Appendix B

Answers and Hints to Problems

CHAPTER 2

2.1 **(a)** $\mathbf{A}+\mathbf{B}=\begin{pmatrix} 7 & 0 & 7 \\ 13 & 14 & 3 \end{pmatrix}$, $\quad \mathbf{A}-\mathbf{B}=\begin{pmatrix} 1 & 4 & -1 \\ 1 & -4 & 13 \end{pmatrix}$

(b) $\mathbf{A}'\mathbf{A}=\begin{pmatrix} 65 & 43 & 68 \\ 43 & 29 & 46 \\ 68 & 46 & 73 \end{pmatrix}$, $\quad \mathbf{A}\mathbf{A}'=\begin{pmatrix} 29 & 62 \\ 62 & 138 \end{pmatrix}$

2.2 **(a)** $(\mathbf{A}+\mathbf{B})'=\begin{pmatrix} 7 & 13 \\ 0 & 14 \\ 7 & 3 \end{pmatrix}$, $\quad \mathbf{A}'+\mathbf{B}'=\begin{pmatrix} 7 & 13 \\ 0 & 14 \\ 7 & 3 \end{pmatrix}$

(b) $\mathbf{A}'=\begin{pmatrix} 4 & 7 \\ 2 & 5 \\ 3 & 8 \end{pmatrix}$, $\quad (\mathbf{A}')'=\begin{pmatrix} 4 & 2 & 3 \\ 7 & 5 & 8 \end{pmatrix}=\mathbf{A}$

*

2.3 **(a)** $\mathbf{AB} = \begin{pmatrix} 5 & 15 \\ 3 & -5 \end{pmatrix}$, $\mathbf{BA} = \begin{pmatrix} 2 & 6 \\ 11 & -2 \end{pmatrix}$

(b) $|\mathbf{AB}| = -70$, $|\mathbf{A}| = -7$, $|\mathbf{B}| = 10$

2.4 **(a)** $\mathbf{A} + \mathbf{B} = \begin{pmatrix} 3 & 3 \\ 3 & 4 \end{pmatrix}$, $\text{tr}(\mathbf{A} + \mathbf{B}) = 7$

(b) $\text{tr}(\mathbf{A}) = 0$, $\text{tr}(\mathbf{B}) = 7$

2.5 **(a)** $\mathbf{AB} = \begin{pmatrix} 4 & 1 \\ 3 & -3 \end{pmatrix}$, $\mathbf{BA} = \begin{pmatrix} -1 & 8 & 7 \\ 2 & 4 & 6 \\ 1 & -3 & -2 \end{pmatrix}$

(b) $\text{tr}(\mathbf{AB}) = 1$, $\text{tr}(\mathbf{BA}) = 1$

2.6 **(b)** $\mathbf{x} = (1 \quad 1 \quad -1)'$

2.7 **(a)** $\mathbf{Bx} = (13, 6, 9)'$ **(b)** $\mathbf{y}'\mathbf{B} = (25, -1, 17)$ **(c)** $\mathbf{x}'\mathbf{Ax} = 16$

(d) $\mathbf{x}'\mathbf{Ay} = 43$ **(e)** $\mathbf{x}'\mathbf{x} = 6$ **(f)** $\mathbf{x}'\mathbf{y} = 3$

(g) $\mathbf{xx}' = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{pmatrix}$ **(h)** $\mathbf{xy}' = \begin{pmatrix} 3 & 2 & 1 \\ -3 & -2 & -1 \\ 6 & 4 & 2 \end{pmatrix}$

(i) $\mathbf{B}'\mathbf{B} = \begin{pmatrix} 62 & 7 & 22 \\ 7 & 14 & 7 \\ 22 & 7 & 41 \end{pmatrix}$

2.8 **(a)** $\mathbf{x} + \mathbf{y} = (4, 1, 3)'$, $\mathbf{x} - \mathbf{y} = (-2, -3, 1)'$

(b) $(\mathbf{x} - \mathbf{y})'\mathbf{A}(\mathbf{x} - \mathbf{y}) = -31$

2.9 $\mathbf{Bx} = \mathbf{b}_1 x_1 + \mathbf{b}_2 x_2 + \mathbf{b}_3 x_3 = (1)\begin{pmatrix} 3 \\ 7 \\ 2 \end{pmatrix} + (-1)\begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix} + (2)\begin{pmatrix} 4 \\ 0 \\ 5 \end{pmatrix} =$

$\begin{pmatrix} 13 \\ 6 \\ 9 \end{pmatrix}$

2.10 **(a)** $(\mathbf{AB})' = \begin{pmatrix} 7 & 16 \\ 8 & 4 \\ 7 & 11 \end{pmatrix}$, $\mathbf{B}'\mathbf{A}' = \begin{pmatrix} 7 & 16 \\ 8 & 4 \\ 7 & 11 \end{pmatrix}$ **(c)** $|\mathbf{A}| = 5$

2.11 **(a)** $\mathbf{a}'\mathbf{b} = 5$, $(\mathbf{a}'\mathbf{b})^2 = 25$ **(b)** $\mathbf{bb}' = \begin{pmatrix} 4 & 2 & 6 \\ 2 & 1 & 3 \\ 6 & 3 & 9 \end{pmatrix}$, $\mathbf{a}'(\mathbf{bb}')\mathbf{a} = 25$

2.12 $\mathbf{DA} = \begin{pmatrix} a & 2a & 3a \\ 4b & 5b & 6b \\ 7c & 8c & 9c \end{pmatrix}$, $\mathbf{AD} = \begin{pmatrix} a & 2b & 3c \\ 4a & 5b & 6c \\ 7a & 8b & 9c \end{pmatrix}$,

$$\mathbf{DAD} = \begin{pmatrix} a^2 & 2ab & 3ac \\ 4ab & 5b^2 & 6bc \\ 7ac & 8bc & 9c^2 \end{pmatrix}$$

2.13 $\mathbf{AB} = \left(\begin{array}{ccc|c} 8 & 9 & 5 & 6 \\ 7 & 5 & 5 & 4 \\ \hline 3 & 4 & 2 & 2 \end{array}\right)$

2.14 $\mathbf{AB} = \begin{pmatrix} 3 & 5 \\ 1 & 4 \end{pmatrix}, \quad \mathbf{CB} = \begin{pmatrix} 3 & 5 \\ 1 & 4 \end{pmatrix}$

2.15 **(a)** $\text{tr}(\mathbf{A}) = 4, \quad \text{tr}(\mathbf{B}) = 5$

(b) $\mathbf{A} + \mathbf{B} = \begin{pmatrix} 6 & 4 & 5 \\ 2 & -2 & 1 \\ 4 & 9 & 5 \end{pmatrix}, \quad \text{tr}(\mathbf{A} + \mathbf{B}) = 9$

(c) $|\mathbf{A}| = 23, \quad |\mathbf{B}| = 2$

(d) $\mathbf{AB} = \begin{pmatrix} 9 & 12 & 17 \\ 3 & -1 & 5 \\ 5 & 11 & 9 \end{pmatrix}, \quad \det(\mathbf{AB}) = 46$

2.16 **(a)** $|\mathbf{A}| = 36$ **(b)** $\mathbf{T} = \begin{pmatrix} 1.7321 & 2.3094 & 1.7321 \\ 0 & 1.6330 & 1.2247 \\ 0 & 0 & 2.1213 \end{pmatrix}$

2.17 **(a)** $\det(\mathbf{A}) = 1$ **(b)** $\mathbf{T} = \begin{pmatrix} 1.7321 & -2.8868 & -.5774 \\ 0 & 2.1602 & -.7715 \\ 0 & 0 & .2673 \end{pmatrix}$

2.18 **(a)** $\mathbf{C} = \begin{pmatrix} .4082 & -.5774 & .7071 \\ .8165 & .5774 & .0000 \\ .4082 & -.5774 & -.7071 \end{pmatrix}$

2.19 **(a)** Eigenvalues: $2, 1, -1$

Eigenvectors: $\begin{pmatrix} .3015 \\ .9045 \\ .3015 \end{pmatrix}, \begin{pmatrix} .7999 \\ .5368 \\ .2684 \end{pmatrix}, \begin{pmatrix} .7071 \\ 0 \\ .7071 \end{pmatrix}$

(b) $\text{tr}(\mathbf{A}) = 2, \quad |\mathbf{A}| = -2$

2.20 **(a)** $\mathbf{C} = \begin{pmatrix} .0000 & .5774 & -.8165 \\ -.7071 & -.5774 & -.4082 \\ .7071 & -.5774 & -.4082 \end{pmatrix}$

(b) $\mathbf{C}'\mathbf{AC} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$

(c) $\mathbf{CDC}' = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix} = \mathbf{A}$

2.21 Eigenvalues: $1, 3$, $\mathbf{C} = \begin{pmatrix} -.7071 & -.7071 \\ -.7071 & .7071 \end{pmatrix}$,

$$\mathbf{A}^{1/2} = \mathbf{CD}^{1/2}\mathbf{C}' = \begin{pmatrix} 1.3660 & -.3660 \\ -.3660 & 1.3660 \end{pmatrix}$$

2.22 (a) The spectral decomposition of $\mathbf{A}$ is given by $\mathbf{A} = \mathbf{CDC}'$, where $\mathbf{C} = \begin{pmatrix} .455 & -.580 & .675 \\ .846 & .045 & -.531 \\ .278 & .813 & .511 \end{pmatrix}$ and $\mathbf{D} = \text{diag}(13.542, 3.935, -2.477)$.

(b) The spectral decomposition of $\mathbf{A}^2$ is given by $\mathbf{A}^2 = \mathbf{CDC}'$, where $\mathbf{C}$ is the same as in part (a) and $\mathbf{D} = \text{diag}(183.378, 15.486, 6.135)$. Note that the diagonal elements of $\mathbf{D}$ are the squares of the diagonal elements of $\mathbf{D}$ in part (a).

(c) The spectral decomposition of $\mathbf{A}^{-1}$ is given by $\mathbf{A}^{-1} = \mathbf{CDC}'$, where $\mathbf{C} = \begin{pmatrix} -.580 & .455 & .675 \\ .045 & .846 & -.531 \\ .813 & .278 & .511 \end{pmatrix}$ and $\mathbf{D} = \text{diag}(.254, .074, -.404)$.
The diagonal elements of $\mathbf{D}$ are the reciprocals of those of $\mathbf{D}$ in part (a). The first two columns of $\mathbf{C}$ have been interchanged to match the interchange of the corresponding elements of $\mathbf{D}$; that is, $\mathbf{D} = (1/\lambda_2, 1/\lambda_1, 1/\lambda_3)$.

2.23 $\mathbf{A} = \mathbf{UDV}'$, where $\mathbf{D} = \text{diag}(13.161, 7{,}000, 3.433)$,

$$\mathbf{U} = \begin{pmatrix} .282 & -.730 & .424 \\ .591 & -.146 & .184 \\ -.225 & .404 & .886 \\ .721 & .531 & -.040 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} .856 & -.015 & .517 \\ -.156 & .946 & .284 \\ .494 & .324 & -.807 \end{pmatrix}.$$

2.24 (a) $\mathbf{j}'\mathbf{a} = (1)a_1 + (1)a_2 + \cdots + (1)a_n = \sum_i a_i = \mathbf{a}'\mathbf{j}$

(b) $\mathbf{j}'\mathbf{A} = [(1)a_{11} + (1)a_{21} + \cdots + (1)a_{n1}, \ldots, (1)a_{1p}$
$+(1)a_{2p} + \cdots + (1)a_{np}]$
$= (\sum_i a_{i1}, \sum_i a_{i2}, \ldots, \sum_i a_{ip})$

(c) $\mathbf{Aj} = \begin{pmatrix} (1)a_{11} & + & (1)a_{12} & + & \cdots & + & (1)a_{1p} \\ (1)a_{21} & + & (1)a_{22} & + & \cdots & + & (1)a_{2p} \\ \vdots & & \vdots & & & & \vdots \\ (1)a_{n1} & + & (1)a_{n2} & + & \cdots & + & (1)a_{np} \end{pmatrix} = \begin{pmatrix} \sum_j a_{1j} \\ \sum_j a_{2j} \\ \vdots \\ \sum_j a_{nj} \end{pmatrix}$

2.25 $(\mathbf{x}-\mathbf{y})'(\mathbf{x}-\mathbf{y}) = (\mathbf{x}'-\mathbf{y}')(\mathbf{x}-\mathbf{y}) = \mathbf{x}'\mathbf{x}-\mathbf{x}'\mathbf{y}-\mathbf{y}'\mathbf{x}+\mathbf{y}'\mathbf{y}$
$= \mathbf{x}'\mathbf{x}-2\mathbf{x}'\mathbf{y}+\mathbf{y}'\mathbf{y}$

2.26 By (2.27), $(\mathbf{A}'\mathbf{A})' = \mathbf{A}'(\mathbf{A}')'$. By (2.6), $(\mathbf{A}')' = \mathbf{A}$. Thus, $(\mathbf{A}'\mathbf{A})' = \mathbf{A}'\mathbf{A}$.

