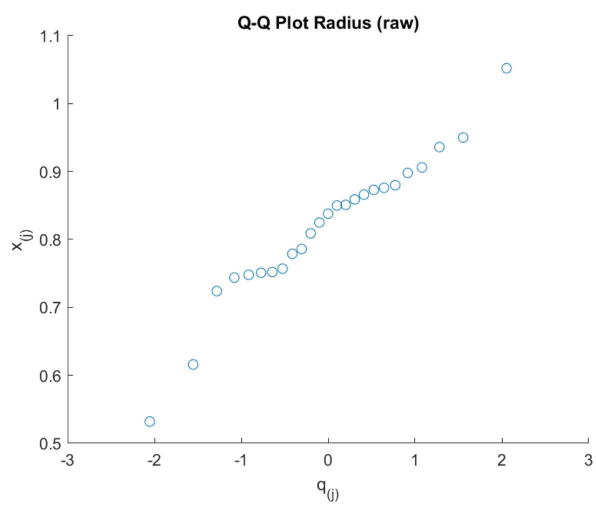






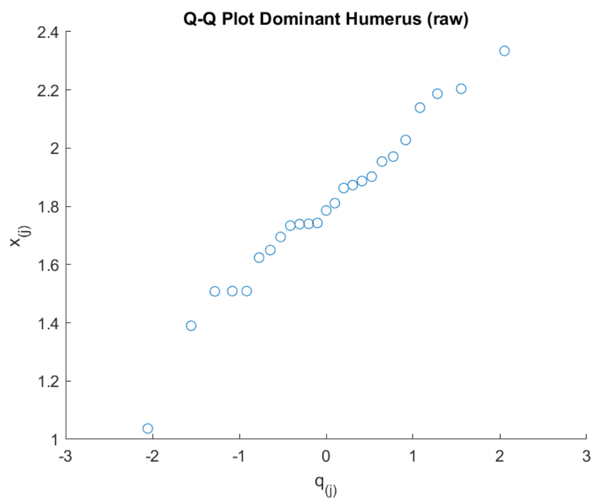
For (x_2), we're using the nondominant side of the radius bone found in the forearm to measure the mineral content via photon absorptiometry, and have 25 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 25 are, 0.9405, 0.9591, and 0.9665, respectively. The Q-Q correlation coefficient using the raw data was 0.9721, is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was 2.0942, so $x_2' = x_2^{2.0942}$. The Q-Q correlation coefficient on the transformed data was 0.9804, so we do get some improvement. Below is the Q-Q plot for the raw data, which looks pretty good.



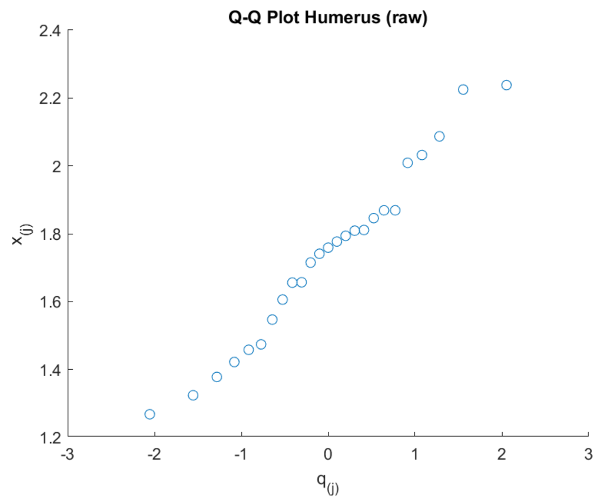
For (x_3), we're using the dominant side of the humerus bone found in the forearm to measure the mineral content via photon absorptiometry, and have 25 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 25 are, 0.9405, 0.9591, and 0.9665, respectively. The Q-Q correlation coefficient using the raw data was 0.9842, is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was 1.7535, so $x_3' = x_3^{1.7535}$. The Q-Q correlation coefficient on the transformed data was 0.9903, so we do get a nice improvement. Below is the Q-Q plot for the raw data, which looks even better than the one for x_2.



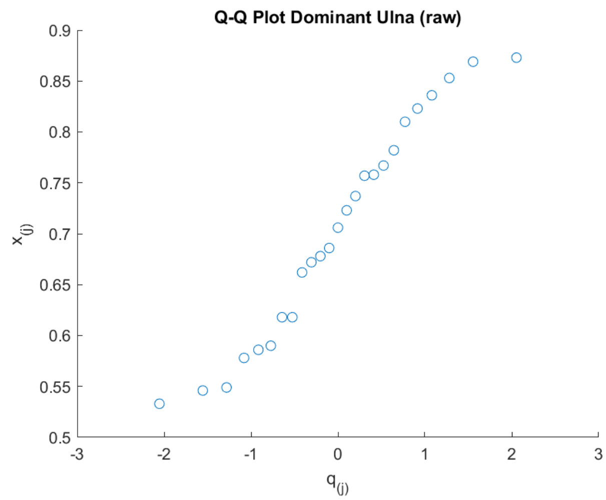
For (x_4), we're using the nondominant side of the humerus bone found in the forearm to measure the mineral content via photon absorptiometry, and have 25 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 25 are, 0.9405, 0.9591, and 0.9665, respectively. The Q-Q correlation coefficient using the raw data was 0.9901, is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at 0.5711, so $x_4' = x_4^{0.5711}$. The Q-Q correlation coefficient on the transformed data was 0.9907, so we do get a very slight improvement. Below is the Q-Q plot for the raw data, which looks nice and linear.



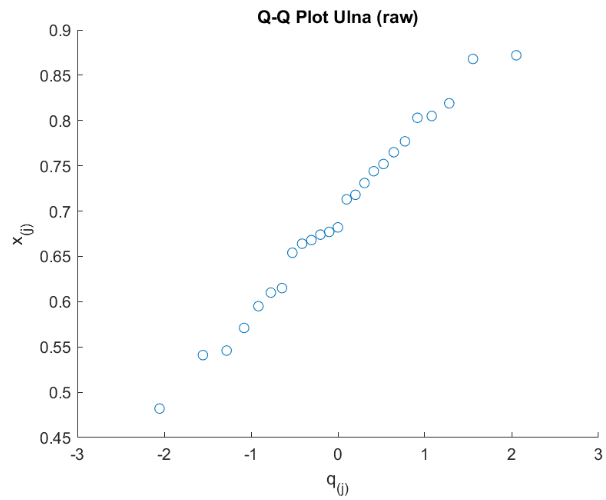
For (x_5), we're using the dominant side of the ulna bone found in the forearm to measure the mineral content via photon absorptiometry, and have 25 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 25 are, 0.9405, 0.9591, and 0.9665, respectively. The Q-Q correlation coefficient using the raw data was 0.9812, is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at 0.8317 (close-ish to 1), so $x_5' = x_5^{0.8317}$. The Q-Q correlation coefficient on the transformed data was 0.9811, so we do get a very slight improvement. Below is the Q-Q plot for the raw data, which looks good.



For (x_6), we're using the nondominant side of the ulna bone found in the forearm to measure the mineral content via photon absorptiometry, and have 25 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 25 are, 0.9405, 0.9591, and 0.9665, respectively. The Q-Q correlation coefficient using the raw data was 0.9940, is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at 1.2926 (close-ish to 1), so $x_6' = x_6^{1.2926}$. The Q-Q correlation coefficient on the transformed data was 0.9945, so we do get a very slight improvement. Below is the Q-Q plot for the raw data, which looks alright.



Checking out the bivariate, we have $\binom{6}{2} = 15$ pairs of variables to evaluate for

$$\left\{\begin{array}{lllll} (x_1,x_2), & (x_1,x_3), & (x_1,x_4), & (x_1,x_5), & (x_1,x_6), \\ (x_2,x_3), & (x_2,x_4), & (x_2,x_5), & (x_2,x_6) & \\ (x_3,x_4), & (x_3,x_5), & (x_3,x_6) & & \\ (x_4,x_5), & (x_4,x_6), & & & \\ (x_5,x_6) & & & & \end{array}\right\}$$

Proportion less than $\chi_2^2(0.50) = 1.3863$

Variables	Raw Data	Transformed Data
(x_1, x_2)	0.7200	0.6800
(x_1, x_3)	0.6000	0.6000
(x_1, x_4)	0.6000	0.5600
(x_1, x_5)	0.6000	0.5200
(x_1, x_6)	0.6000	0.6000
(x_2, x_3)	0.7200	—
(x_2, x_4)	0.5600	—
(x_2, x_5)	0.5200	—
(x_2, x_6)	0.6000	—
(x_3, x_4)	0.5200	—
(x_3, x_5)	0.4000	—
(x_3, x_6)	0.5200	—
(x_4, x_5)	0.4000	—
(x_4, x_6)	0.4400	—
(x_5, x_7)	0.4400	—

I opted not to transform x_2 through x_6 and for the most part bivariate normality looks okay. The exceptions are x_1 and x_2, whose proportion of squared distance values less than 1.3862 is 72% for raw data and 68% for transformed x_1, when we'd expect to see 50%. Also, x_2 and x_3 has a percentage of 73%. It might be worth it to transform x_2.

4.35

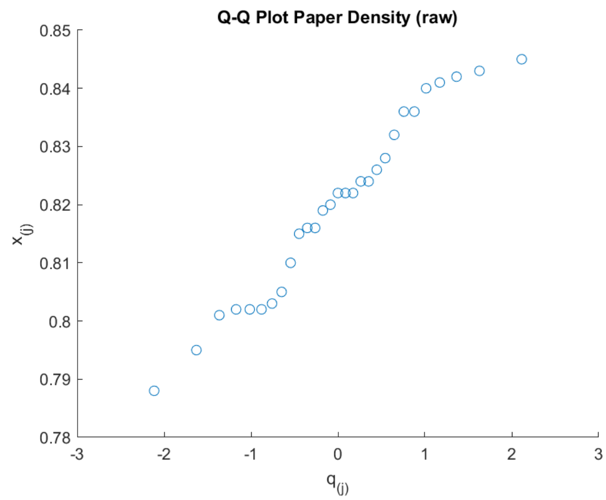
Examine the data on paper-quality measurements in Table 1.2 for marginal and multivariate normality.

Paper is manufactured in continuous sheets several feet wide.Because of the orientation of fibers within the paper, it has a different strength when measured in the direction produced by the machine than when measured across, or at right angles to, the machine direction. Table 1.2 shows the measured values of

First, as mentioned on page 20, observations that are from old paper 16–21, 34, and 38–41 were deleted. Also, observation 25, which is a massive outlier was deleted.

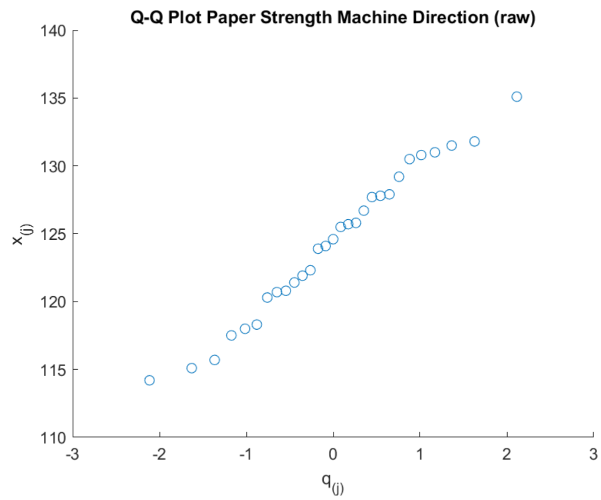
For (x_1), we're using the measuring the paper density in grams/centimeter and have 29 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 29 are, 0.9473, 0.9637, and 0.9703, respectively. The Q-Q correlation coefficient using the raw data was 0.9838. This value is greater than all three critical values, so the data would be considered normally distributed at all three levels. If we left in the 12 observations deleted, this would absolutely not be the case.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was 4.3988, so $x_2' = x_2^{4.3988}$. The Q-Q correlation coefficient on the transformed data was 0.9804, so we do get some improvement. Below is the Q-Q plot for the raw data, which looks pretty good.