2.27 **(a)**
$$\begin{aligned}\textstyle\sum_i \mathbf{a}'\mathbf{x}_i &= \mathbf{a}'\mathbf{x}_1+\mathbf{a}'\mathbf{x}_2+\cdots+\mathbf{a}'\mathbf{x}_n\\ &= \mathbf{a}'(\mathbf{x}_1+\mathbf{x}_2+\cdots+\mathbf{x}_n) \qquad \text{[by (2.21)]}\\ &= \mathbf{a}'\textstyle\sum_i \mathbf{x}_i\end{aligned}$$

(b)
$$\begin{aligned}\textstyle\sum_i \mathbf{A}\mathbf{x}_i &= \mathbf{A}\mathbf{x}_1+\mathbf{A}\mathbf{x}_2+\cdots+\mathbf{A}\mathbf{x}_n\\ &= \mathbf{A}(\mathbf{x}_1+\mathbf{x}_2+\cdots+\mathbf{x}_n) \qquad \text{[by (2.21)]}\\ &= \mathbf{A}\textstyle\sum_i \mathbf{x}_i\end{aligned}$$

(c)
$$\begin{aligned}\textstyle\sum_i(\mathbf{a}'\mathbf{x}_i)^2 &= \textstyle\sum_i \mathbf{a}'(\mathbf{x}_i\mathbf{x}_i')\mathbf{a} \qquad \text{[by (2.40)]}\\ &= \mathbf{a}'(\textstyle\sum_i \mathbf{x}_i\mathbf{x}_i')\mathbf{a} \qquad \text{[by (2.29)]}\end{aligned}$$

(d) $\sum_i \mathbf{A}\mathbf{x}_i(\mathbf{A}\mathbf{x}_i)' = \sum_i \mathbf{A}\mathbf{x}_i\mathbf{x}_i'\mathbf{A}' = \mathbf{A}(\sum_i \mathbf{x}_i\mathbf{x}_i')\mathbf{A}'$ [by (2.29)]

2.28 **(a)** $\mathbf{A}\mathbf{x} = \begin{pmatrix}\mathbf{a}_1'\\ \mathbf{a}_2'\end{pmatrix}\mathbf{x} = \begin{pmatrix}\mathbf{a}_1'\mathbf{x}\\ \mathbf{a}_2'\mathbf{x}\end{pmatrix}$

(b)
$$\begin{aligned}\mathbf{ASA}' &= \begin{pmatrix}\mathbf{a}_1'\\ \mathbf{a}_2'\end{pmatrix}\mathbf{S}(\mathbf{a}_1,\mathbf{a}_2) = \begin{pmatrix}\mathbf{a}_1'\\ \mathbf{a}_2'\end{pmatrix}(\mathbf{Sa}_1,\mathbf{Sa}_2) \qquad \text{[by (2.48)]}\\ &= \begin{pmatrix}\mathbf{a}_1'\mathbf{Sa}_1 & \mathbf{a}_1'\mathbf{Sa}_2\\ \mathbf{a}_2'\mathbf{Sa}_1 & \mathbf{a}_2'\mathbf{Sa}_2\end{pmatrix}\end{aligned}$$

2.29 **(a)** If $\mathbf{A} = \begin{pmatrix}\mathbf{a}_1'\\ \mathbf{a}_2'\\ \vdots\\ \mathbf{a}_n'\end{pmatrix}$, then by (2.68), $\mathbf{A}' = (\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n)$ and

$$\begin{aligned}\mathbf{A}'\mathbf{A} &= (\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n)\begin{pmatrix}\mathbf{a}_1'\\ \mathbf{a}_2'\\ \vdots\\ \mathbf{a}_n'\end{pmatrix}\\ &= \mathbf{a}_1\mathbf{a}_1'+\mathbf{a}_2\mathbf{a}_2'+\cdots+\mathbf{a}_n\mathbf{a}_n' \qquad \text{[by (2.66)].}\end{aligned}$$

2.30
$$\begin{aligned}&\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}\\ &(\mathbf{A}^{-1}\mathbf{A})' = \mathbf{I}' = \mathbf{I}\\ &\mathbf{A}'(\mathbf{A}^{-1})' = \mathbf{I}\\ &(\mathbf{A}')^{-1}\mathbf{A}'(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}\mathbf{I} = (\mathbf{A}')^{-1}\\ &(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}\end{aligned}$$

2.31 $\frac{1}{b}\begin{pmatrix} b\mathbf{A}_{11}^{-1}+\mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{a}_{12}'\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{a}_{12} \\ -\mathbf{a}_{12}'\mathbf{A}_{11}^{-1} & 1 \end{pmatrix}\begin{pmatrix} \mathbf{A}_{11} & \mathbf{a}_{12} \\ \mathbf{a}_{12}' & a_{22} \end{pmatrix}$

$$= \frac{1}{b}\begin{pmatrix} b\mathbf{I}+\mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{a}_{12}'-\mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{a}_{12}' & b\mathbf{A}_{11}^{-1}\mathbf{a}_{12}+\mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{a}_{12}'\mathbf{A}_{11}^{-1}\mathbf{a}_{12}-\mathbf{A}_{11}^{-1}\mathbf{a}_{12}a_{22} \\ -\mathbf{a}_{12}'+\mathbf{a}_{12}' & -\mathbf{a}_{12}'\mathbf{A}_{11}^{-1}\mathbf{a}_{12}+a_{22} \end{pmatrix}$$

$$= \frac{1}{b}\begin{pmatrix} b\mathbf{I}, & \mathbf{0} \\ \mathbf{0}' & b \end{pmatrix}, \text{ where } b = a_{22}-\mathbf{a}_{12}'\mathbf{A}_{11}^{-1}\mathbf{a}_{12}$$

$$= \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}' & 1 \end{pmatrix}$$

2.32 $(\mathbf{B}+\mathbf{c}\mathbf{c}')\left(\mathbf{B}^{-1}-\dfrac{\mathbf{B}^{-1}\mathbf{c}\mathbf{c}'\mathbf{B}^{-1}}{1+\mathbf{c}'\mathbf{B}^{-1}\mathbf{c}}\right)$

$$= \mathbf{I}-\frac{\mathbf{c}\mathbf{c}'\mathbf{B}^{-1}}{1+\mathbf{c}'\mathbf{B}^{-1}\mathbf{c}}+\mathbf{c}\mathbf{c}'\mathbf{B}^{-1}-\frac{\mathbf{c}(\mathbf{c}'\mathbf{B}^{-1}\mathbf{c})\mathbf{c}'\mathbf{B}^{-1}}{1+\mathbf{c}'\mathbf{B}^{-1}\mathbf{c}} \qquad \text{[by (2.26)]}$$

$$= \mathbf{I}-\mathbf{c}\mathbf{c}'\mathbf{B}^{-1}\left(\frac{1+\mathbf{c}'\mathbf{B}^{-1}\mathbf{c}}{1+\mathbf{c}'\mathbf{B}^{-1}\mathbf{c}}\right)+\mathbf{c}\mathbf{c}'\mathbf{B}^{-1} = \mathbf{I}$$

2.33 $|c\mathbf{A}| = |c\mathbf{I}\mathbf{A}|$

$= |c\mathbf{I}|\,|\mathbf{A}|$ [by (2.89)]

$= c^n|\mathbf{A}|$ [by (2.84)]

2.34 $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$

$|\mathbf{A}\mathbf{A}^{-1}| = |\mathbf{I}|$

$|\mathbf{A}||\mathbf{A}^{-1}| = 1$ [by (2.83)]

$|\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|}$

2.35 In (2.93) and (2.94), let $\mathbf{A}_{11} = \mathbf{B}, \mathbf{A}_{12} = \mathbf{c}, \mathbf{A}_{21} = -\mathbf{c}'$, and $\mathbf{A}_{22} = 1$. Then equate the right-hand sides of (2.93) and (2.94) to obtain (2.95).

2.36 By (2.52), $\mathrm{tr}(\mathbf{A}\mathbf{A}') = \sum_{i=1}^n \mathbf{a}_i'\mathbf{a}_i = \sum_{i=1}^n (a_{i1}^2+a_{i2}^2+\cdots+a_{in}^2)$
$= \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2$.

2.37 Show that $|\mathbf{C}| \neq 0$ by taking the determinant of both sides of $\mathbf{C}'\mathbf{C} = \mathbf{I}$. Thus $\mathbf{C}$ is nonsingular and $\mathbf{C}^{-1}$ exists. Multiply $\mathbf{C}'\mathbf{C} = \mathbf{I}$ on the right by $\mathbf{C}^{-1}$ and on the left by $\mathbf{C}$.

2.38 Multiply $\mathbf{A}\mathbf{B}\mathbf{x} = \lambda\mathbf{x}$ on the left by $\mathbf{B}$. Then λ is an eigenvalue of $\mathbf{B}\mathbf{A}$ and $\mathbf{B}\mathbf{x}$ is an eigenvector.

2.39 **(a)** $(\mathbf{A}^{1/2})^2 = (\mathbf{C}\mathbf{D}^{1/2}\mathbf{C}')^2 = \mathbf{C}\mathbf{D}^{1/2}\mathbf{C}'\mathbf{C}\mathbf{D}^{1/2}\mathbf{C}'$

$= \mathbf{C}\mathbf{D}\mathbf{C}'$ [by (2.101)]

$= \mathbf{A}$ [by (2.109)]

(b) By (2.114), $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$. By (2.89),

$$|\mathbf{A}^{1/2}\mathbf{A}^{1/2}| = |\mathbf{A}|$$

$$|\mathbf{A}^{1/2}||\mathbf{A}^{1/2}| = |\mathbf{A}|$$

$$|\mathbf{A}^{1/2}|^2 = |\mathbf{A}|$$

(c) Since $\mathbf{A}$ is positive definite, we have, from part (b), $|\mathbf{A}^{1/2}| = |\mathbf{A}|^{1/2}$.

2.40 Use properties (2.123) and (2.124). Then use property (2.121) (twice). Finally, use property (2.124).

CHAPTER 3

3.1 $\overline{z} = \sum_{i=1}^{n} z_i/n = \sum_i ay_i/n = (ay_1 + \cdots + ay_n)/n$. Now factor a out of the sum.

3.2 The numerator of s_z^2 is $\sum_{i=1}^{n}(z_i - \overline{z})^2 = \sum_i (ay_i - a\overline{y})^2 = \sum_i [a(y_i - \overline{y})]^2$.

3.3 $\overline{x} = 4, \overline{y} = 4$:

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})(y-\overline{y})$
2	2	−2	−2	4
2	4	−2	0	0
2	6	−2	2	−4
4	2	0	−2	0
4	4	0	0	0
4	6	0	2	0
6	2	2	−2	−4
6	4	2	0	0
6	6	2	2	4
				Sum = 0

3.4 $$\mathbf{x} - \overline{x}\mathbf{j} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} - \overline{x}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} \overline{x} \\ \overline{x} \\ \vdots \\ \overline{x} \end{pmatrix} = \begin{pmatrix} x_1 - \overline{x} \\ x_2 - \overline{x} \\ \vdots \\ x_n - \overline{x} \end{pmatrix}$$

3.5 $$\mathbf{y}_i - \overline{\mathbf{y}} = \begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{pmatrix} - \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \\ \overline{y}_3 \end{pmatrix} = \begin{pmatrix} y_{i1} - \overline{y}_1 \\ y_{i2} - \overline{y}_2 \\ y_{i3} - \overline{y}_3 \end{pmatrix}$$

$$\sum_{i=1}^{n}(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})' = \sum_{i=1}^{n} \begin{pmatrix} y_{i1} - \overline{y}_1 \\ y_{i2} - \overline{y}_2 \\ y_{i3} - \overline{y}_3 \end{pmatrix} (y_{i1} - \overline{y}_1, y_{i2} - \overline{y}_2, y_{i3} - \overline{y}_3)$$

$$= \sum_{i=1}^{n} \begin{pmatrix} (y_{i1} - \overline{y}_1)^2 & (y_{i1} - \overline{y}_1)(y_{i2} - \overline{y}_2) & (y_{i1} - \overline{y}_1)(y_{i3} - \overline{y}_3) \\ (y_{i2} - \overline{y}_2)(y_{i1} - \overline{y}_1) & (y_{i2} - \overline{y}_2)^2 & (y_{i2} - \overline{y}_2)(y_{i3} - \overline{y}_3) \\ (y_{i3} - \overline{y}_3)(y_{i1} - \overline{y}_1) & (y_{i3} - \overline{y}_3)(y_{i2} - \overline{y}_2) & (y_{i3} - \overline{y}_3)^2 \end{pmatrix}$$

3.6 $\overline{z} = \sum_{i=1}^{n} z_i/n = \sum_i \mathbf{a}'\mathbf{y}_i/n = (\mathbf{a}'\mathbf{y}_1 + \cdots + \mathbf{a}'\mathbf{y}_n)/n$. Now factor out $\mathbf{a}'$ on the left. See also (2.42).