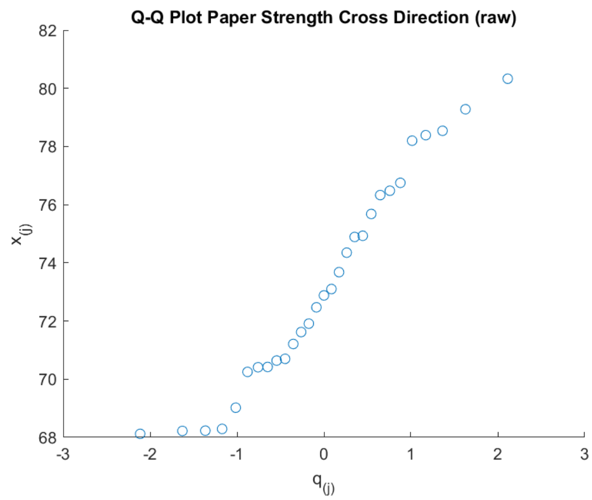
For (x_2), we're using the measuring the paper strength in the machine direction and have 29 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 29 are, 0.9473, 0.9637, and 0.9703, respectively. The Q-Q correlation coefficient using the raw data was 0.9911. This value is greater than all three critical values, so the data would be considered normally distributed at all three levels. If we left in the 12 observations deleted, this would absolutely not be the case.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was 1.7936, so $x_2' = x_2^{1.7936}$. The Q-Q correlation coefficient on the transformed data was 0.9914, so we do get a very small improvement. Below is the Q-Q plot for the raw data, which looks pretty good.



For (x_3), we're using the measuring the paper strength in the machine direction and have 29 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 29 are, 0.9473, 0.9637, and 0.9703, respectively. The Q-Q correlation coefficient using the raw data was 0.9789. This value is greater than all three critical values, so the data would be considered normally distributed at all three levels. If we left in the 12 observations deleted, this would absolutely not be the case.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was -2.4549, so $x_2' = x_2^{-2.4549}$. The Q-Q correlation coefficient on the transformed data was 0.9810, so we do get a very small improvement. Below is the Q-Q plot for the raw data, which looks pretty good.



All of our variables are considered univariate normal, now to check the bivariate, we have $\binom{3}{2} = 3$ pairs of variables to evaluate for

$$\left\{\begin{matrix}(x_1, x_2), & (x_1, x_3), \\ (x_2, x_3) & \end{matrix}\right\}$$

$$\text{Proportion less than } \chi_2^2(0.50) = 1.3863$$

Variables	Raw Data
(x_1, x_2)	0.3793
(x_1, x_3)	0.3103
(x_2, x_3)	0.4483

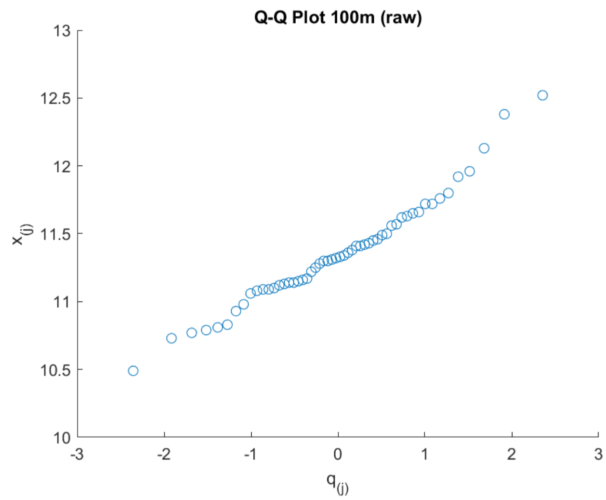
We'd expect to see about 50% of the data with distance values less than 1.3863, but all 3 variables have proportions are are fairly low. We only have 29 observations, so that might have something to do with it.

4.36

Examine the data on women's national track records in Table 1.9 for marginal and multivariate normality.

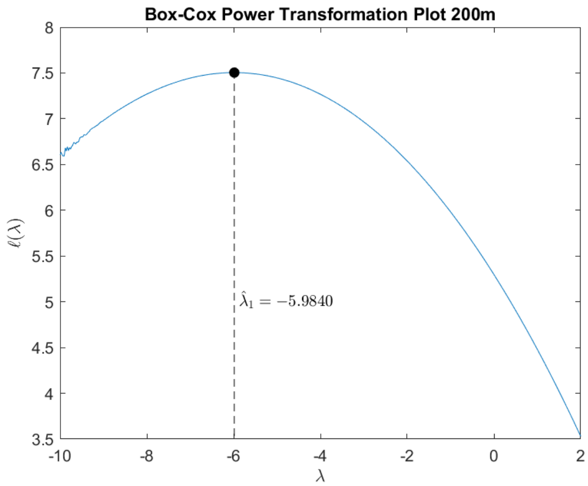
For (x_1), we're looking at the Womens 100m national tract record (in seconds) for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively. The Q-Q correlation coefficient using the raw data is 0.9836, which is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

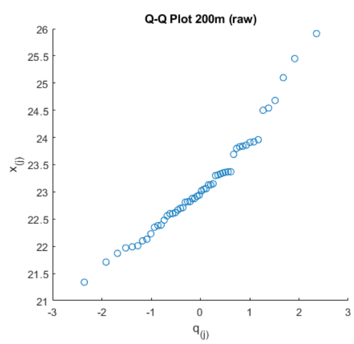
We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at -4.0381, so $x_1' = x_1^{-4.0381}$. The Q-Q correlation coefficient on the transformed data was 0.9932, so we do get a slight improvement. Below is the Q-Q plot for the raw data, which looks alright.

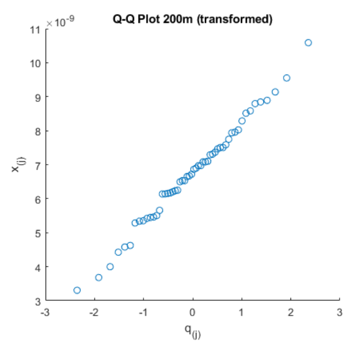


For (x_2), we're looking at the Womens 200m national tract record (in seconds) for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively. The Q-Q correlation coefficient using the raw data was 0.9516, so the data would not be considered normally distributed at the 0.05 and 0.10 levels.

The transformation suggested by Box-Cox power transformation was -5.9840, but I rounded that to -6, so $x_2' = x_2^{-6}$. The Q-Q correlation coefficient on the transformed data was 0.9963, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original plot shows a bit of a curve with a few large values separated out. The Q-Q plot of the transformed data was able to make the data much more linear.



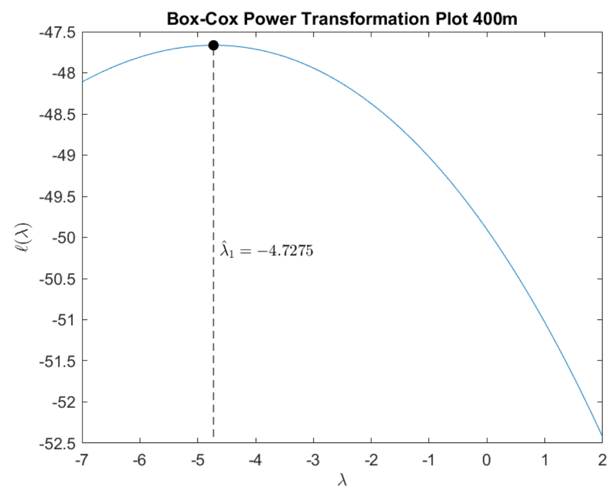


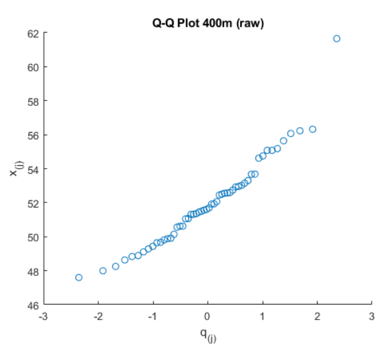


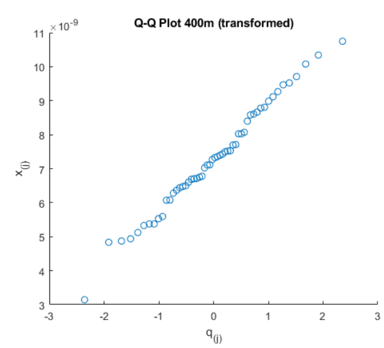
For (x_3), we're looking at the Womens 400m national tract record (in seconds) for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively. The Q-Q correlation coefficient using the raw data was 0.9698, so the data would not be considered normally distributed at the 0.05 and 0.10 levels.

The transformation suggested by Box-Cox power transformation was -4.7275, but I rounded that to -4.75, so $x_3' = x_3^{-4.75}$. The Q-Q correlation coefficient on the transformed data was 0.9949, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original Q-Q plot looks okay with the exception of a single outlier for the country of Cooks Island who has a 200m record of 61.65

seconds. There is no good reason to exclude this observation, so it stays. The Q-Q plot of the transformed data was able to make the data much more linear though.

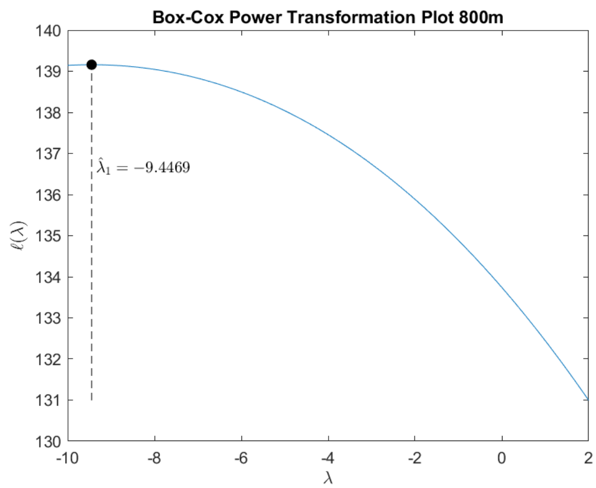


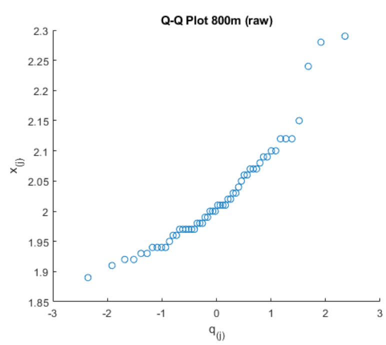


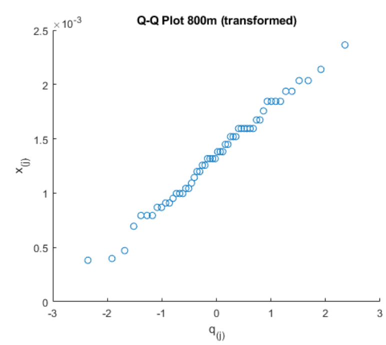


For (x_4), we're looking at the Womens 800m national tract record (in minutes) for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively. The Q-Q correlation coefficient using the raw data was 0.9512, so the data would not be considered normally distributed at any of the three levels.