3.7 The numerator of s_z^2 is $\sum_{i=1}^n (z_i - \overline{z})^2 = \sum_i (\mathbf{a}'\mathbf{y}_i - \mathbf{a}'\overline{\mathbf{y}})^2 = \sum_i (\mathbf{a}'\mathbf{y}_i - \mathbf{a}'\overline{\mathbf{y}})(\mathbf{a}'\mathbf{y}_i - \mathbf{a}'\overline{\mathbf{y}})$. The scalar $\mathbf{a}'\mathbf{y}_i$ is equal to its transpose, as in (2.39). Thus $\mathbf{a}'\mathbf{y}_i = (\mathbf{a}'\mathbf{y}_i)' = \mathbf{y}_i'\mathbf{a}$, and $\sum_i (\mathbf{a}'\mathbf{y}_i - \mathbf{a}'\overline{\mathbf{y}})(\mathbf{a}'\mathbf{y}_i - \mathbf{a}'\overline{\mathbf{y}}) = \sum_i (\mathbf{a}'\mathbf{y}_i - \mathbf{a}'\overline{\mathbf{y}})(\mathbf{y}_i'\mathbf{a} - \overline{\mathbf{y}}'\mathbf{a})$. By (2.22) and (2.24), this becomes $\sum_i \mathbf{a}'(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})'\mathbf{a}$. Now factor out $\mathbf{a}'$ on the left and $\mathbf{a}$ on the right. See also (2.44).

3.8 By (3.63) and (3.64),

$$\mathbf{ASA}' = \begin{pmatrix} \mathbf{a}_1'\mathbf{Sa}_1 & \mathbf{a}_1'\mathbf{Sa}_2 & \cdots & \mathbf{a}_1'\mathbf{Sa}_k \\ \mathbf{a}_2'\mathbf{Sa}_1 & \mathbf{a}_2'\mathbf{Sa}_2 & \cdots & \mathbf{a}_2'\mathbf{Sa}_k \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_k'\mathbf{Sa}_1 & \mathbf{a}_k'\mathbf{Sa}_2 & \cdots & \mathbf{a}_k'\mathbf{Sa}_k \end{pmatrix},$$

from which the result follows immediately.

3.9
$$\begin{aligned} \mathrm{cov}(\mathbf{z}) &= \mathrm{cov}[(\boldsymbol{\Sigma}^{1/2})^{-1}\overline{\mathbf{y}} - (\boldsymbol{\Sigma}^{1/2})^{-1}\boldsymbol{\mu}] \\ &= (\boldsymbol{\Sigma}^{1/2})^{-1}\mathrm{cov}(\overline{\mathbf{y}})[(\boldsymbol{\Sigma}^{1/2})^{-1}]' \qquad \text{[by (3.76)]} \\ &= (\boldsymbol{\Sigma}^{1/2})^{-1}\left(\frac{\boldsymbol{\Sigma}}{n}\right)(\boldsymbol{\Sigma}^{1/2})^{-1} \\ &= \tfrac{1}{n}(\boldsymbol{\Sigma}^{1/2})^{-1}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}(\boldsymbol{\Sigma}^{1/2})^{-1} \qquad \text{[by (2.114)]} \\ &= \tfrac{1}{n}\mathbf{I} \end{aligned}$$

3.10 Answers are given in Examples 3.7 and 3.8.

3.11 **(a)** $|\mathbf{S}| = 459.956$ **(b)** $\mathrm{tr}(\mathbf{S}) = 213.043$

3.12 **(a)** $|\mathbf{S}| = 27{,}236{,}586$ **(b)** $\mathrm{tr}(\mathbf{S}) = 292.891$

3.13
$$\mathbf{R} = \begin{pmatrix} 1.000 & .614 & .757 & .575 & .413 \\ .614 & 1.000 & .547 & .750 & .548 \\ .757 & .547 & 1.000 & .605 & .692 \\ .575 & .750 & .605 & 1.000 & .524 \\ .413 & .548 & .692 & .524 & 1.000 \end{pmatrix}$$

3.14 $\overline{z} = 83.298, \quad s_z^2 = 1048.659$

3.15 $r_{zw} = -.6106$

3.16 $y_1 = (1,0,0)\mathbf{y} = \mathbf{a}'\mathbf{y}, \quad \frac{1}{2}(y_2 + y_3) = (0, \frac{1}{2}, \frac{1}{2})\mathbf{y} = \mathbf{b}'\mathbf{y}$. Use (3.57) to obtain $r_{zw} = .4873$.

3.17 **(a)**
$$\overline{\mathbf{z}} = \begin{pmatrix} 38.369 \\ 40.838 \\ -51.727 \end{pmatrix}, \quad \mathbf{S}_z = \begin{pmatrix} 323.64 & 19.25 & -460.98 \\ 19.25 & 588.67 & 104.07 \\ -460.98 & 104.07 & 686.27 \end{pmatrix}$$

(b)
$$\mathbf{R}_z = \begin{pmatrix} 1.0000 & .0441 & -.9781 \\ .0441 & 1.0000 & .1637 \\ -.9781 & .1637 & 1.0000 \end{pmatrix}$$

3.18 **(a)** $\overline{\mathbf{y}} = \begin{pmatrix} 48.655 \\ 49.625 \\ 50.570 \\ 51.445 \end{pmatrix}$, $\mathbf{S} = \begin{pmatrix} 6.3300 & 6.1891 & 5.7770 & 5.5348 \\ 6.1891 & 6.4493 & 6.1534 & 5.9057 \\ 5.7770 & 6.1534 & 6.9180 & 6.9267 \\ 5.5348 & 5.9057 & 6.9267 & 7.4331 \end{pmatrix}$,

$$\mathbf{R} = \begin{pmatrix} 1.0000 & .9687 & .8730 & .8069 \\ .9687 & 1.0000 & .9212 & .8530 \\ .8730 & .9212 & 1.0000 & .9659 \\ .8069 & .8530 & .9659 & 1.0000 \end{pmatrix}$$

(b) $|\mathbf{S}| = 1.0865, \quad \mathrm{tr}(\mathbf{S}) = 27.1304$

3.19 **(a)** $\overline{z} = 44.1400, \quad s_z^2 = 21.2309, \quad \overline{w} = 103.8850, \quad s_w^2 = 30.8161$

(b) $s_{zw} = 6.5359, \quad r_{zw} = .2555$

3.20 $\overline{\mathbf{z}} = \begin{pmatrix} 401.40 \\ -47.55 \\ 150.48 \end{pmatrix}$, $\mathbf{S}_z = \begin{pmatrix} 398.33 & -44.35 & 148.35 \\ -44.35 & 12.36 & -16.90 \\ 148.35 & -16.90 & 59.46 \end{pmatrix}$,

$$\mathbf{R}_z = \begin{pmatrix} 1.00 & -.63 & .96 \\ -.63 & 1.00 & -.62 \\ .96 & -.62 & 1.00 \end{pmatrix}$$

3.21 **(a)** $\begin{pmatrix} \overline{\mathbf{y}} \\ \overline{\mathbf{x}} \end{pmatrix} = \left(\begin{array}{c} 185.72 \\ 151.12 \\ \hline 183.84 \\ 149.24 \end{array}\right)$

(b) $\mathbf{S} = \left(\begin{array}{cc|cc} 95.29 & 52.87 & 69.66 & 46.11 \\ 52.87 & 54.36 & 51.31 & 35.05 \\ \hline 69.6 & 51.31 & 100.81 & 56.54 \\ 46.11 & 35.05 & 56.54 & 45.02 \end{array}\right)$

3.22 $\begin{pmatrix} \overline{\mathbf{y}} \\ \overline{\mathbf{x}} \end{pmatrix} = \left(\begin{array}{c} 70.08 \\ 73.54 \\ 75.10 \\ \hline 109.68 \\ 104.24 \\ 109.98 \end{array}\right)$,

$$\mathbf{S} = \left(\begin{array}{ccc|ccc} 95.54 & 17.61 & 12.18 & 60.52 & 23.00 & 62.84 \\ 17.61 & 73.19 & 14.25 & 5.73 & 61.28 & -1.66 \\ 12.18 & 14.25 & 76.17 & 46.75 & 32.77 & 69.84 \\ \hline 60.52 & 5.73 & 46.75 & 808.63 & 320.59 & 227.36 \\ 23.00 & 61.28 & 32.77 & 320.59 & 505.86 & 167.35 \\ 62.84 & -1.66 & 69.84 & 227.36 & 167.35 & 508.71 \end{array}\right)$$

CHAPTER 4

4.1 $|\mathbf{\Sigma}_1| = 1$, $\text{tr}(\mathbf{\Sigma}_1) = 20$, $|\mathbf{\Sigma}_2| = 4$, $\text{tr}(\mathbf{\Sigma}_2) = 15$. Thus $\text{tr}(\mathbf{\Sigma}_1) > \text{tr}(\mathbf{\Sigma}_2)$, but $|\mathbf{\Sigma}_1| < |\mathbf{\Sigma}_2|$. When converted to correlations, we have

$$\mathbf{P}_{\rho 1} = \begin{pmatrix} 1 & .96 & .80 \\ .96 & 1 & .89 \\ .80 & .89 & 1 \end{pmatrix}, \quad \mathbf{P}_{\rho 2} = \begin{pmatrix} 1 & .87 & .41 \\ .87 & 1 & .71 \\ .41 & .71 & 1 \end{pmatrix}.$$

As noted at the end of Section 4.1.3, a decrease in intercorrelations or an increase in the variances will lead to a larger $|\mathbf{\Sigma}|$. In this case, the decrease in correlations from $\mathbf{\Sigma}_1$ to $\mathbf{\Sigma}_2$ outweighed the increase in the variances (the increase in trace).

4.2
$$\begin{aligned} E(\mathbf{z}) &= (\mathbf{T}')^{-1}[E(\mathbf{y}) - \boldsymbol{\mu}] \quad \text{[by (3.75)]} \\ &= (\mathbf{T}')^{-1}[\boldsymbol{\mu} - \boldsymbol{\mu}] = \mathbf{0}, \\ \text{cov}(\mathbf{z}) &= (\mathbf{T}')^{-1}\mathbf{\Sigma}[(\mathbf{T}')^{-1}]' \quad \text{[by (3.76)]} \\ &= (\mathbf{T}')^{-1}\mathbf{T}'\mathbf{T}\mathbf{T}^{-1} \quad \text{[by (2.75) and (2.79)]} \\ &= \mathbf{I} \end{aligned}$$

4.3 By the last expression in Section 2.3.1,

$$\prod_{i=1}^{n} \frac{1}{(\sqrt{2\pi})^p|\mathbf{\Sigma}|^{1/2}} = \frac{1}{(\sqrt{2\pi})^{np}|\mathbf{\Sigma}|^{n/2}}.$$

The sum in the exponent of (4.13) follows from the basic algebra of exponents.

4.4 Since $(\mathbf{y} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})$ is a scalar, we have $E[(\mathbf{y} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})] = E\{\text{tr}[(\mathbf{y} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})]\} = E\{\text{tr}[\mathbf{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})']\} = \text{tr}[\mathbf{\Sigma}^{-1}E(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})'] = \text{tr}(\mathbf{\Sigma}^{-1}\mathbf{\Sigma}) = \text{tr}(\mathbf{I}_p) = p$.

4.5 The other two terms are of the form $\frac{1}{2}\sum_{i=1}^{n}(\overline{\mathbf{y}} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{y}_i - \overline{\mathbf{y}})$, which is equal to $\frac{1}{2}[(\overline{\mathbf{y}} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}]\sum_{i=1}^{n}(\overline{\mathbf{y}}_i - \overline{\mathbf{y}})$. This vanishes because $\sum_{i=1}^{n}(\mathbf{y}_i - \overline{\mathbf{y}}) = n\overline{\mathbf{y}} - n\overline{\mathbf{y}} = \mathbf{0}$.