The transformation suggested by Box-Cox power transformation was -9.4469, but I rounded that to -9.5, so $x_4' = x_4^{-9.5}$. The Q-Q correlation coefficient on the transformed data was 0.9952, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original Q-Q plot has a few large time values, where one is considered an outlier, that are reduced after the transformation. The Q-Q plot of the transformed data looks much better.

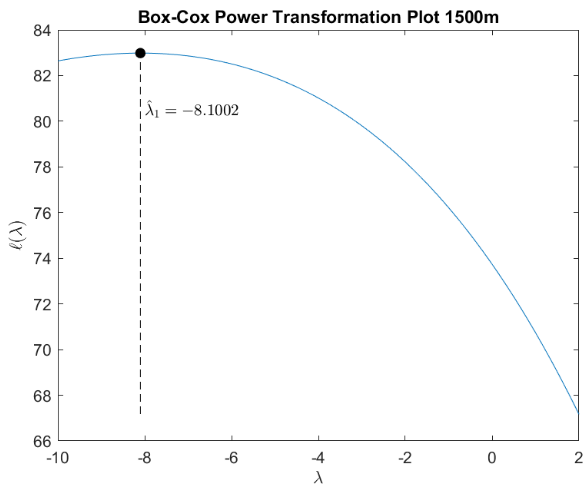


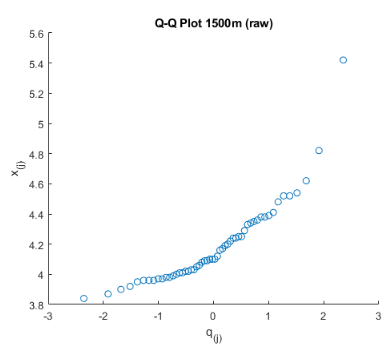


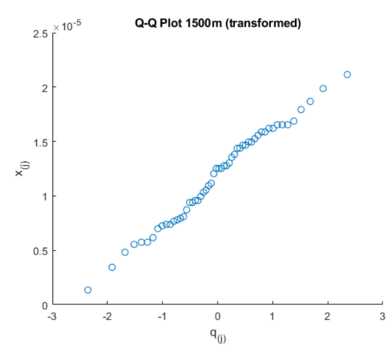


For (x_5), we're looking at the Womens 1500m national tract record (in minutes) for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively. The Q-Q correlation coefficient using the raw data was 0.91, so the data would not be considered normally distributed at any of the three levels.

The transformation suggested by Box-Cox power transformation was -8.1002, but I rounded that to -8, so $x_5' = x_5^{-8}$. The Q-Q correlation coefficient on the transformed data was 0.9921, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original Q-Q plot has an outlier for the country of Samoa (5.42 min), that's reduced after the transformation. The Q-Q plot of the transformed data looks much better.

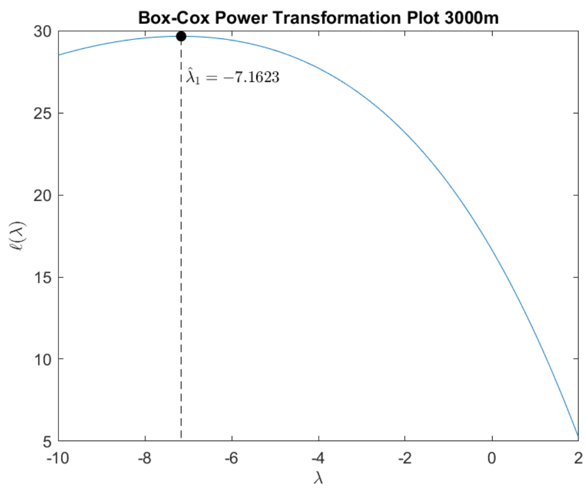


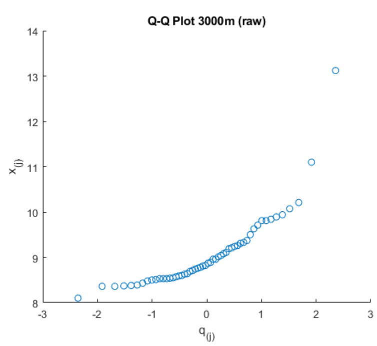


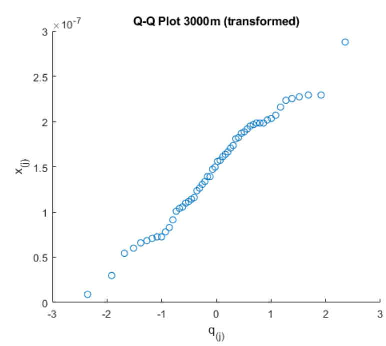


For (x_6), we're looking at the Womens 3000m national tract record (in minutes) for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively. The Q-Q correlation coefficient using the raw data was 0.8677, so the data would not be considered normally distributed at any of the three levels.

The transformation suggested by Box-Cox power transformation was -7.1623, but I rounded that to -7.2, so $x_6' = x_6^{-7.2}$. The Q-Q correlation coefficient on the transformed data was 0.9897, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original Q-Q plot has an outlier for the country of Samoa (13.12 min), that's reduced after the transformation. The Q-Q plot of the transformed data looks somewhat better.

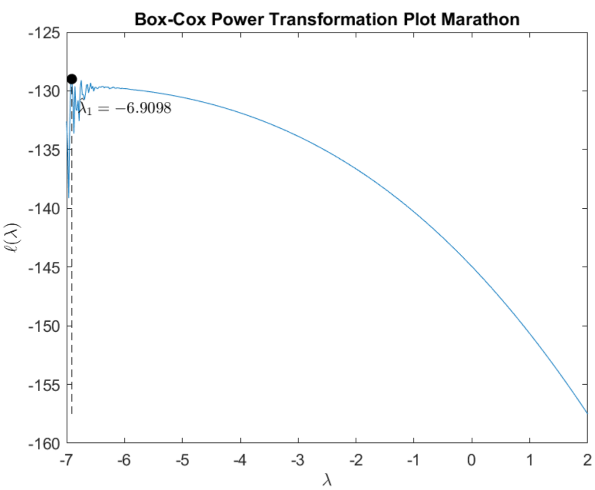


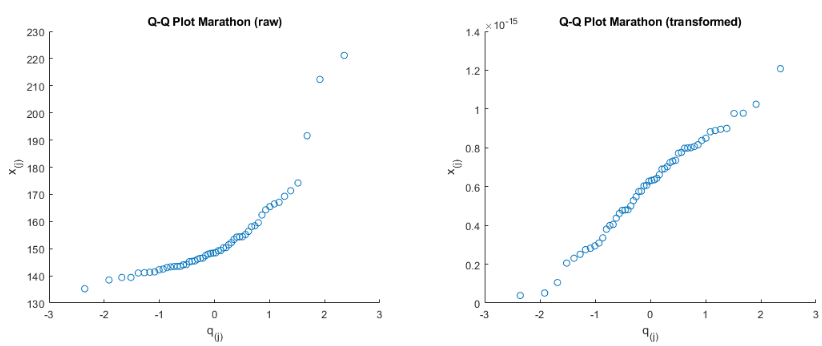




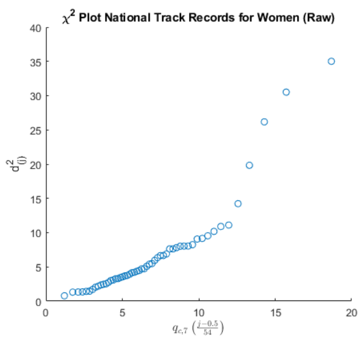
For (x_7), we're looking at the Womens marathon national tract record (in minutes) for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively. The Q-Q correlation coefficient using the raw data was 0.8606, so the data would not be considered normally distributed at any of the three levels.

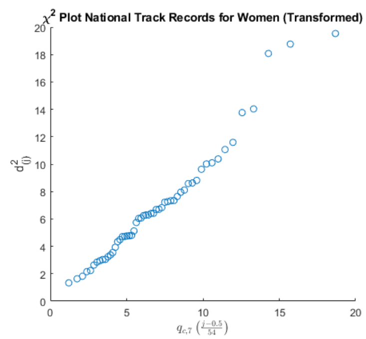
The transformation suggested by Box-Cox power transformation was -6.9098, but I rounded that to -7, so $x_7' = x_7^{-7}$. The Q-Q correlation coefficient on the transformed data was 0.9935, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original Q-Q plot has two outliers. One for Papua New Guinea (221.14 min), and the other for Cook Islands (212.33 min). Both are reduced after the transformation. The Q-Q plot of the transformed data looks somewhat better.





Computing the statistical distance based on our 7 covariates and creating a Chi-Squared plot (below). The raw data on the left shows 5 observations with statistical distances larger than most of the data. The transformed data on the right does show some improvement, but there are 3 observation with larger statistical distances than the others They are, North Korea (19.5334), Czech Republic (18.7619), and Mexico (18.0769). Without them the overall data would appear much more normal, but assuming these are not measurement errors, these observations will have to stay.



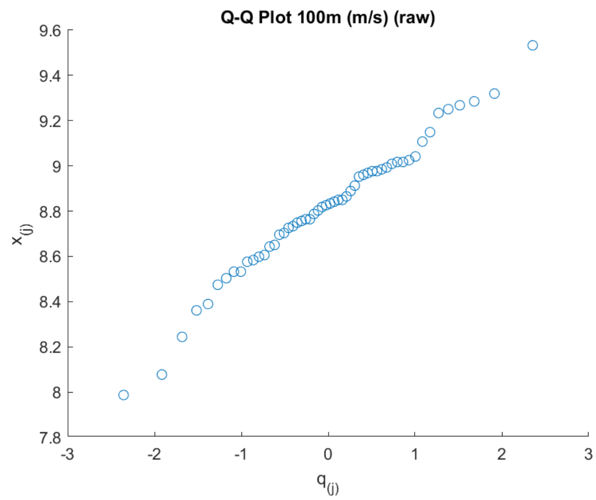


4.37

Refer to Exercise 1.18. Convert the women's track records in Table 1.9 to speeds measured in meters per second. Examine the data on speeds for marginal and multivariate normality.

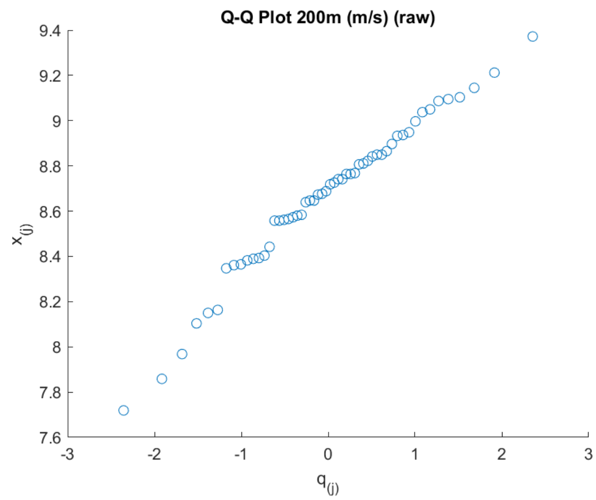
For (x_1), we're looking at the Womens 100m national tract record. This time in meters/seconds for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively and the same as in exercise 4.36. The Q-Q correlation coefficient using the raw data is 0.9898, which is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at 4.0381, so $x_1' = x_1^{4.0381}$. The Q-Q correlation coefficient on the transformed data was 0.9932, so we do get a slight improvement. Below is the Q-Q plot for the raw data, which looks alright.



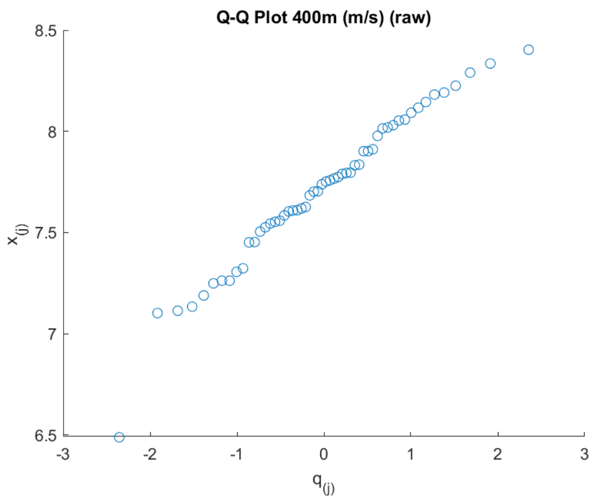
For (x_2), we're looking at the Womens 200m national tract record. This time in meters/seconds for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively and the same as in exercise 4.36. The Q-Q correlation coefficient using the raw data is 0.9859, which is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels. This is a different (better) conclusion than in exercise 4.36.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at 5.9842 and rounded to 6, so $x_2' = x_2^6$. The Q-Q correlation coefficient on the transformed data was 0.9963, so we do get a slight improvement. Below is the Q-Q plot for the raw data, which looks alright.



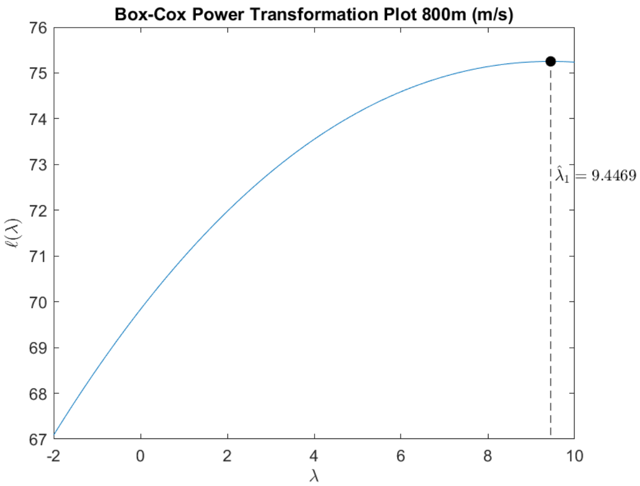
For (x_3), we're looking at the Womens 400m national tract record. This time in meters/seconds for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively and the same as in exercise 4.36. The Q-Q correlation coefficient using the raw data is 0.9846, which is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels. This is a different (better) conclusion than in exercise 4.36. There is still the one outlier for the Cook Islands, but it's in-line with the other points.

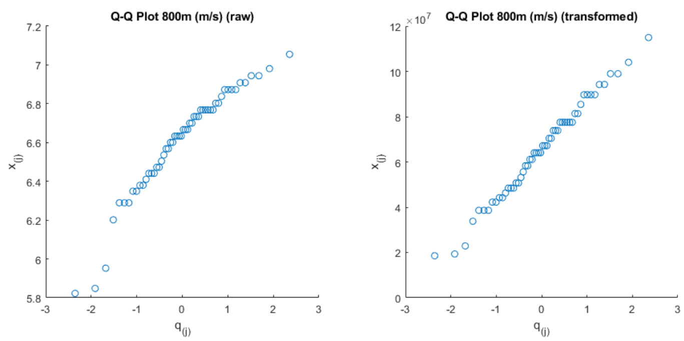
We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at 4.7275 and rounded to 4.75, so $x_3' = x_3^{4.75}$. The Q-Q correlation coefficient on the transformed data was 0.9949, so we do get a slight improvement. Below is the Q-Q plot for the raw data, which looks alright.



For (x_4), we're looking at the Womens 800m national tract record. This time in meters/seconds for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively and the same as in exercise 4.36. The Q-Q correlation coefficient using the raw data was 0.9665, so the data would not be considered normally distributed at any of the three levels.

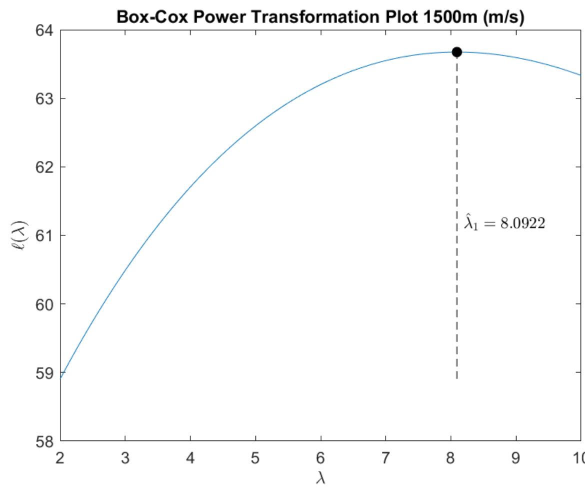
The transformation suggested by Box-Cox power transformation was 9.4469, but I rounded that to 9.5, so $x_4' = x_4^{9.5}$. The Q-Q correlation coefficient on the transformed data was 0.9952, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original Q-Q plot has a few large time values that are reduced after the transformation. The Q-Q plot of the transformed data looks much better.

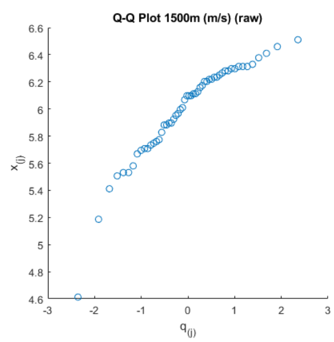


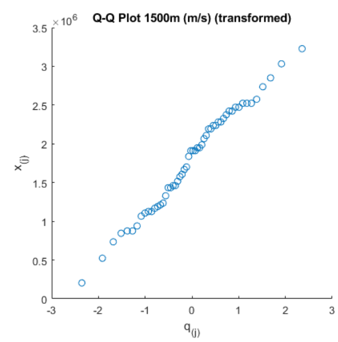


For (x_5), we're looking at the Womens 1500m national tract record. This time in meters/seconds for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively and the same as in exercise 4.36. The Q-Q correlation coefficient using the raw data was 0.9474, so the data would not be considered normally distributed at any of the three levels.

The transformation suggested by Box-Cox power transformation was 8.0922, but I rounded that to 8, so $x_5' = x_5^8$. The Q-Q correlation coefficient on the transformed data was 0.9921, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original Q-Q plot has an outlier for the country of Samoa (5.42 min), that's reduced after the transformation. The Q-Q plot of the transformed data looks much better.

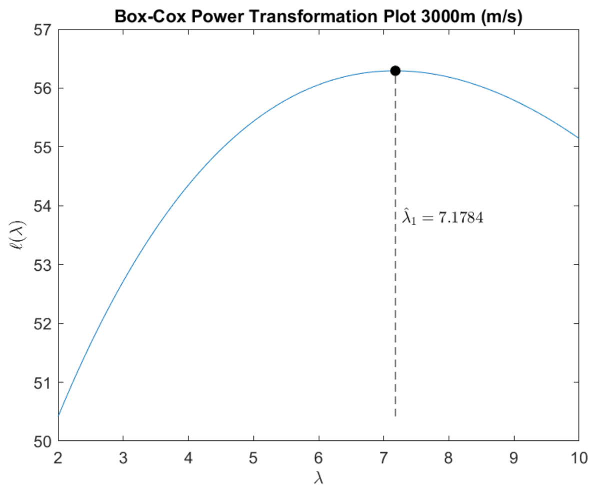


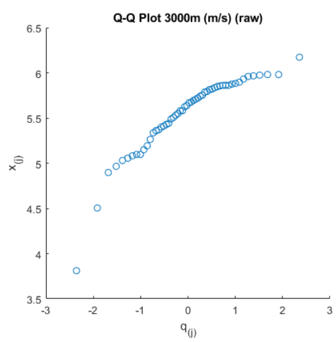


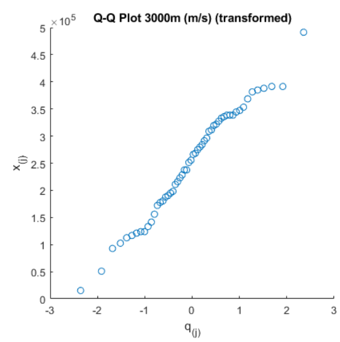


For (x_6), we're looking at the Womens 3000m national tract record. This time in meters/seconds for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively and the same as in exercise 4.36. The Q-Q correlation coefficient using the raw data was 0.8677, so the data would not be considered normally distributed at any of the three levels.

The transformation suggested by Box-Cox power transformation was 7.1784, but I rounded that to 7.2, so $x_6' = x_6^{7.2}$. The Q-Q correlation coefficient on the transformed data was 0.9897, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original Q-Q plot has an outlier for the country of Samoa (13.12 min), that's reduced after the transformation. The Q-Q plot of the transformed data looks somewhat better.

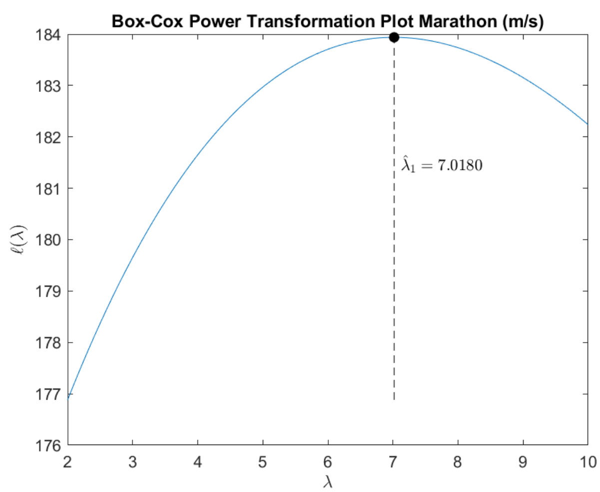


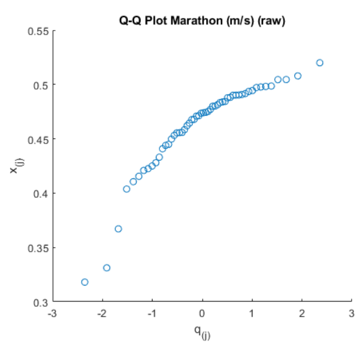


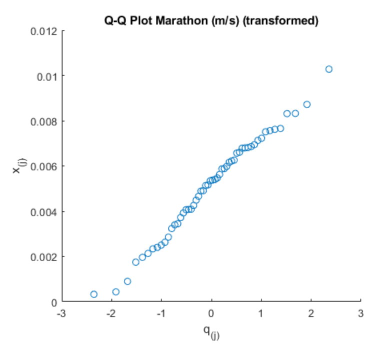


For (x_7), we're looking at the Womens marathon national tract record. This time in meters/seconds for 54 countries. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 54 are, 0.9691, 0.9784, and 0.9822, respectively and the same as in exercise 4.36. The Q-Q correlation coefficient using the raw data was 0.9221, so the data would not be considered normally distributed at any of the three levels.