4.6 We replace y_i in $\sqrt{b_1}$ by $z_i = ay_i + b$. By an extension of (3.3), $\bar{z} = a\bar{y} + b$. Then (4.18) becomes

$$\begin{aligned}\frac{\sqrt{n}\sum_{i=1}^{n}(z_i-\bar{z})^3}{[\sum_{i=1}^{n}(z_i-\bar{z})^2]^{3/2}} &= \frac{\sqrt{n}\sum_i(ay_i+b-a\bar{y}-b)^3}{[\sum_i(ay_i+b-a\bar{y}-b)^2]^{3/2}} \\ &= \frac{\sqrt{n}a^3\sum_i(y_i-\bar{y})^3}{[a^2\sum_i(y_i-\bar{y})^2]^{3/2}} = \frac{\sqrt{n}\sum_i(y_i-\bar{y})^3}{[\sum_i(y_i-\bar{y})^2]^{3/2}} = \sqrt{b_1}.\end{aligned}$$

Similarly, if (4.19) is expressed in terms of $z_i = ay_i + b$, it reduces to b_2 in terms of y_i.

4.7 $\beta_{2,p} = E[(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})]^2$ by (4.33). But when $\mathbf{y}$ is $N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$, $v = (\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})$ is distributed as $\chi^2(p)$ by property 3 in Section 4.2. Then $E(v^2) = \text{var}(v) + [E(v)]^2$.

4.8 To show that $b_{1,p}$ and $b_{2,p}$ are invariant under the transformation $\mathbf{z} = \mathbf{A}\mathbf{y}_i + \mathbf{b}$, where $\mathbf{A}$ is nonsingular, it is sufficient to show that $g_{ij}(\mathbf{z}) = (\mathbf{y}_i-\bar{\mathbf{y}})'\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_j-\bar{\mathbf{y}})$. By (3.67) and (3.68), $\bar{\mathbf{z}} = \mathbf{A}\bar{\mathbf{y}} + \mathbf{b}$ and $\hat{\boldsymbol{\Sigma}}_z = \mathbf{A}\hat{\boldsymbol{\Sigma}}\mathbf{A}'$. Then g_{ij} for $\mathbf{z}$ becomes

$$\begin{aligned}g_{ij}(\mathbf{z}) &= (\mathbf{z}_i-\bar{\mathbf{z}})'\hat{\boldsymbol{\Sigma}}_z^{-1}(\mathbf{z}_j-\bar{\mathbf{z}}) \\ &= (\mathbf{A}\mathbf{y}_i+\mathbf{b}-\mathbf{A}\bar{\mathbf{y}}-\mathbf{b})'(\mathbf{A}\hat{\boldsymbol{\Sigma}}\mathbf{A}')^{-1}(\mathbf{A}\mathbf{y}_j+\mathbf{b}-\mathbf{A}\bar{\mathbf{z}}-\mathbf{b}) \\ &= (\mathbf{y}_i-\bar{\mathbf{y}})'\mathbf{A}'(\mathbf{A}')^{-1}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{A}^{-1}\mathbf{A}(\mathbf{y}_j-\bar{\mathbf{y}}) \\ &= (\mathbf{y}_i-\bar{\mathbf{y}})'\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_j-\bar{\mathbf{y}}) = g_{ij}(\mathbf{y})\end{aligned}$$

4.9 Let $i = (n)$ in (4.47); then solve for $D^2_{(n)}$ in (4.46) and substitute into (4.47) to obtain $F_{(n)}$ in terms of w, as in (4.48).

4.10 **(a)** $\mathbf{a}' = (2,-1,3)$, $\quad z = \mathbf{a}'\mathbf{y}$ is $N(17,21)$

(b) $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix}$, $\quad \mathbf{z} = \mathbf{A}\mathbf{y}$ is $N_2\left[\begin{pmatrix} 8 \\ 10 \end{pmatrix}, \begin{pmatrix} 29 & -1 \\ -1 & 9 \end{pmatrix}\right]$

(c) By property 4b in Section 4.2, y_2 is $N(1,13)$.

(d) By property 4a in Section 4.2, $\begin{pmatrix} y_1 \\ y_3 \end{pmatrix}$ is $N_2\left[\begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 6 & -2 \\ -2 & 4 \end{pmatrix}\right]$.

(e) $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$, $\quad \mathbf{A}\mathbf{y}$ is $N_3\left[\begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 6 & -2 & 3.5 \\ -2 & 4 & 1 \\ 3.5 & 1 & 5.25 \end{pmatrix}\right]$

4.11 **(a)** $\mathbf{z} = \begin{pmatrix} .408 & 0 & 0 \\ -.047 & .279 & 0 \\ .285 & -.247 & .731 \end{pmatrix}\begin{pmatrix} y-3 \\ y-1 \\ y-4 \end{pmatrix}$

(b) $\mathbf{z} = \begin{pmatrix} .465 & -.070 & .170 \\ -.070 & .326 & -.166 \\ .170 & -.166 & .692 \end{pmatrix}\begin{pmatrix} y-3 \\ y-1 \\ y-4 \end{pmatrix}$

(c) By (4.6), $(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})$ is distributed as χ^2_3.

4.12 **(a)** $\mathbf{a}' = (4, -2, 1, -3)$, $z = \mathbf{a}'\mathbf{y}$ is $N(-30, 153)$

(b) $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -2 & 3 & 1 & -2 \end{pmatrix}$, $\mathbf{z} = \mathbf{Ay}$ is $N_2\left[\begin{pmatrix} 5 \\ 2 \end{pmatrix}, \begin{pmatrix} 27 & -79 \\ -79 & 361 \end{pmatrix}\right]$

(c) $\mathbf{A} = \begin{pmatrix} 3 & 1 & -4 & -1 \\ -1 & -3 & 1 & -2 \\ 2 & 2 & 4 & -5 \end{pmatrix}$,

$$\mathbf{z} = \mathbf{Ay} \text{ is } N_3\left[\begin{pmatrix} -4 \\ -18 \\ -27 \end{pmatrix}, \begin{pmatrix} 35 & -18 & -6 \\ -18 & 46 & 14 \\ -6 & 14 & 93 \end{pmatrix}\right]$$

(d) By property 4b in Section 4.2, y_3 is $N(-1, 2)$.

(e) By property 4a in Section 4.2, $\begin{pmatrix} y_2 \\ y_4 \end{pmatrix}$ is $N_2\left[\begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 9 & -6 \\ -6 & 9 \end{pmatrix}\right]$.

(f) $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$,

$$\mathbf{Ay} \text{ is } N_4\left[\begin{pmatrix} -2 \\ .5 \\ 0 \\ 1.25 \end{pmatrix}, \begin{pmatrix} 11 & 1.5 & 2 & 3.75 \\ 1.5 & 1 & .67 & .875 \\ 2 & .67 & .67 & 1 \\ 3.75 & .875 & 1 & 1.688 \end{pmatrix}\right]$$

4.13 **(a)** $\mathbf{z} = \begin{pmatrix} .302 & 0 & 0 & 0 \\ .408 & .561 & 0 & 0 \\ -.087 & .261 & 1.015 & 0 \\ -.858 & -.343 & -.686 & .972 \end{pmatrix}\begin{pmatrix} y+2 \\ y-3 \\ y+1 \\ y-5 \end{pmatrix}$

(b) $\mathbf{z} = \begin{pmatrix} .810 & .305 & .143 & -.480 \\ .305 & .582 & .249 & -.083 \\ .143 & .249 & 1.153 & -.298 \\ -.480 & -.083 & -.298 & .787 \end{pmatrix}\begin{pmatrix} y+2 \\ y-3 \\ y+1 \\ y-5 \end{pmatrix}$

(c) $(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu}) = (\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2}(\mathbf{y}-\boldsymbol{\mu}) = \mathbf{z}'\mathbf{z}$, which is $\chi^2(p) = \chi^2(4)$

4.14 The variables in (b), (c), and (d) are independent.

4.15 The variables in (a), (c), (d), (f), (i), (j), and (n) are independent.

4.16 **(a)** $E(\mathbf{y}|\mathbf{x}) = \boldsymbol{\mu}_y + \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}(\mathbf{x}-\boldsymbol{\mu}_x)$

$$= \begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} -3 & 2 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 5 & -2 \\ -2 & 4 \end{pmatrix}^{-1}\begin{pmatrix} x_1-3 \\ x_2-1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} -.5 & .25 \\ .5 & 1.25 \end{pmatrix} \begin{pmatrix} x_1 - 3 \\ x_2 - 1 \end{pmatrix}$$

$$= \begin{pmatrix} 3.25 \\ -3.75 \end{pmatrix} + \begin{pmatrix} -.5 & .25 \\ .5 & 1.25 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

(b) $\mathrm{cov}(\mathbf{y}|\mathbf{x}) = \mathbf{\Sigma}_{yy} - \mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy}$

$$= \begin{pmatrix} 7 & 3 \\ 3 & 6 \end{pmatrix} - \begin{pmatrix} -3 & 2 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -2 & 4 \end{pmatrix}^{-1} \begin{pmatrix} -3 & 0 \\ 2 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 7 & 3 \\ 3 & 6 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$$

4.17 **(a)** $E(\mathbf{y}|\mathbf{x}) = \boldsymbol{\mu}_y + \mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}(\mathbf{x} - \boldsymbol{\mu}_x)$

$$= \begin{pmatrix} 3 \\ -2 \end{pmatrix} + \begin{pmatrix} 15 & 0 & 3 \\ 8 & 6 & -2 \end{pmatrix} \begin{pmatrix} 50 & 8 & 5 \\ 8 & 4 & 0 \\ 5 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1 - 4 \\ x_2 + 3 \\ x_3 - 5 \end{pmatrix}$$

$$= \begin{pmatrix} 3 \\ -2 \end{pmatrix} - \begin{pmatrix} 15 \\ -24.5 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 3 \\ .67 & .167 & -5.33 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= \begin{pmatrix} -12 \\ 22.5 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 3 \\ .67 & .167 & -5.33 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

(b) $\mathrm{cov}(\mathbf{y}|\mathbf{x}) = \mathbf{\Sigma}_{yy} - \mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy}$

$$= \begin{pmatrix} 14 & -8 \\ -8 & 18 \end{pmatrix} - \begin{pmatrix} 15 & 0 & 3 \\ 8 & 6 & -2 \end{pmatrix} \begin{pmatrix} 50 & 8 & 5 \\ 8 & 4 & 0 \\ 5 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 15 & 8 \\ 0 & 6 \\ 3 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} 14 & -8 \\ -8 & 18 \end{pmatrix} - \begin{pmatrix} 9 & -6 \\ -6 & 17 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}$$

4.18 **(a)** By the central limit theorem in Section 4.3.2, $\sqrt{n}(\overline{\mathbf{y}} - \boldsymbol{\mu})$ is approximately $N_p(\mathbf{0}, \mathbf{\Sigma})$.

(b) $\overline{\mathbf{y}}$ is approximately $N_p(\boldsymbol{\mu}, \mathbf{\Sigma}/n)$.

4.19 **(a)** The plots show almost no deviation from normality.

(b)

Variable	y_1	y_2	y_3	y_4
$\sqrt{b_1}$	.3069	.3111	.0645	.0637
b_2	1.932	2.107	1.792	1.570

The values of $\sqrt{b_1}$ show a small amount of positive skewness, but none exceeds the upper 2.5% critical value for $\sqrt{b_1}$ given in Table A.1 as .942. The values of b_2 show negative kurtosis. For y_4, the kurtosis is significant, since $b_2 < 1.74$, the lower 2.5 percentile in Table A.3.

(c)

Variable	y_1	y_2	y_3	y_4
D	.2848	.2841	.2866	.2851
Y	.4021	.2934	.6730	.4491

From Table A.4, the lower 2.5 percentile for Y is -3.04 and the upper 97.5 percentile is .628. We reject the hypothesis of normality only for y_3.

(d) z defined in (4.24) is approximately $N(0, 3/n)$. To obtain a $N(0,1)$ statistic, we calculate $z^* = z/\sqrt{3/n}$.

Variable	y_1	y_2	y_3	y_4
z^*	$-.3366$	$-.3095$	$-.0737$	$-.0856$

4.20 (a)

i	1	2	3	4	5	6	7	8	9	10
D_i^2	1.06	1.60	7.54	3.54	4.61	.63	.81	2.47	.95	3.78

(b) The .05 critical value from Table A.6 is 7.01. $D_{(10)}^2 = 7.54 > 7.01$.

(c)

i	1	2	3	4	5	6	7	8	9	10
$u_{(i)}$	.08	.10	.12	.13	.20	.30	.44	.47	.57	.93
v_i	.07	.13	.18	.23	.28	.34	.40	.47	.55	.68

The plot of $(v_i, u_{(i)})$ shows some evidence of nonlinearity and an outlier.