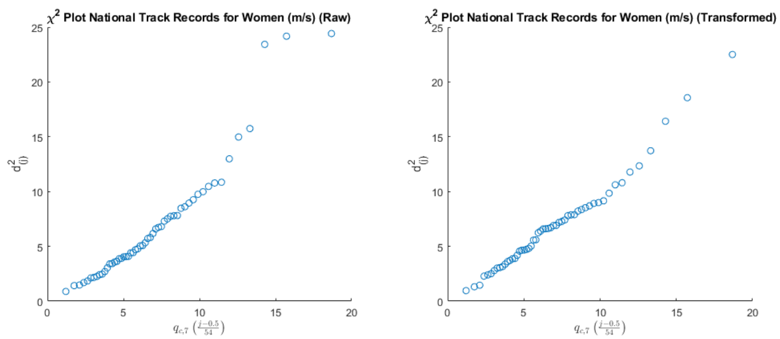
The transformation suggested by Box-Cox power transformation was 7.0180, but I rounded that to 7, so $x_7' = x_7^7$. The Q-Q correlation coefficient on the transformed data was 0.9935, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original Q-Q plot has two outliers. One for Papua New Guinea (221.14 min), and the other for Cook Islands (212.33 min). Both are reduced after the transformation. The Q-Q plot of the transformed data looks somewhat better.







Computing the statistical distance based on our 7 covariates and creating a Chi-Squared plot (below). The raw data on the left shows 3 observations with statistical distances larger than most of the data. The transformed data on the right does show some improvement, but there are 3 observation with larger statistical distances than the others They are, North Korea (22.5017), Czech Republic (18.5597), and Mexico (16.4095). Without them the overall data would appear much more normal, but assuming these are not measurement errors, these observations will have to stay. Converting the columns so they're all in meters per seconds did help. The optimal transformation for most of the variables was the same but with opposite sign.

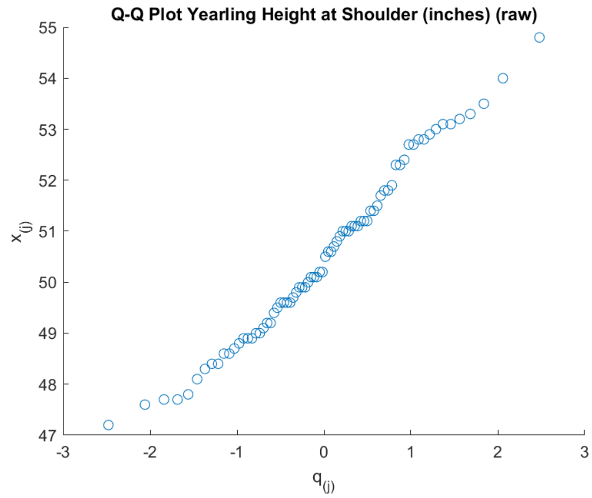


4.38

Examine the data on bulls in Table 1.10 for marginal and multivariate normality. Consider only the variables YrHgt, FtFrBody, PrctFFB, BkFat, SaleHt, and SaleWt

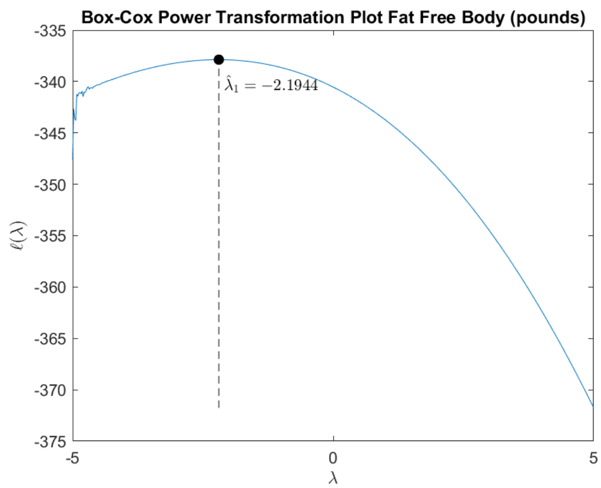
For (x_1), we're looking at the bull yearling height at shoulder (inches) for 76 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9772, 0.9839, and 0.9867, respectively. The Q-Q correlation coefficient using the raw data is 0.9916, which is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

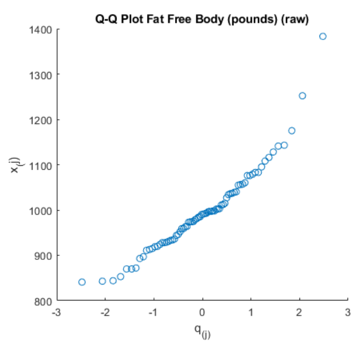
We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at -2.4549, and rounded to -2.5, so $x_1' = x_1^{-2.5}$. The Q-Q correlation coefficient on the transformed data was 0.9938, so we do get a slight improvement. Below is the Q-Q plot for the raw data, which looks good.



For (x_2), we're looking at the bull fat free body (pounds) for 76 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9772, 0.9839, and 0.9867, respectively. The Q-Q correlation coefficient using the raw data was 0.9631, so the data would not be considered normally distributed at the 0.05 and 0.10 levels.

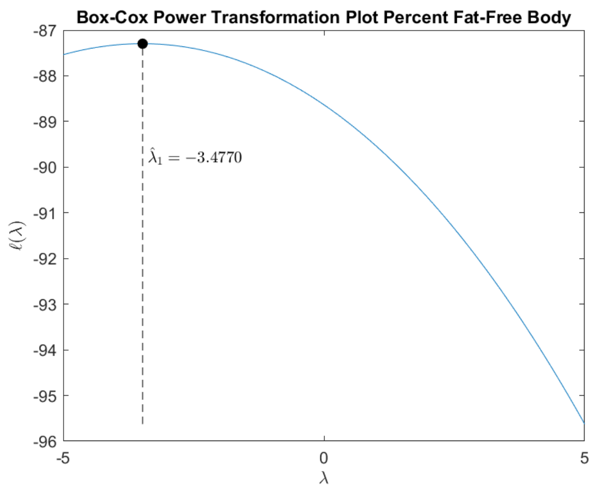
The transformation suggested by Box-Cox power transformation was -2.1944, but I rounded that to -2.2, so $x_2' = x_2^{-2.2}$. The Q-Q correlation coefficient on the transformed data was 0.9937, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original plot shows an outlier separated out. The Q-Q plot of the transformed data was able to make the data much more linear.

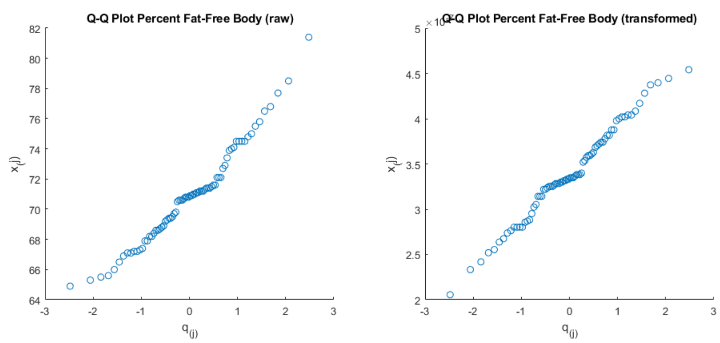




For (x_3), we're looking at the bull percent fat free body for 76 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9772, 0.9839, and 0.9867, respectively. The Q-Q correlation coefficient using the raw data was 0.9847, so the data would be considered normally distributed at the 0.01 and 0.05 levels, but not the 0.10 level.

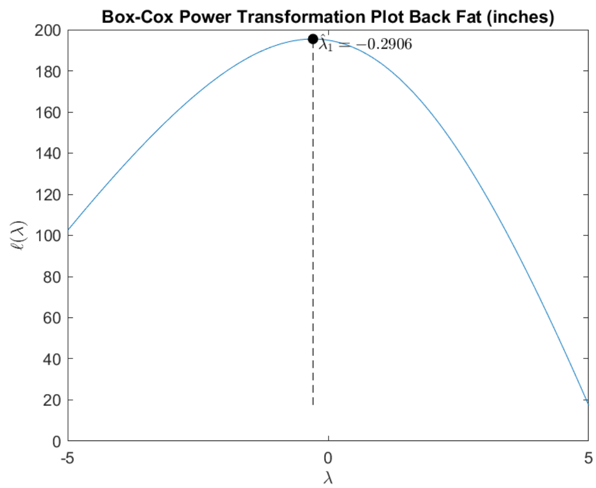
The transformation suggested by Box-Cox power transformation was -3.477, but I rounded that to -3.5, so $x_3' = x_2^{-3.5}$. The Q-Q correlation coefficient on the transformed data was 0.9946, which is larger than all three critical values, so the data is now considered normally distributed at all three levels. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The original plot shows a curvi-linear pattern, but the Q-Q plot of the transformed data was able to make the data much more linear.

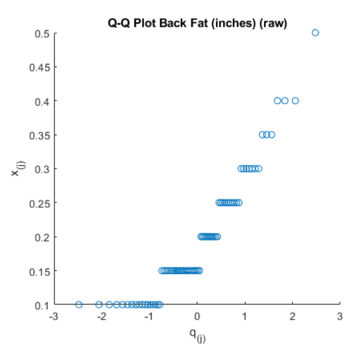


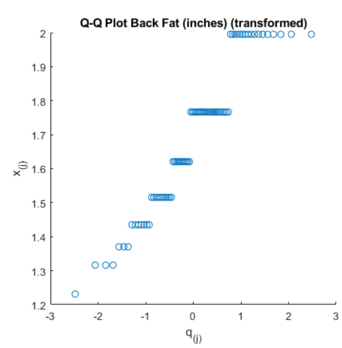


For (x_4), we're looking at the bull back fat (inches) for 76 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9772, 0.9839, and 0.9867, respectively. The Q-Q correlation coefficient using the raw data was 0.9376, less than any of the three critical values. Because of this the data would not be considered normally distributed at any of the three levels.

The transformation suggested by Box-Cox power transformation was -0.2906, but I rounded that to -0.3, so $x'_3 = x_2^{-0.3}$. The Q-Q correlation coefficient on the transformed data was 0.96, which is an improvement, but not good enough to considered the data to be normally distributed. Below are the results of the power transformation and the Q-Q plots of the raw and transformed data. The data has lots of repeated values, so isn't exactly continuous. I don't think there'll be a common transformation to make this data normal. It might be better converted to a factor variable instead.



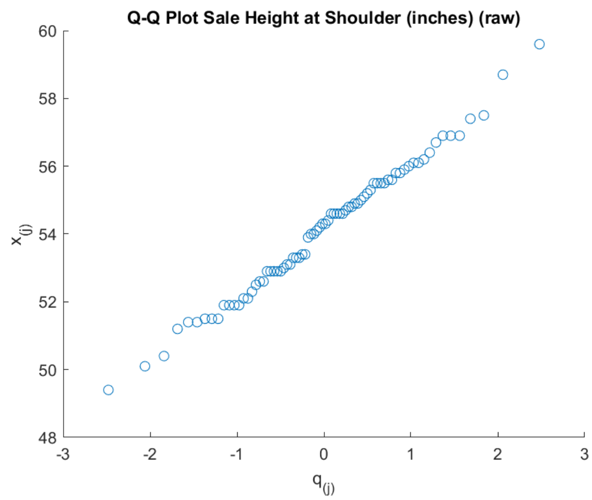




For (x_5), we're looking at the bull sale height at shoulder (inches) for 76 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9772, 0.9839, and 0.9867, respectively. The Q-Q correlation coefficient using the raw data is 0.9956, which is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

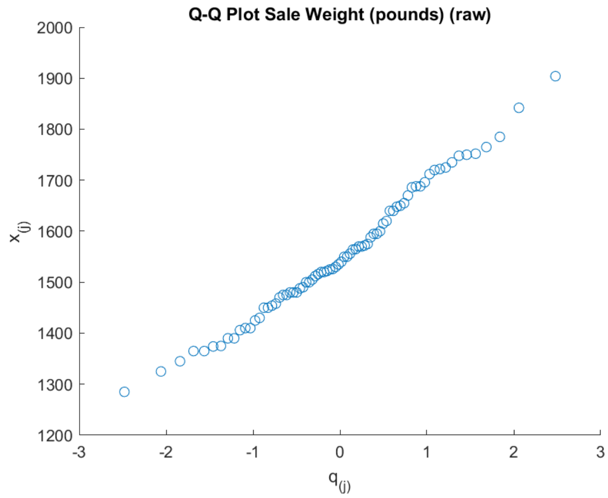
We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at 0.01, and rounded to 0, so $x_5' = \ln\{x_5\}$. The Q-Q correlation coefficient on the transformed data was

0.9959, so we do get a slight improvement. Below is the Q-Q plot for the raw data, which looks really good.

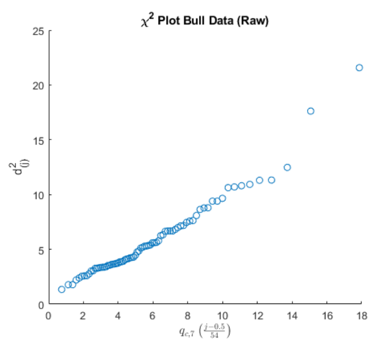


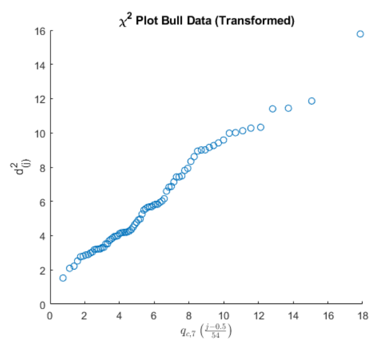
For (x_6), we're looking at the bull sale weight (pounds) for 76 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9772, 0.9839, and 0.9867, respectively. The Q-Q correlation coefficient using the raw data is 0.9956, which is larger than all three of these values, so the data would be considered normally distributed at the 0.01, 0.05, and 0.10 levels.

We don't really need to transform the data, but just to see how much better it can get, the Box-Cox power transformation max was at -0.6112, and rounded to -0.5, so $x_6' = \frac{1}{\sqrt{x_6}}$. The Q-Q correlation coefficient on the transformed data was 0.9968, so we do get a slight improvement. Below is the Q-Q plot for the raw data, which looks really good.



Computing the statistical distance based on our 6 covariates and creating a Chi-Squared plot (below). The raw data on the left shows two observations with statistical distances larger than most of the data. The transformed data on the right does show some improvement, but observation 15 has a larger statistical distance than the others (15.777). Without them the overall data would appear much more normal, but assuming it's not a measurement error, this observations will stay. Even with the outlier in the transformed data, it still looks fairly linear. I would think this data is multivariate normal.





4.39

The data in Table 4.6 (see the psychological profile data: www.prenhall.com/statistics) consist of 130 observations generated by scores on a psychological test administered to Peruvian teenagers (ages 15, 16, and 17). For each of these teenagers the gender (male = 1, female = 2) and socioeconomic status (low = 1, medium = 2) were also recorded The scores were accumulated into five subscale scores labeled *independence* (indep), *support* (supp), *benevolence* (benev), *conformity* (conform), and *leadership* (leader).

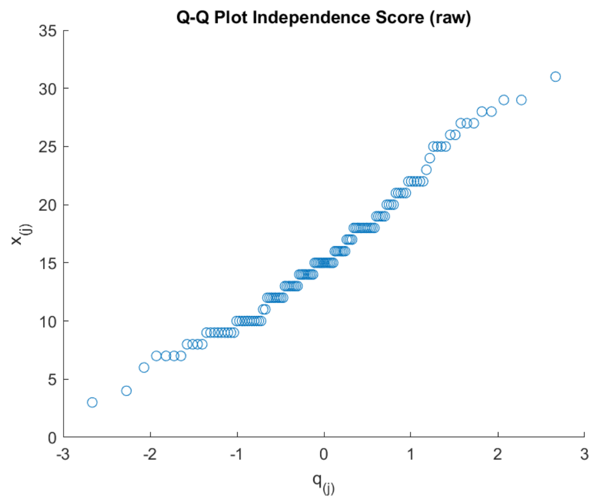
Table 4.6 Psychological Profile Data

Indep	Supp	Benev	Conform	Leader	Gender	Socio
27	13	14	20	11	2	1
12	13	24	25	6	2	1
14	20	15	16	7	2	1
18	20	17	12	6	2	1
9	22	22	21	6	2	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	11	26	17	10	1	1
14	12	14	11	29	1	2
19	11	23	18	13	2	2
27	19	22	7	9	2	2
10	17	22	22	8	2	2

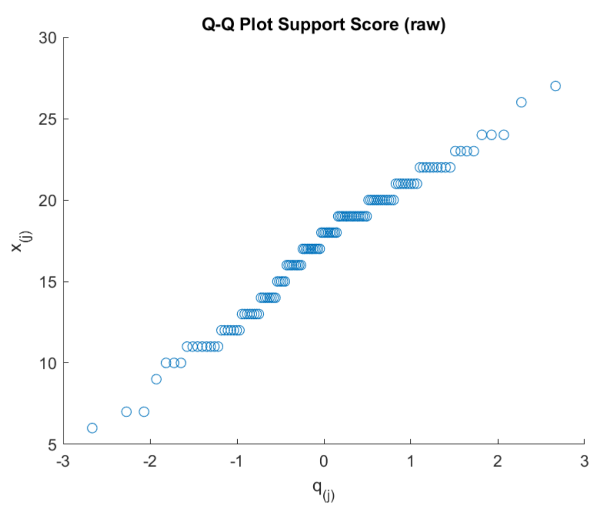
Source: Data courtesy of C. Soto.

(a) Examine each of the variables independence, support, benevolence, conformity and leadership for marginal normality.

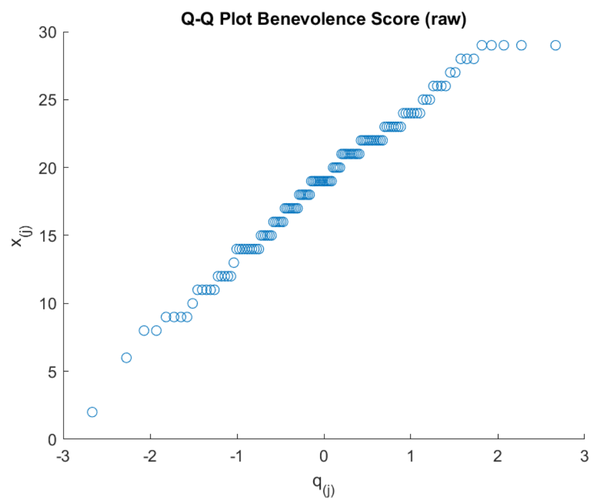
For (x_1), we're looking at the independence score for 130 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9860, 0.9900, and 0.9916, respectively. The Q-Q correlation coefficient using the raw data is 0.9881, which is larger than the 0.01-level critical point, but not the 0.05 and 0.10, and so would be considered normal at the 0.01-level but not the 0.05 and 0.10. The Q-Q plot for the raw data is below. There is some slight curvature, but a transformation might help.



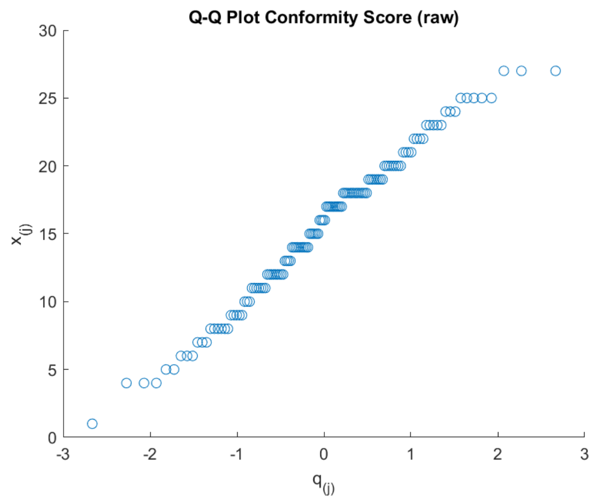
For (x_2), we're looking at the support score for 130 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9860, 0.9900, and 0.9916, respectively. The Q-Q correlation coefficient using the raw data is 0.9893, which is larger than the 0.01-level critical point, but not the 0.05 and 0.10, and so would be considered normal at the 0.01-level but not the 0.05 and 0.10. The Q-Q plot for the raw data is below. There is some slight curvature, but a transformation also might help.



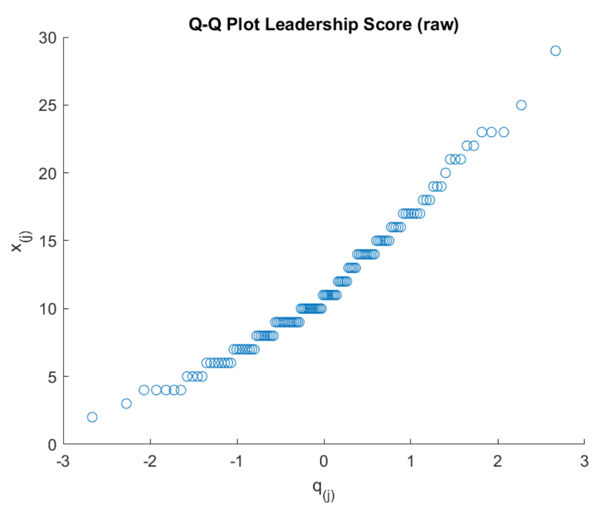
For (x_3), we're looking at the benevolence score for 130 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9860, 0.9900, and 0.9916, respectively. The Q-Q correlation coefficient using the raw data is 0.9925, which is larger than all three critical points, so the benevolence score would be considered normal at the 0.01, 0.05, and 0.10 levels. The Q-Q plot for the raw data is below. There are five observations with the same score value of 29 causing a right tail effect, but overall x_3 looks marginally normal.