(d) $b_{1,p} = 7.255, b_{2,p} = 14.406$. Both (barely) exceed upper .05 critical values in Table A.5.

4.21 (b)

Variable	y_1	y_2	y_3	y_4	y_5
$\sqrt{b_1}$	.2176	.5857	.7461	$-.3327$	$-.1772$
b_2	2.079	1.681	2.583	1.774	2.456

None of the values of $\sqrt{b_1}$ exceeds 1.134 (from Table A.1) or is less than -1.134. None of the values of b_2 is less than 1.53 (from Table A.3). Thus there is no significant departure from normality.

(c)

Variable	y_1	y_2	y_3	y_4	y_5
D	.279	.269	.275	.281	.276
Y	$-.305$	-1.399	$-.805$	$-.114$	$-.669$

(d) $z^* = z/\sqrt{3/n}$, where z is defined in (4.24).

Variable	y_1	y_2	y_3	y_4	y_5
z^*	$-.4848$	-1.7183	-1.3627	.8091	.3686

4.22 (a)

i	1	2	3	4	5	6	7	8	9	10	11
D_i^2	5.20	2.15	7.63	5.34	5.54	1.73	5.21	5.90	2.72	6.02	2.56

(c)

i	1	2	3	4	5	6	7	8	9	10	11
$u_{(i)}$	.19	.24	.28	.30	.57	.57	.59	.61	.65	.66	.84
v_i	.18	.27	.34	.39	.45	.50	.55	.61	.66	.73	.82

The plot shows a sharp break from the fourth to the fifth points.

(d) $b_{1,p} = 12.985$, $b_{2,p} = 29.072$

4.23 **(a)** The $Q - Q$ plots for y_1 and y_5 show little departure from normality. The $Q - Q$ plots for y_2 and y_3 show some evidence of heavier tails than the normal. The $Q - Q$ plots for y_4 and y_6 show some evidence of positive skewness.

(b)

Variable	y_1	y_2	y_3	y_4	y_5	y_6
$\sqrt{b_1}$	.5521	.0302	.7827	1.4627	.2219	.9974
b_2	3.160	3.275	2.772	6.675	2.176	4.528

(c)

Variable	y_1	y_2	y_3	y_4	y_5	y_6
D	.276	.274	.275	.260	.286	.271
Y	-1.469	-1.845	-1.675	-5.249	.889	-2.741

(d)

Variable	y_1	y_2	y_3	y_4	y_5	y_6
z^*	-1.640	$-.062$	-2.803	-2.961	$-.870$	-2.456

4.24 **(a)** $D_i^2 = 7.816, 3.640, 5.730, \ldots, 6.433$

(b) $D_{(51)}^2 = 25.628$. By extrapolation in Table A.6, the .05 critical value for $p = 6$ is approximately 19. Thus we reject the hypothesis of multivariate normality.

(c) $(v_i, u_{(i)}) = (.021, .024), (.029, .028), \ldots, (.306, .523)$. The plot shows nonlinearity for the last four points.

(d) $b_{1,p} = 16.287, b_{2,p} = 58.337$. By extrapolation to $p = 6$ in Table A.5, both appear to exceed their critical values.

4.25 **(a)** $\lambda_1 = 0.5$, $\lambda_2 = -0.3$, $\lambda_3 = 1.1$

(b) $\boldsymbol{\lambda} = (0.7, 0, 1.1)'$

4.26 **(a)** $\lambda_1 = -0.1$, $\lambda_2 = -0.4$, $\lambda_3 = 0.4$

(b) $\boldsymbol{\lambda} = (-0.2, -0.5, 0.5)'$

CHAPTER 5

5.1 By (5.6), we have

$$\begin{aligned}(\bar{\mathbf{y}} - \boldsymbol{\mu}_0)' \left(\frac{\mathbf{S}}{n}\right)^{-1} (\bar{\mathbf{y}} - \boldsymbol{\mu}_0) &= (\bar{\mathbf{y}} - \boldsymbol{\mu}_0)' \left(\frac{1}{n}\right)^{-1} \mathbf{S}^{-1}(\bar{\mathbf{y}} - \boldsymbol{\mu}_0) \\ &= n(\bar{\mathbf{y}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\mathbf{y}} - \boldsymbol{\mu}_0).\end{aligned}$$

5.2 From (5.9), we have

$$\begin{aligned}
&\frac{n_1 n_2}{n_1+n_2}(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1}(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2) \\
&= (\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2)'\left(\frac{n_1+n_2}{n_1 n_2}\right)^{-1}\mathbf{S}_{\text{pl}}^{-1}(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2) \\
&= (\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2)'\left(\frac{n_1+n_2}{n_1 n_2}\mathbf{S}_{\text{pl}}\right)^{-1}(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2) \\
&= (\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2)'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathbf{S}_{\text{pl}}\right]^{-1}(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2).
\end{aligned}$$

5.3 By (5.13) and (5.14),

$$t^2(\mathbf{a}) = \frac{[\mathbf{a}'(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2)]^2}{[(n_1+n_2)/n_1n_2]\mathbf{a}'\mathbf{S}_{\text{pl}}\mathbf{a}} = \frac{n_1n_2}{n_1+n_2}\,\frac{[(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1}(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2)]^2}{(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1}\mathbf{S}_{\text{pl}}\mathbf{S}_{\text{pl}}^{-1}(\overline{\mathbf{y}}_1-\overline{\mathbf{y}}_2)}.$$

5.4 It is assumed that y and x have a bivariate normal distribution. Let $\mathbf{y}_i = \begin{pmatrix} y_i \\ x_i \end{pmatrix}$. Then d_i can be expressed as $d_i = y_i - x_i = \mathbf{a}'\mathbf{y}_i$, where $\mathbf{a}' = (1,-1)$. By property 1a in Section 4.2, d_i is $N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})$. Show that $\mathbf{a}'\overline{\mathbf{y}} = \overline{y}-\overline{x}, \mathbf{a}'\mathbf{S}\mathbf{a} = s_y^2 - 2s_{yx} + s_x^2 = s_d^2$, and that $T^2 = n(\mathbf{a}'\overline{\mathbf{y}})'(\mathbf{a}'\mathbf{S}\mathbf{a})^{-1}(\mathbf{a}'\overline{\mathbf{y}})$ is the square of $t = \overline{d}/(s_d/\sqrt{n})$.

5.5 $\overline{d} = \frac{1}{n}\sum_{i=1}^n d_i = \frac{1}{n}\sum_{i=1}^n (y_i - x_i) = \frac{1}{n}\sum_i y_i - \frac{1}{n}\sum_i x_i = \overline{y}-\overline{x}$,
$s_d^2 = \frac{1}{n-1}\sum_{i=1}^n (d_i-\overline{d})^2 = \frac{1}{n-1}\sum_i (y_i - x_i - \overline{y} + \overline{x})^2$
$= \frac{1}{n-1}\sum_i [(y_i-\overline{y}) - (x_i - \overline{x})]^2$

When this is expanded, we obtain $s_d^2 = s_y^2 + s_x^2 - 2s_{yx}$.

5.6 Similar to Problem 5.1.

5.7 By (5.7), $[(\nu-p+1)/\nu p]T^2_{p,\nu} = F_{p,\nu-p+1}$. By (5.29), $(\nu-q)(T^2_{p+q} - T^2_p)/(\nu+T^2)$ is $T^2_{q,\nu-p}$. Replacing p by q and ν by $\nu-p$ in (5.7), we obtain $\frac{(\nu-p)-q+1}{(\nu-p)q}\,(\nu-p)\,\frac{T^2_{p+q}-T^2_p}{\nu+T^2_p}$ is $F_{q,(\nu-p)-q+1}$.

5.9 Under H_{03}, we have $\mathbf{C}\boldsymbol{\mu}_1 = 0$ and $\mathbf{C}\boldsymbol{\mu}_2 = \mathbf{0}$. Then

$$E(\mathbf{C}\overline{\mathbf{y}}) = \mathbf{C}E(\overline{\mathbf{y}}) = \mathbf{C}E\left(\frac{n_1\overline{\mathbf{y}}_1 + n_2\overline{\mathbf{y}}_2}{n_1+n_2}\right) = \frac{n_1\mathbf{C}\boldsymbol{\mu}_1 + n_2\mathbf{C}\boldsymbol{\mu}_2}{n_1+n_2} = \mathbf{0}.$$

Since $\overline{\mathbf{y}}_1$ and $\overline{\mathbf{y}}_2$ are independent,

$$\begin{aligned}
\text{cov}(\overline{\mathbf{y}}) &= \text{cov}\left(\frac{n_1\overline{\mathbf{y}}_1 + n_2\overline{\mathbf{y}}_2}{n_1+n_2}\right) = \frac{n_1^2\boldsymbol{\Sigma}/n_1 + n_2^2\boldsymbol{\Sigma}/n_2}{(n_1+n_2)^2} \\
&= \frac{(n_1+n_2)\boldsymbol{\Sigma}}{(n_1+n_2)^2}.
\end{aligned}$$

5.10 $\mathbf{CS}_{\mathrm{p1}}\mathbf{C}'/(n_1+n_2)$ is the sample covariance matrix of $\mathbf{C}\overline{\mathbf{y}}$. Hence the equation immediately above (5.39) exhibits the characteristic form of the T^2-distribution.

5.11 $T^2 = .061$

5.12 **(a)** $T^2 = 85.3327$

(b) $t_1 = 2.5039, t_2 = .2665, t_3 = -2.5157, t_4 = .9510, t_5 = .3161$

5.13 $T^2 = 30.2860$

5.14 **(a)** $T^2 = 1.8198$

(b) $t_1 = 1.1643, t_2 = 1.1006, t_3 = .9692, t_4 = .7299$. None of these is significant. In fact, ordinarily they would not have been examined because the T^2-test in part (a) did not reject H_0.

5.15 $T^2 = 79.5510$

5.16 **(a)** $T^2 = 133.4873$

(b) $t_1 = 3.8879, t_2 = -3.8652, t_3 = -5.6911, t_4 = -5.0426$

(c) $\mathbf{a}' = (.345, -.130, -.106, -.143)$

(d) $t^2(\mathbf{a}) = 133.4873$

(e) $R^2 = .782975, T^2 = 133.4873$

(f) By (5.32), $t^2(y_1|y_2, y_3, y_4) = 35.9336, t^2(y_2|y_1, y_3, y_4) = 5.7994,$ $t^2(y_3|y_1, y_2, y_4) = 1.7749, t^2(y_4|y_1, y_2, y_3) = 8.2592$

(g) By (5.29), $T^2(y_3, y_4|y_1, y_2) = 12.5206, F(y_3, y_4|y_1, y_2) = 6.0814$

5.17 By (5.34), the test for parallelism gives $T^2 = 132.6863$. The discriminant function coefficient vector is given by (5.35) as $\mathbf{a}' = (-.362, -.223, -.137)$.

5.18 **(a)** $T^2 = 66.6604$

(b) $t_1 = -.6556, t_2 = 2.6139, t_3 - 3.2884, t_4 = -4.6315, t_5 = 1.8873, t_6 = -3.2205$

(c) By (5.32),

$$t^2(y_1|y_2, y_3, y_4, y_5, y_6) = .0758, t^2(y_2|y_1, y_3, y_4, y_5, y_6) = 6.4513,$$
$$t^2(y_3|y_1, y_2, y_4, y_5, y_6) = 6.9518, t^2(y_4|y_1, y_2, y_3, y_5, y_6) = 6.0309,$$
$$t^2(y_5|y_1, y_2, y_3, y_4, y_6) = 3.7052, t^2(y_6|y_1, y_2, y_3, y_4, y_5) = 6.2619$$

(d) By (5.29), $T^2(y_4, y_5, y_6|y_1, y_2, y_3) = 27.547$

5.19 **(a)** $T^2 = 70.5679$

(b) $T^2(y_5, y_6|y_3, y_4) = 13.1517$

(c) $T^2(y_1, y_2|y_3, y_4, y_5, y_6) = 8.5162$

5.20 **(a)** $T^2 = 18.4625$

(b) $\mathbf{a}' = (-.057, -.010, -.242, -.071)$

(c) By (5.32),

$$t^2(y_1|y_2, y_3, y_4) = 3.3315, t^2(y_2|y_1, y_3, y_4) = .0102,$$
$$t^2(y_3|y_1, y_2, y_4) = 1.4823, t^2(y_4|y_1, y_2, y_3) = .0013$$

5.21 **(a)** $T^2 = 15.1912$

(b) $\mathbf{a}' = (-.036, .048)$

(c) $t_1 = -3.8371, t_2 = -2.4362$

5.22 $T^2 = 22.3238$

5.23 **(a)** $T^2 = 206.1188$

(b) $t^2(d_1|d_2, d_3) = 59.0020, t^2(d_2|d_1, d_3) = 53.4507, t^2(d_3|d_1, d_2) = 80.9349$

CHAPTER 6

6.1 **(a)** Using $\overline{y}_{i.} = y_{i.}/n$, we have

$$\begin{aligned}
\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \overline{y}_{i.})^2 &= \sum_{ij}(y_{ij}^2 - 2y_{ij}\overline{y}_{i.} + \overline{y}_{i.}^2) \\
&= \sum_{ij} y_{ij}^2 - \sum_i \overline{y}_{i.} \sum_j y_{ij} + n\sum_i \overline{y}_{i.}^2 \\
&= \sum_{ij} y_{ij}^2 - 2\sum_i \frac{y_{i.}}{n} y_{i.} + n\sum_i \left(\frac{y_{i.}}{n}\right)^2 \\
&= \sum_{ij} y_{ij}^2 - 2\sum_i \frac{y_{i.}^2}{n} + \sum_i \frac{y_{i.}^2}{n}.
\end{aligned}$$

6.2 $\dfrac{|\mathbf{E}^{-1}|\,|\mathbf{E}|}{|\mathbf{E}^{-1}|\,|\mathbf{E}+\mathbf{H}|} = \dfrac{|\mathbf{E}^{-1}\mathbf{E}|}{|\mathbf{E}^{-1}(\mathbf{E}+\mathbf{H})|} = \dfrac{|\mathbf{I}|}{|\mathbf{I}+\mathbf{E}^{-1}\mathbf{H}|} = \dfrac{1}{\prod_{i=1}^{s}(1+\lambda_i)}$; see Section 2.11.2.