For (x_4), we're looking at the conformity score for 130 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9860, 0.9900, and 0.9916, respectively. The Q-Q correlation coefficient using the raw data is 0.9934, which is larger than all three critical points, so the conformity score would be considered normal at the 0.01, 0.05, and 0.10 levels. The Q-Q plot for the raw data is below. There are three observations with the same score value of 27 causing a right tail effect, similar to x_3, but overall x_4 looks marginally normal.

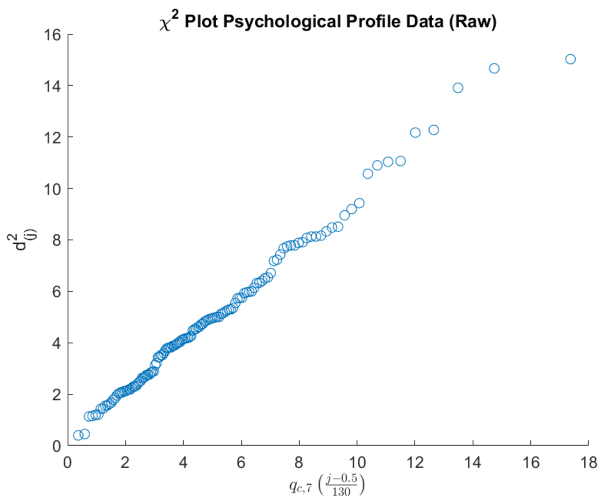


For (x_5), we're looking at the leadership score for 130 valid observations. The simulated 0.01, 0.05, and 0.10 level critical correlation coefficient test values for a sample size of 76 are, 0.9860, 0.9900, and 0.9916, respectively. The Q-Q correlation coefficient using the raw data is 0.9813, which is smaller than all three critical values, and so would not be considered normal at the 0.01, 0.05, or 0.10 levels. The Q-Q plot for the raw data is below. There is curvature in the plot. A transformation would absolutely help.



(b) Using all five variables, check for multivariate normality.

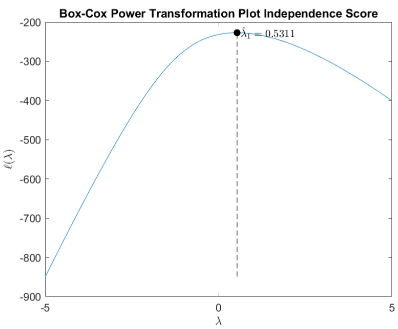
The χ^2 plot for the five covariates is below. It looks fairly linear. However, there are three observations with large d^2 values that stand out as possible outliers.

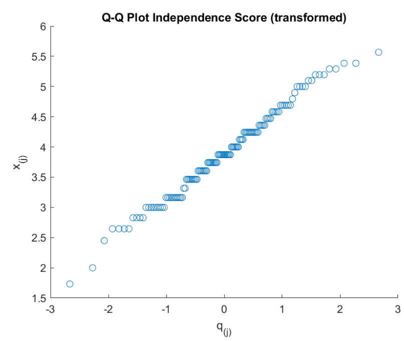


(c) Refer to part (a). For those variables that are nonnormal, determine the transformation that makes them more nearly normal.

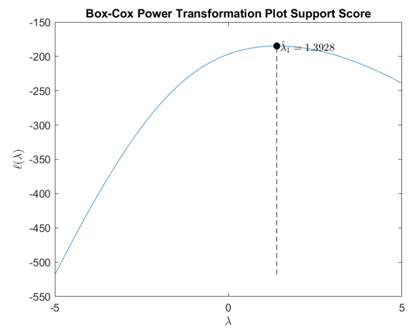
From part (a) we found that independence (x_1), support (x_2) and leadership (x_5) could use a transformation.

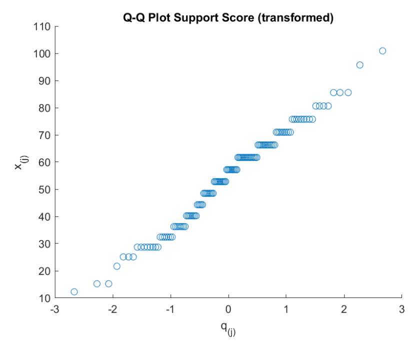
For independence score, x_1, the optimal Box-Cox power transformation value found was 0.5311. I rounded that to 0.5, so $x_1' = \sqrt{x_1}$. This transformation increased the Q-Q correlation coefficient to 0.9949, which is larger than the critical point values of 0.9860, 0.9900, and 0.9916 who represent the 0.01, 0.05, and 0.10 levels, respectively. Because of this the data would be considered normally distributed at all the usual levels. Below are the results of the power transformation and the Q-Q plots of the transformed data. The original plot shows an outlier separated out. The Q-Q plot of the transformed data was able to make the data appear more linear.



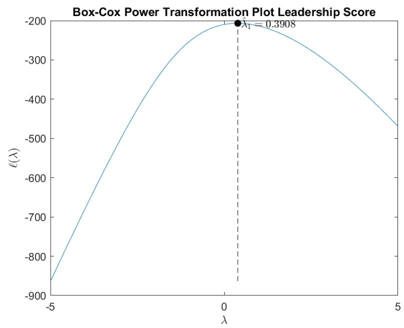


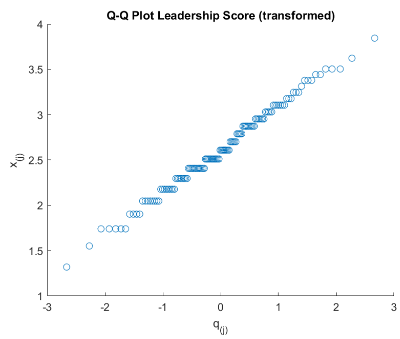
For support score, x_2, the optimal Box-Cox power transformation value found was 1.3928. I rounded that to 1.4, so $x_2' = x_2^{1.4}$. This transformation increased the Q-Q correlation coefficient to 0.9920, which is larger than the critical point values of 0.9860, 0.9900, and 0.9916 who represent the 0.01, 0.05, and 0.10 levels, respectively. Because of this the data would be considered normally distributed at all the usual levels. Below are the results of the power transformation and the Q-Q plots of the transformed data. The original plot shows an outlier separated out. The Q-Q plot of the transformed data was able to make the data appear more linear.





For leadership score, x_5, the optimal Box-Cox power transformation value found was 0.3908. I rounded that to 0.4, so $x_5' = x_5^{0.4}$. This transformation increased the Q-Q correlation coefficient to 0.9965, which is larger than the critical point values of 0.9860, 0.9900, and 0.9916 who represent the 0.01, 0.05, and 0.10 levels, respectively. Because of this the data would be considered normally distributed at all the usual levels. Below are the results of the power transformation and the Q-Q plots of the transformed data. The original plot shows an outlier separated out. The Q-Q plot of the transformed data was able to make the data appear more linear.



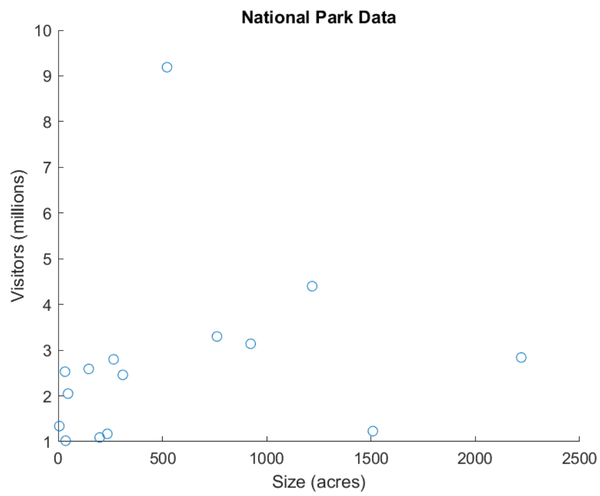


4.40

Consider the data on national parks in Exercise 1.27.

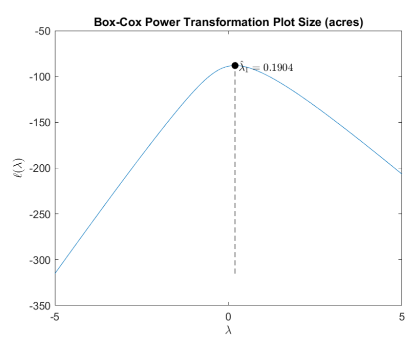
(a) Comment on any possible outliers in a scatter plot of the original variables.

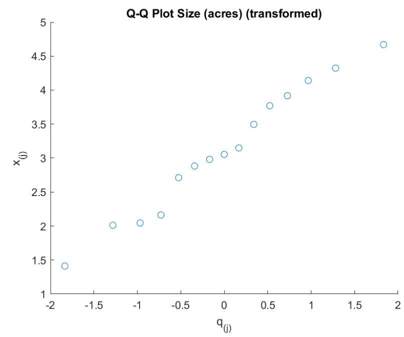
The scatter plot of the data is below. We only have 15 observations, so the sample size is small. In the x-drection, park size in acres, it looks like there's one observation, corresponding to Yellowstone, which has a very large size of 2219.8 acres. In the y-direction, number of visitors in millions, there also appears to be a park with a very large numer of visitors, corresponding to Great Smoky (9.19 million).



(b) Determine the power transformation $\hat{\lambda}_1$ the makes the x_1 values approximately normal. Construct a Q-Q plot of the transformed observations.

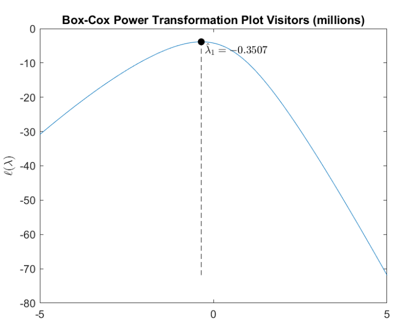
For park size in acres, x_1, the optimal Box-Cox power transformation value found was $\hat{\lambda}_1 = 0.1904$. I rounded that to 0.2, so $x_1' = x_1^{\hat{\lambda}_1} = x_1^{0.20}$. This transformation increased the Q-Q correlation coefficient to 0.9909, which is larger than the critical point values of 0.9110, 0.9392, and 0.9506 who represent the 0.01, 0.05, and 0.10 levels, respectively, for a sample size of 15. Because of this the data would be considered normally distributed at all the usual levels. Below are the results of the power transformation and the Q-Q plots of the transformed data. For only having 15 observations, the Q-Q plot of the transformed data looks fairly linear.

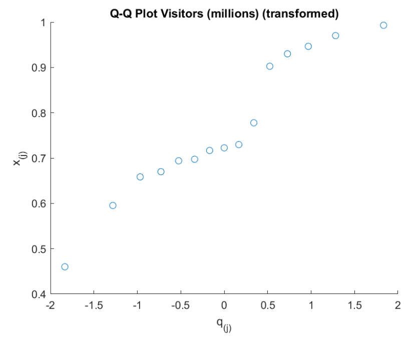




(c) Determine the power transformation $\hat{\lambda}_2$ the makes the x_2 values approximately normal. Construct a Q-Q plot of the transformed observations.