6.3 $(\mathbf{E}^{-1}\mathbf{H} - \lambda\mathbf{I})\mathbf{a} = \mathbf{0}$
$[(\mathbf{E}^{1/2}\mathbf{E}^{1/2})^{-1}\mathbf{H} - \lambda\mathbf{I}]\mathbf{a} = \mathbf{0}$
$[(\mathbf{E}^{1/2})^{-1}(\mathbf{E}^{1/2})^{-1}\mathbf{H} - \lambda\mathbf{I}]\mathbf{a} = \mathbf{0}$
$[(\mathbf{E}^{1/2})^{-1}\mathbf{H} - \lambda\mathbf{E}^{1/2}]\mathbf{a} = \mathbf{0}$
$[(\mathbf{E}^{1/2})^{-1}\mathbf{H} - \lambda\mathbf{E}^{1/2}](\mathbf{E}^{1/2})^{-1}\mathbf{E}^{1/2}\mathbf{a} = \mathbf{0}$
$[(\mathbf{E}^{1/2})^{-1}\mathbf{H}(\mathbf{E}^{1/2})^{-1} - \lambda\mathbf{I}]\mathbf{E}^{1/2}\mathbf{a} = \mathbf{0}$

6.4 We need to show that $(2N+s+1)/(2m+s+1) = (\nu_E - p + s)/d$. Using the definitions $N = \frac{1}{2}(\nu_E - p - 1), m = \frac{1}{2}(|\nu_H - p| - 1), d = \max(p, \nu_H)$, and $s = \min(p, \nu_H)$, we have $2N + s + 1 = 2(\frac{1}{2})(\nu_E - p - 1) + s + 1 = \nu_E - p - 1 + s + 1 = \nu_E - p + s$. For the denominator, we have $2m + s + 1 = 2(\frac{1}{2})(|\nu_H - p| - 1) + s + 1 = |\nu_H - p| + s$. Suppose $\nu_H > p$. Then $|\nu_H - p| + s = \nu_H - p + p = \nu_H = d$. On the other hand, if $\nu_H < p$, then $|\nu_H - p| + s = p - \nu_H + \nu_H = p = d$.

6.5 If $p \leq \nu_H$, we have $s = p$ and $|\nu_H - p| = \nu_H - p$. Then (6.29) becomes

$$\begin{aligned}\frac{2(sN+1)U^{(s)}}{s^2(2m+s+1)} &= \frac{2\left[p\left(\frac{1}{2}\right)(\nu_E - p - 1) + 1\right]U^{(s)}}{p^2[2(\frac{1}{2})(\nu_H - p - 1) + p + 1]} \\ &= \frac{[p(\nu_E - p - 1) + 2]U^{(s)}}{p^2(\nu_H - p - 1 + p + 1)} \\ &= \frac{[p(\nu_E - p - 1) + 2]U^{(s)}}{p^2\nu_H},\end{aligned}$$

which is the same as (6.30) because $p = s$. If $\nu_H \leq p$, then $s = \nu_H$, $|\nu_H - p| = p - \nu_H$, and (6.29) can be shown to equal (6.30) in a similar manner.

6.6 When $s = 1$, we have $V^{(1)} = \lambda_1/(1+\lambda_1), U^{(1)} = \lambda_1, \Lambda = 1/(1+\lambda_1)$, and $\theta = \lambda_1/(1+\lambda_1)$. Solving the last of these for λ_1 gives $\lambda_1 = \theta/(1-\theta)$, and the results in (6.33), (6.34), and (6.35) follow immediately.

6.7 With $T^2 = (n_1 + n_2 - 2)U^{(1)}$ and $U^{(1)} = \theta/(1-\theta)$, we obtain (5.19). We obtain (5.18) from (5.19) by $V^{(1)} = \theta$. A similar argument leads to (5.16).

6.8 **(a)** With $\overline{\mathbf{y}}_{i.} = \mathbf{y}_{i.}/n_i$ and $\overline{\mathbf{y}}_{..} = \mathbf{y}_{..}/N$, we obtain

$$\begin{aligned}\mathbf{H} &= \textstyle\sum_{i=1}^k n_i(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})' \\ &= \textstyle\sum_i n_i(\overline{\mathbf{y}}_{i.}\overline{\mathbf{y}}_{i.}' - \overline{\mathbf{y}}_{i.}\overline{\mathbf{y}}_{..}' - \overline{\mathbf{y}}_{..}\overline{\mathbf{y}}_{i.}' + \overline{\mathbf{y}}_{..}\overline{\mathbf{y}}_{..}') \\ &= \textstyle\sum_i n_i\overline{\mathbf{y}}_{i.}\overline{\mathbf{y}}_{i.}' - (\sum_i n_i\overline{\mathbf{y}}_{i.})\overline{\mathbf{y}}_{..}' - \overline{\mathbf{y}}_{..}\sum_i n_i\overline{\mathbf{y}}_{i.}' + \overline{\mathbf{y}}_{..}\overline{\mathbf{y}}_{..}'\sum_i n_i \\ &= \sum_i n_i\frac{\mathbf{y}_{i.}\mathbf{y}_{i.}'}{n_i^2} - \frac{(\sum_i \mathbf{y}_{i.})\mathbf{y}_{..}'}{N} - \frac{\mathbf{y}_{..}}{N}\sum_i \mathbf{y}_{i.}' + \frac{N\mathbf{y}_{..}\mathbf{y}_{..}'}{N^2} \\ &= \sum_i \frac{\mathbf{y}_{i.}\mathbf{y}_{i.}'}{n_i} - \frac{\mathbf{y}_{..}\mathbf{y}_{..}'}{N} - \frac{\mathbf{y}_{..}\mathbf{y}_{..}'}{N} + \frac{\mathbf{y}_{..}\mathbf{y}_{..}'}{N}.\end{aligned}$$

6.9 $\overline{\mathbf{y}}_{1.} - \overline{\mathbf{y}}_{..}$ becomes

$$\overline{\mathbf{y}}_{1.} - \frac{n_1\overline{\mathbf{y}}_{1.} + n_2\overline{\mathbf{y}}_{2.}}{n_1+n_2} = \frac{n_1\overline{\mathbf{y}}_{1.} + n_2\overline{\mathbf{y}}_{1.} - n_1\overline{\mathbf{y}}_{1.} - n_2\overline{\mathbf{y}}_{2.}}{n_1+n_2} = \frac{n_2(\overline{\mathbf{y}}_{1.} - \overline{\mathbf{y}}_{2.})}{n_1+n_2}.$$

The first term in the sum is

$$\frac{n_1 n_2^2}{(n_1+n_2)^2}(\overline{\mathbf{y}}_{1.} - \overline{\mathbf{y}}_{2.})(\overline{\mathbf{y}}_{1.} - \overline{\mathbf{y}}_{2.})'.$$

The second term in the sum is

$$\frac{n_1^2 n_2}{(n_1+n_2)^2}(\overline{\mathbf{y}}_{1.} - \overline{\mathbf{y}}_{2.})(\overline{\mathbf{y}}_{1.} - \overline{\mathbf{y}}_{2.})'.$$

6.10 $\theta = \dfrac{\lambda_1}{1+\lambda_1} = \dfrac{\text{SSH}(z)/\text{SSE}(z)}{1+\text{SSH}(z)/\text{SSE}(z)} = \dfrac{\text{SSH}(z)}{\text{SSE}(z)+\text{SSH}(z)}$

6.11 From $r_i^2 = \lambda_i/(1+\lambda_i)$, obtain $\lambda_i = r_i^2/(1-r_i^2)$. Substitute this into $1/(1+\lambda_i)$ to obtain the result.

6.12 Substitute $A_p = V^{(s)}/s$ into (6.49) to obtain (6.26).

6.13 When $s = 1$, (6.50) becomes

$$A_{\text{LH}} = \frac{U^{(1)}}{1+U^{(1)}}.$$

By (6.33), $U^{(1)} = \lambda_1$.

6.14 Substitute $A_{\text{LH}} = U^{(s)}/(s+U^{(s)})$ from (6.50) into (6.51) to obtain F_3 in (6.30).

6.15 To show $\text{cov}(c_i\overline{\mathbf{y}}_{i.}) = c_i^2\mathbf{\Sigma}/n$, use (3.74), $\text{cov}(\mathbf{Ay}) = \mathbf{A\Sigma A}'$, with $\mathbf{A} = c_i\mathbf{I}$.

6.16 By (6.9),

$$\begin{aligned}
\mathbf{H}_z &= n\textstyle\sum_{i=1}^k(\overline{\mathbf{z}}_{i.} - \overline{\mathbf{z}}_{..})(\overline{\mathbf{z}}_{i.} - \overline{\mathbf{z}}_{..})' \\
&= n\textstyle\sum_i(\mathbf{C}\overline{\mathbf{y}}_{i.} - \mathbf{C}\overline{\mathbf{y}}_{..})(\mathbf{C}\overline{\mathbf{y}}_{i.} - \mathbf{C}\overline{\mathbf{y}}_{..})' \\
&= n\textstyle\sum_i[\mathbf{C}(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})][\mathbf{C}(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})]' \\
&= n\mathbf{C}\left[\textstyle\sum_i(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})'\right]\mathbf{C}' \quad [\text{by}(2.45)]
\end{aligned}$$

6.17 $\mathbf{C}$ is not square.

6.18
$$\begin{aligned}
E(\mathbf{C}\overline{\mathbf{y}}_{..}) &= \mathbf{C}E(\overline{\mathbf{y}}_{..}) = \mathbf{C}E(\textstyle\sum_{i=1}^k \overline{\mathbf{y}}_{i.}/k) \\
&= \mathbf{C}\textstyle\sum_i E(\overline{\mathbf{y}}_{i.})/k = \mathbf{C}\sum_i \boldsymbol{\mu}_i/k \\
&= \mathbf{0} \quad [\text{by } H_{03} \text{ in } (6.82)]
\end{aligned}$$

$\text{cov}(\mathbf{C}\overline{\mathbf{y}}_{..}) = \mathbf{C\Sigma C}'/kn$ if there are no differences in the group means, $\mathbf{C}\boldsymbol{\mu}_1, \mathbf{C}\boldsymbol{\mu}_2, \ldots$. This condition is assured by H_{01} in (6.77).

6.19 For our purposes, it will suffice to show that T^2 has the characteristic form of the T^2-distribution in (5.6).

6.20 If $\mathbf{\Sigma} = \sigma^2\mathbf{I}$, (6.88) becomes

$$\varepsilon = \frac{[\text{tr}(\sigma^2\mathbf{I} - \mathbf{J}\sigma^2\mathbf{I}/p)]^2}{(p-1)\text{tr}(\sigma^2\mathbf{I} - \mathbf{J}\sigma^2\mathbf{I}/p)^2} = \frac{[\sigma^2\text{tr}(\mathbf{I} - \mathbf{J}/p)]^2}{\sigma^4(p-1)\text{tr}(\mathbf{I} - \mathbf{J}/p)^2}.$$

Show that $(\mathbf{I} - \mathbf{J}/p)^2 = \mathbf{I} - \mathbf{J}/p$. Then

$$\varepsilon = \frac{\sigma^4(p - p/p)^2}{\sigma^4(p-1)(p-p/p)} = \frac{(p-1)^2}{(p-1)^2} = 1.$$

6.21 The (univariate) expected mean square corresponding to $\overline{\mu}.$ in a one-way ANOVA is $\sigma^2 + N\mu^2$. Thus the mean square for $\overline{\mu}.$ is tested with MSE. The corresponding multivariate test therefore uses $\mathbf{H}^*$ and $\mathbf{E}$.