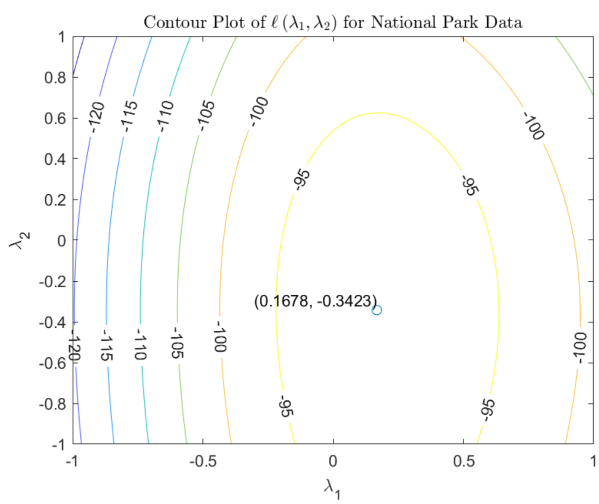
For number of visitors in millions, x_2, the optimal Box-Cox power transformation value found was $\hat{\lambda}_2 = -0.3507$. I rounded that to -0.35, so $x_1' = x_2^{\hat{\lambda}_2} = x_2^{-0.35}$. This transformation increased the Q-Q correlation coefficient to 0.9675, which is larger than the critical point values of 0.9110, 0.9392, and 0.9506 who represent the 0.01, 0.05, and 0.10 levels, respectively, for a sample size of 15. Because of this the data would be considered normally distributed at all the usual levels. Below are the results of the power transformation and the Q-Q plots of the transformed data. The transformed Q-Q plot doesn't look strongly linear, but our sample size is only 15, so the bar isn't very high.

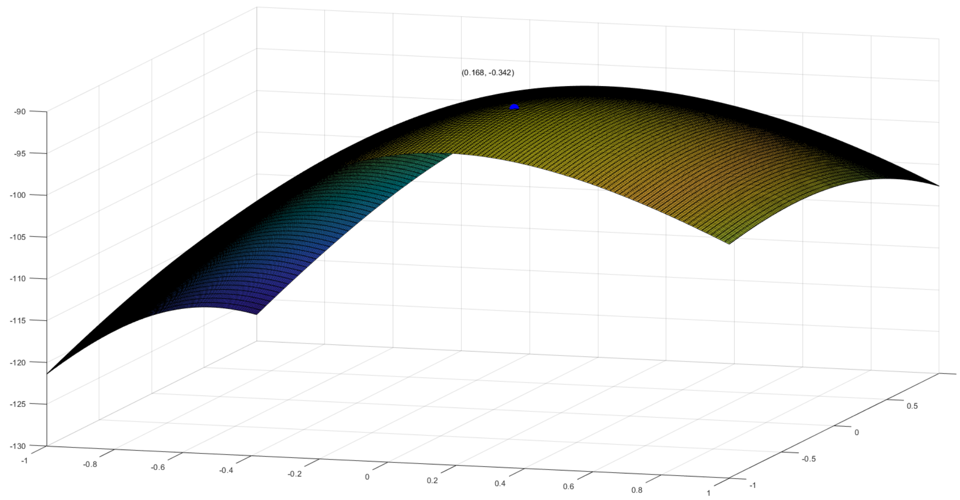




(d) Determine the power transformation for approximate bivariate normality using (4–40).

The optimum found was at $\left(\hat{\lambda}_1, \hat{\lambda}_2\right) = (0.1678, -0.3423)$. These values are close to those found in part (b) and (c), where the result of optimizing the univariate transform using Box-Cox was $\left(\hat{\lambda}_1, \hat{\lambda}_2\right) = (0.1904, -0.3507)$. The contour and surface plot for $\ell\left(\lambda_1, \lambda_2\right)$ is below.



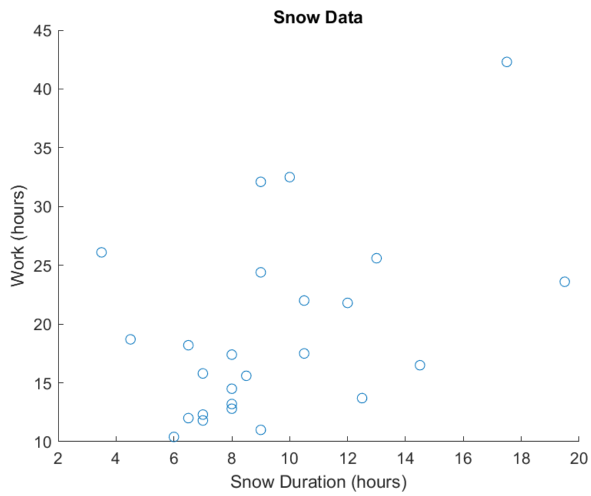


4.41

Consider the data on snow removal in Exercise 3.20.

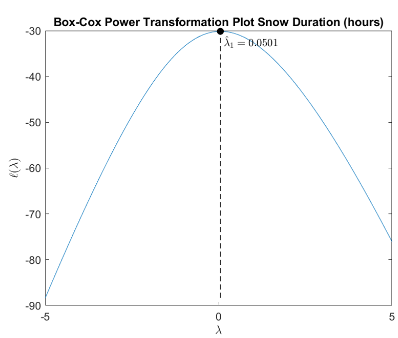
(a) Comment on any possible outliers in a scatter plot of the original variables.

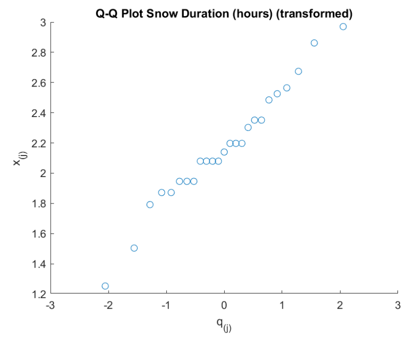
The scatter plot of the data is below for 25 observations. There's a couple observations with large snow durations. Observation 5 and observation 21 with snow direations of 19.5 hours and 17.5 hours, respectively. Observation 21 also has a high number of work hours to clear the snow (42.3 hours).



(b) Determine the power transformation $\hat{\lambda}_1$ the makes the x_1 values approximately normal. Construct a Q-Q plot of the transformed observations.

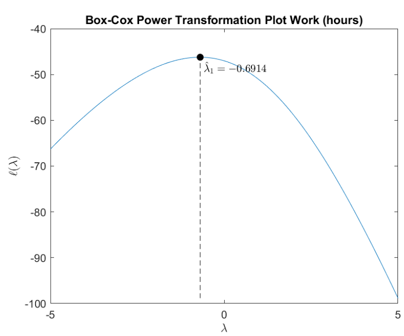
For the snow duration in hours, x_1, the optimal Box-Cox power transformation value found was $\hat{\lambda}_1 = 0.0501$. I rounded that to 0, so $x_1' = x_1^{\hat{\lambda}_1} = \ln\{x_1\}$. This transformation increased the Q-Q correlation coefficient to 0.9872, which is larger than the critical point values of 0.9406, 0.9591, and 0.9665 who represent the 0.01, 0.05, and 0.10 levels, respectively, for a sample size of 25. Because of this the data would be considered normally distributed at all the usual levels. Below are the results of the power transformation and the Q-Q plots of the transformed data. The Q-Q plot of the transformed data looks pretty good.

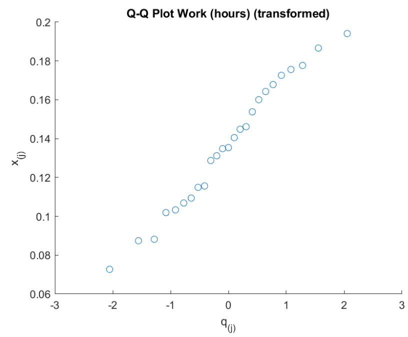




(c) Determine the power transformation $\hat{\lambda}_2$ the makes the x_2 values approximately normal. Construct a Q-Q plot of the transformed observations.

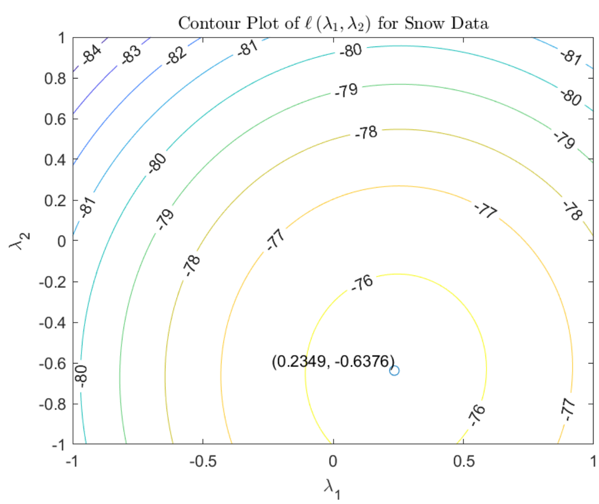
For amount of time it took to clear the snow in hours, x_2, the optimal Box-Cox power transformation value found was $\hat{\lambda}_2 = -0.6914$. I rounded that to -0.70, so $x_1' = x_2^{\hat{\lambda}_2} = x_2^{-0.70}$. This transformation increased the Q-Q correlation coefficient to 0.9909, which is larger than the critical point values of 0.9406, 0.9591, and 0.9665 who represent the 0.01, 0.05, and 0.10 levels, respectively, for a sample size of 25. Because of this the data would be considered normally distributed at all the usual levels. Below are the results of the power transformation and the Q-Q plots of the transformed data. The transformed Q-Q plot as the liear appearance we'd hope for.

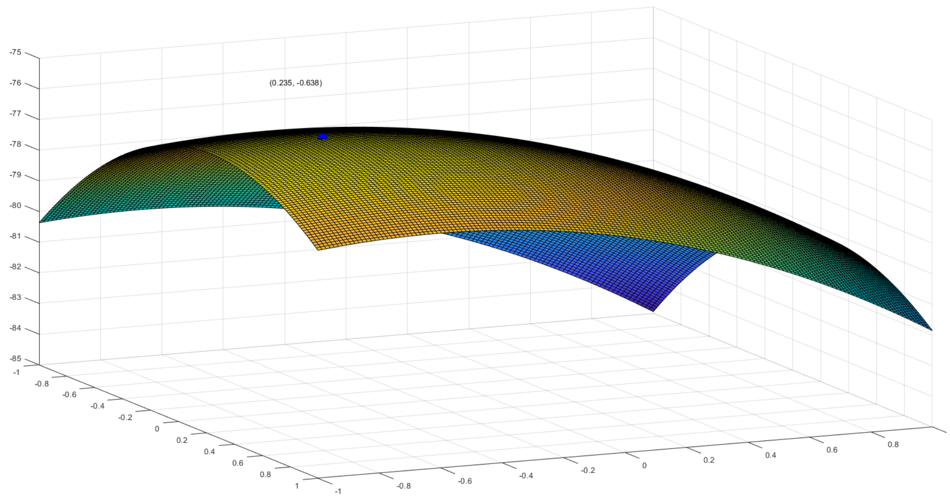




(d) Determin,e the power transformation for approximate bivariate normality using (4–40).

The optimum found was at $\left(\hat{\lambda}_1, \hat{\lambda}_2\right) = (0.2349, -0.6376)$. These values sort of close to those found in part (b) and (c), at least for $\hat{\lambda}_2$ anyway, where the result of optimizing the univariate transform using Box-Cox was $\left(\hat{\lambda}_1, \hat{\lambda}_2\right) = (0.0501, -0.6914)$. The contour and surface plot for $\ell(\lambda_1, \lambda_2)$ is below.




(0.235, -0.638)
-75
-76
-77
-78
-79
-80
-81
-82
-83
-84
-85
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1