6.22 From (6.104) we have

$$\Lambda = \frac{|\mathbf{AEA}'|}{|\mathbf{A}(\mathbf{E}+\mathbf{H}^*)\mathbf{A}'|} = \frac{|\mathbf{AEA}'|}{|\mathbf{AEA}' + \mathbf{AH}^*\mathbf{A}'|}.$$

Substitute $\mathbf{H}^* = kn\overline{\mathbf{y}}..\overline{\mathbf{y}}'..$ to obtain

$$\Lambda = \frac{|\mathbf{AEA}'|}{|\mathbf{AEA}' + \sqrt{kn}\mathbf{A}\overline{\mathbf{y}}..(\sqrt{kn}\mathbf{A}\overline{\mathbf{y}}..)'|}.$$

Now use (2.95) with $\mathbf{B} = \mathbf{AEA}'$ and $\mathbf{c} = \sqrt{kn}\mathbf{A}\overline{\mathbf{y}}..$ to obtain

$$\Lambda = \frac{1}{1 + kn(\mathbf{A}\overline{\mathbf{y}}..)'(\mathbf{AEA}')^{-1}(\mathbf{A}\overline{\mathbf{y}}..)}.$$

Multiply and divide by ν_E and use (6.100) to obtain (6.105).

6.23 Solve for T^2 in (6.105).

6.24 In $\mathbf{C}_1\mathbf{B}'$ the rows of $\mathbf{C}_1$ are multiplied by the rows of $\mathbf{B}$. Show that $\mathbf{C}_1\mathbf{B}' = \mathbf{O}$.

6.25 As noted, the function $(\overline{\mathbf{y}} - \mathbf{A}\hat{\boldsymbol{\beta}})'\mathbf{S}^{-1}(\overline{\mathbf{y}} - \mathbf{A}\hat{\boldsymbol{\beta}})$ is similar to SSE $= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$ in (10.4) and (10.6). By an argument similar to that used in Section 10.2.2 to obtain $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, it follows that $\hat{\boldsymbol{\beta}} = (\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{S}^{-1}\overline{\mathbf{y}}$. An alternative approach (for those familiar with differentiation with respect to a vector) is to expand $(\overline{\mathbf{y}} - \mathbf{A}\hat{\boldsymbol{\beta}})'\mathbf{S}^{-1}(\overline{\mathbf{y}} - \mathbf{A}\hat{\boldsymbol{\beta}})$ to four terms, differentiate with respect to $\hat{\boldsymbol{\beta}}$, and set the result equal to $\mathbf{0}$.

6.26 Expand $n(\overline{\mathbf{y}} - \mathbf{A}\hat{\boldsymbol{\beta}})'\mathbf{S}^{-1}(\overline{\mathbf{y}} - \mathbf{A}\hat{\boldsymbol{\beta}})$ to four terms and substitute

$$\hat{\boldsymbol{\beta}} = (\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{S}^{-1}\overline{\mathbf{y}}$$

into the last one.

6.27 **(a)** $$\mathbf{E} = \begin{pmatrix} 13.41 & 7.72 & 8.68 & 5.86 \\ 7.72 & 8.48 & 7.53 & 6.21 \\ 8.68 & 7.53 & 11.61 & 7.04 \\ 5.86 & 6.21 & 7.04 & 10.57 \end{pmatrix}$$

$$\mathbf{H} = \begin{pmatrix} 1.05 & 2.17 & -1.38 & -.76 \\ 2.17 & 4.88 & -2.37 & -1.26 \\ -1.38 & -2.37 & 2.38 & 1.38 \\ -.76 & -1.26 & 1.38 & .81 \end{pmatrix}$$

$\Lambda = .224, V^{(s)} = .860, U^{(s)} = 3.08$, and $\theta = .747$ All four are significant.

(b) $\eta^2_\Lambda = 1 - \Lambda = .776, \eta^2_\theta = \theta = .747, A_\Lambda = 1 - \Lambda^{1/s} = .526, A_{\text{LH}} = .606, A_p = V^{(s)}/s = .430$

(c) The eigenvalues of $\mathbf{E}^{-1}\mathbf{H}$ are 2.9515 and .1273. The essential dimensionality of the space of the mean vectors is 1.

(d) For 1, 2 vs. 3 we have $\Lambda = .270, V^{(s)} = .730, U^{(s)} = 2.702$, and $\theta = .730$. All four are significant. For 1 vs. 2 we obtain $\Lambda = .726, V^{(s)} = .274, U^{(s)} = .377$, and $\theta = .274$. All four are significant.

(e)

Variable	y_1	y_2	y_3	y_4
F	1.29	9.50	3.39	1.27

The F's for y_2 and y_3 are significant. For the discriminant function $z = \mathbf{a}'\mathbf{y}$, where $\mathbf{a}$ is the first eigenvector of $\mathbf{E}^{-1}\mathbf{H}$, we have $\mathbf{a}' = (-.032, -.820, .533, .208)$. Again y_2 and y_3 contribute most to separation of groups.

(f) By (6.126), $\Lambda(y_3, y_4|y_1, y_2) = \Lambda(y_1, y_2, y_3, y_4)/\Lambda(y_1, y_2) = .224/.568 = .395 < \Lambda_{.05} = .725$

(g) By (6.127),

$$\begin{aligned} \Lambda(y_1|y_2, y_3, y_4) &= \Lambda(y_1, y_2, y_3, y_4)/\Lambda(y_2, y_3, y_4) \\ &= .224/.240 = .934 > \Lambda_{.05} = .819 \\ \Lambda(y_2|y_1, y_3, y_4) &= .224/.538 = .417 < .819 \\ \Lambda(y_3|y_1, y_2, y_4) &= .224/.369 = .609 < .819 \\ \Lambda(y_4|y_1, y_2, y_3) &= .224/.243 = .924 > .819 \end{aligned}$$

6.28 **(a)** S effect: $\Lambda = .00065, V^{(s)} = 2.357, U^{(s)} = 142.304, \theta = .993$. All are significant.

V effect: $\Lambda = .065, V^{(s)} = 1.107, U^{(s)} = 11.675, \theta = .920$. All are significant.

SV interaction: $\Lambda = .138, V^{(s)} = 1.321, U^{(s)} = 3.450, \theta = .726$. All are significant.

(b) Contrast on V comparing 2 vs. 1, 3: $\Lambda = .0804, V^{(s)} = .920, U^{(s)} = 11.445, \theta = .920$. All are significant.

(c) Linear contrast for S: $\Lambda = .0073$, $V^{(S)} = .993, U^{(s)} = 135.273, \theta = .993$. All are significant.

Quadratic contrast for S: $\Lambda = .168, V^{(s)} = .832, U^{(s)} = 4.956, \theta = .832$. All are significant.

Cubic contrast for S: $\Lambda = .325, V^{(s)} = .675, U^{(s)} = 2.076, \theta = .675$. All are significant.

(d) The ANOVA F's for each variable are as follows:

Source	y_1	y_2	y_3	y_4
S	980.21	214.24	876.13	73.91
V	251.22	9.47	14.77	27.12
SV	20.37	2.84	3.44	2.08

All F's are significant except the last one, 2.08.

(e) Test of significance of y_3 and y_4 adjusted for y_1 and y_2:

	S	V	SV
$\Lambda(y_3, y_4 \mid y_1, y_2)$	.1226	.9336	.6402

(f) Test of significance of each variable adjusted for the other three:

	S	V	SV
$\Lambda(y_1 \mid y_2, y_3, y_4)$	.1158	.2099	.3082
$\Lambda(y_2 \mid y_1, y_3, y_4)$	.5586	.8134	.7967
$\Lambda(y_3 \mid y_1, y_2, y_4)$	.2271	.9627	.7604
$\Lambda(y_4 \mid y_1, y_2, y_3)$	.6692	.9795	.8683

6.29 V = velocity (fixed), L = lubricant (random)

V effect (using $\mathbf{H}_{VL}$ for error matrix): $\Lambda = .0492, V^{(s)} = .951, U^{(s)} = 19.315, \theta = .951$. With $p = 2, \nu_H = 1$, and $\nu_E = 3, \Lambda_{.05} = .050, V^{(s)}_{.05} = .950, U^{(s)}_{.05} = T^2_{.05}/\nu_E = 19.00, \theta_{.05} = .950$. Thus all four test statistics are significant.

L effect (using $\mathbf{E}$ for error matrix): $\Lambda = .692, V^{(s)} = .314, U^{(s)} = .438, \theta = .295$. None is significant.

VL interaction (using $\mathbf{E}$ for error matrix): $\Lambda = .932, V^{(s)} = .069, U^{(s)} = .073, \theta = .061$. None is significant.

6.30

Source	Λ	$V^{(s)}$	$U^{(s)}$	θ	Significant?
(a) Reagent	.0993	1.126	6.911	.868	Yes
(b) Contrast 1 vs. 2, 3, 4	.146	.854	5.871	.854	Yes
Subjects	.00000082	2.847	1091.127	.999	Yes

6.31 P = proportion of filler, T = surface treatment, F = filler:

Source	Λ	$V^{(s)}$	$U^{(s)}$	θ	Significant?
P	.138	.977	5.441	.841	Yes
T	.080	.920	11.503	.920	Yes
PT	.712	.295	.396	.271	No
F	.019	.980	51.180	.981	Yes
PF	.179	.958	3.835	.784	Yes
TF	.355	.645	1.815	.645	Yes
PTF	.752	.264	.309	.172	No

6.32 A = period; P, T, and F are defined in Problem 6.31:

Source	Λ	$V^{(s)}$	$U^{(s)}$	θ	Significant?
A	.021	.979	47.099	.979	Yes
AP	.475	.545	1.063	.505	No
AT	.142	.858	6.049	.858	Yes
APT	.777	.228	.282	.208	No
AF	.095	.905	9.486	.905	Yes
APF	.622	.387	.594	.363	No
ATF	.387	.613	1.586	.613	Yes
$APTF$	.781	.229	.267	.169	No

For the between-subject factors and interactions, we have

Source	df	F	p-Value
P	2	21.79	$< .0001$
T	1	78.34	$< .0001$
PT	2	1.28	.3143
F	1	345.04	$< .0001$
PF	2	15.79	.0004
TF	1	5.36	.0392
PTF	2	.48	.6294
Error	12		

6.33 For parallelism, we use (6.78) to obtain $\Lambda = .2397$. For levels, we use (6.80) and (6.81) to obtain $\Lambda = .9651$ and $F = .597$. For flatness we use (6.83) to obtain $T^2 = 110.521$.

6.34 **(a)** By (6.89), $T^2 = 20.7420$. By (6.104) or (6.105), $\Lambda = .5655$.

(b) For each row $\mathbf{c}_i'$ of $\mathbf{C}$, we use $T_i^2 = n(\mathbf{c}_i'\overline{\mathbf{y}})'(\mathbf{c}_i'\mathbf{S}\mathbf{c}_i)^{-1}\mathbf{c}_i'\overline{\mathbf{y}}$, as in Example 6.9.2: $T_1^2 = 17.0648, T_2^2 = .3238, T_3^2 = .2714$. This can also be done by Wilks' Λ using $\Lambda_i = \mathbf{c}_i'\mathbf{E}\mathbf{c}_i/\mathbf{c}_i'(\mathbf{E}+\mathbf{H}^*)\mathbf{c}_i$: $\Lambda_1 = .6127, \Lambda_2 = .9882, \Lambda_3 = .9900$.

6.35 The six variables represent two within-subjects factors: y_1 is A_1B_1, y_2 is A_1B_2, y_3 is A_1B_3, x_1 is A_2B_1, x_2 is A_2B_2, and x_3 is A_2B_2. Using linear and quadratic effects (other orthogonal contrasts could be used), the matrices

$\mathbf{A}, \mathbf{B}$, and $\mathbf{G}$ in (6.97), (6.98), and (6.99) become

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

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & -1 & 1 & 0 & -1 \\ 1 & -2 & 1 & 1 & -2 & 1 \end{pmatrix}$$

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 1 \\ 1 & -2 & 1 & -1 & 2 & -1 \end{pmatrix}.$$

Using these in T^2 as given by (6.100), (6.101), and (6.102), we obtain $T_A^2 = 193.0901, T_B^2 = 2.8000$, and $T_{AB}^2 = 6.8676$. Using MANOVA tests for the same within-subjects factors, we obtain

Source	Λ	$V^{(s)}$	$U^{(s)}$	θ	Significant?
A	.202	.798	3.941	.798	Yes
B	.946	.054	.057	.054	No
AB	.877	.123	.140	.123	Yes

6.36 MANOVA tests for the within-subjects effect T (time), and interactions of time with the between-subjects effects C (cancer) and G (gender):

Source	Λ	$V^{(s)}$	$U^{(s)}$	θ
T	.258	.742	2.874	.742
TC	.363	.809	1.299	.444
TG	.929	.071	.077	.071
TCG	.809	.201	.225	.130

ANOVA F-tests for between-subjects factors and interactions:

Source	df	F	p-Value
C	5	4.16	.003
G	1	2.69	.107
CG	5	.37	.869

6.37 **(a)** $T^2 = 79.551$

(b) Using $t_i = \mathbf{c}_i'\overline{\mathbf{y}}/\sqrt{\mathbf{c}_i'\mathbf{S}\mathbf{c}_i/n}$, where $\mathbf{c}_i'$ is the ith row of $\mathbf{C}$, we obtain $t_1 = 7.155, t_2 = -.445, t_3 = -.105$.

6.38 **(a)** $T^2 = 1712.2201$

(b) Using $t_i = \mathbf{c}_i'\overline{\mathbf{y}}/\sqrt{\mathbf{c}_i'\mathbf{S}\mathbf{c}_i/n}$, we obtain $t_1 = 332.358, t_2 = 54.589, t_3 = .056, t_4 = 7.637, t_5 = 4.344, t_6 = 1.968$.

6.39 **(a)** Using $T^2 = N(\mathbf{C}\overline{\mathbf{y}}_{..})'(\mathbf{C}\mathbf{S}_{p1}\mathbf{C}')^{-1}(\mathbf{C}\overline{\mathbf{y}}_{..})$ in (6.121), we obtain $T^2 = 17.582 < T^2_{.05,3,9} = 27.202$.

(b) $t_1 = .951, t_2 = 1.606, t_3 = .127$ [Since the T^2-test in part (a) did not reject H_0, these would ordinarily not be calculated.]

(c) Using $\Lambda = |\mathbf{CEC}'|/|\mathbf{C}(\mathbf{E}+\mathbf{H})\mathbf{C}'|$ in (6.123), we obtain $\Lambda = .3107 > \Lambda_{.05,3,2,9} = .203$.

(d) To compare groups using each row of $\mathbf{C}$, we use $\Lambda_i = \mathbf{c}_i'\mathbf{E}\mathbf{c}_i/\mathbf{c}_i'(\mathbf{E}+\mathbf{H})\mathbf{c}_i$ to obtain $\Lambda_1 = .833, \Lambda_2 = .988, \Lambda_3 = .650$. [Since the Λ-test in part (c) did not reject H_0, we would ordinarily not have calculated these.]

6.40 (a) Using $T^2 = N(\mathbf{C}\overline{\mathbf{y}}_{..})'(\mathbf{C}\mathbf{S}_{\text{p}1}\mathbf{C}')^{-1}(\mathbf{C}\overline{\mathbf{y}}_{..})$ in (6.121), we obtain $T^2 = 33.802 > T^2_{.05,4,24} = 12.983$.

(b) Using $t_i^2 = N(\mathbf{c}_i'\overline{\mathbf{y}})^2/\mathbf{c}_i'\mathbf{S}_{\text{p}1}\mathbf{c}_i$, we obtain $t_1^2 = .675, t_2^2 = .393, t_3^2 = 32.626$. Only the cubic effect is significant.

(c) For an overall test comparing groups, we use (6.123),

$$\Lambda = |\mathbf{C}\mathbf{E}\mathbf{C}'|/|\mathbf{C}(\mathbf{E}+\mathbf{H})\mathbf{C}'| = .4361.$$

(d) To compare groups using each row of $\mathbf{C}$, we use $\Lambda_i = \mathbf{c}_i'\mathbf{E}\mathbf{c}_i/\mathbf{c}_i'(\mathbf{E}+\mathbf{H})\mathbf{c}_i$: $\Lambda_1 = .534, \Lambda_2 = .764, \Lambda_3 = .941$.

6.41 (a) Using $T^2 = N(\mathbf{C}\overline{\mathbf{y}}_{..})'(\mathbf{C}\mathbf{S}_{\text{p}1}\mathbf{C}')^{-1}(\mathbf{C}\overline{\mathbf{y}}_{..})$ in (6.121), we obtain $T^2 = 45.500$.

(b) Using $t_i^2 = N(\mathbf{c}_i'\overline{\mathbf{y}})^2/\mathbf{c}_i'\mathbf{S}_{\text{p}1}\mathbf{c}_i$, we obtain $t_1^2 = 18.410, t_2^2 = 8.385, t_3^2 = 3.446, t_4^2 = .011, t_5^2 = .098, t_6^2 = 2.900$.

(c) For an overall test comparing groups, we use (6.123),

$$\Lambda = |\mathbf{C}\mathbf{E}\mathbf{C}'|/|\mathbf{C}(\mathbf{E}+\mathbf{H})\mathbf{C}'| = .304.$$

(d) To compare groups using each row of $\mathbf{C}$, we use $\Lambda_i = \mathbf{c}_i'\mathbf{E}\mathbf{c}_i/\mathbf{c}_i'(\mathbf{E}+\mathbf{H})\mathbf{c}_i$: $\Lambda_1 = .695, \Lambda_2 = .925, \Lambda_3 = .731, \Lambda_4 = .814, \Lambda_5 = .950, \Lambda_6 = .894$.

6.42 (a) Combined groups (pooled covariance matrix). Using t = number of minutes -30, we obtain, by (6.114),

$$\hat{\boldsymbol{\beta}}' = (98.1, .981, .0418, -.00101, -.000048)$$

By (6.115), we obtain $T^2 = .216$. By (6.117), we have

$$\hat{\boldsymbol{\mu}}' = (95.5, 96.7, 95.6, 93.8, 98.1, 99.2)$$

(b) Group 1: $\hat{\boldsymbol{\beta}}_1' = (100.7, .819, .040, -.00085, -.000038), T^2 = .0113$, $\hat{\boldsymbol{\mu}}_1' = (105.2, 104.4, 101.5, 98.6, 100.6, 108.1)$

(c) Groups 2–4: $\hat{\boldsymbol{\beta}}_2' = (97.4, 1.010, .0403, -.00103, -.000049), T^2 = .2554, \hat{\boldsymbol{\mu}}_2' = (92.6, 94.4, 93.8, 92.4, 97.4, 96.6)$

6.43 (a) For the control group, the overall test is

$$T^2 = n_1(\mathbf{C}\overline{\mathbf{y}}_{1.})'(\mathbf{C}\mathbf{S}_1\mathbf{C}')^{-1}(\mathbf{C}\overline{\mathbf{y}}_{1.}) = 554.749.$$

For each row of $\mathbf{C}$ (linear, quadratic, etc.), we have

$$t_i^2 = n_1(\mathbf{c}_i'\overline{\mathbf{y}}_{1.})^2/\mathbf{c}_i'\mathbf{S}_1\mathbf{c}_i:$$

$t_1^2 = 5.714, t_2^2 = 50.111, t_3^2 = 50.767, t_4^2 = 8.011, t_5^2 = .508.$

(b) For the obese group, we obtain $T^2 = n_2(\mathbf{C}\overline{\mathbf{y}}_{2.})'(\mathbf{C}\mathbf{S}_2\mathbf{C}')^{-1}(\mathbf{C}\overline{\mathbf{y}}_{2.}) = 128.552$. For the five rows of $\mathbf{C}$, we obtain $t_1^2 = 4.978, t_2^2 = 107.129, t_3^2 = 5.225, t_4^2 = 10.750, t_5^2 = 3.572$.

(c) For the combined groups ($\mathbf{S}_{\text{pl}}$ = pooled covariance matrix), we use $T^2 = N(\mathbf{C}\overline{\mathbf{y}}_{..})'(\mathbf{C}\mathbf{S}_{\text{pl}}\mathbf{C}')^{-1}(\mathbf{C}\overline{\mathbf{y}}_{..})$ in (6.121) to obtain $T^2 = 247.0079$. We test for linear, quadratic, etc., trends using the rows of $\mathbf{C}$ in $t_i^2 = N(\mathbf{c}_i'\overline{\mathbf{y}}_{..})^2/\mathbf{c}_i'\mathbf{S}\mathbf{c}_i$; $t_1^2 = 1.162, t_2^2 = 155.017, t_3^2 = 30.540, t_4^2 = 1.319, t_5^2 = .506$. To compare groups, we use $\Lambda = |\mathbf{CEC}'|/|\mathbf{C}(\mathbf{E}+\mathbf{H})\mathbf{C}'|$ in (6.123) and $\Lambda_i = \mathbf{c}_i'\mathbf{E}\mathbf{c}_i/\mathbf{c}_i'(\mathbf{E}+\mathbf{H})\mathbf{c}_i$: $\Lambda = .4902, \Lambda_1 = .7947, \Lambda_2 = .9940, \Lambda_3 = .7987, \Lambda_4 = .6228, \Lambda_5 = .9172$.

6.44 Control group: By (6.114), $\hat{\boldsymbol{\beta}}_1' = (3.129, .656, -.283, -.334, .192, .037, -0.20)$. By (6.115), $T^2 = .7633$. By (6.117), $\hat{\boldsymbol{\mu}}_1' = (\hat{\mu}_{11}, \hat{\mu}_{12}, \ldots, \hat{\mu}_{18}) = (4.11, 3.29, 2.71, 2.71, 3.04, 3.39, 3.54, 3.95)$.

Obese group: $\hat{\boldsymbol{\beta}}_2' = (3.207, -.187, .463, .056, -.102, -.010, .010), T^2 = .3943, \hat{\boldsymbol{\mu}}_2' = (4.51, 4.12, 3.81, 3.48, 3.24, 3.37, 3.70, 4.02)$

Combined groups (pooled covariance matrix): $\hat{\boldsymbol{\beta}}' = (3.15, .162, .183, -.115, .012, .010, -.002)$, $T^2 = .0158$, $\hat{\boldsymbol{\mu}}' = (4.36, 3.80, 3.36, 3.15, 3.13, 3.37, 3.63, 3.98)$

6.45 A = activator, T = time, C = group. In (6.100), (6.101), and (6.102), we use

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

$$\mathbf{T} = \begin{pmatrix} -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \\ 1 & -2 & 1 & 1 & -2 & 1 & 1 & -2 & 1 \end{pmatrix}$$

$$\mathbf{G} = \begin{pmatrix} -2 & 0 & 2 & 1 & 0 & -1 & 1 & 0 & -1 \\ 2 & -4 & 2 & -1 & 2 & -1 & -1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -2 & 1 & -1 & 2 & -1 \end{pmatrix}$$

$T_A^2 = 5072.579, T_T^2 = 268.185, T_{AT}^2 = 143.491$. The same within-sample factors and interaction can be tested with Wilks' Λ using (6.104) and the other three MANOVA tests:

Source	Λ	$V^{(s)}$	$U^{(s)}$	θ	Significant?
A	.003	.997	317.04	.997	Yes
T	.056	.944	16.76	.944	Yes
AT	.100	.900	8.97	.900	Yes

The interactions of the within factors with the between factor G are tested with Wilks' Λ (Section 6.9.5) and with the other three MANOVA tests:

Source	Λ	$V^{(s)}$	$U^{(s)}$	θ	Significant?
AC	.884	.116	.131	.116	No
TC	.889	.111	.125	.111	No
ATC	.795	.205	.258	.205	No

The between-subjects factor C is tested with an ANOVA F-test: $F = .47, p$-value $=$ .504.

CHAPTER 7

7.1 If $\mathbf{\Sigma}_0$ is substituted for $\mathbf{S}$ in (7.1), we have

$$u = \nu[\ln|\mathbf{\Sigma}_0| - \ln|\mathbf{\Sigma}_0| + \text{tr}(\mathbf{I}) - p] = \nu[0 + p - p] = 0$$

7.2
$$\begin{aligned}\ln|\mathbf{\Sigma}_0| - \ln|\mathbf{S}| &= -\ln|\mathbf{\Sigma}_0|^{-1} - \ln|\mathbf{S}| \\ &= -\ln|\mathbf{\Sigma}_0^{-1}| - \ln|\mathbf{S}| \quad \text{[by (2.91)]} \\ &= -(\ln|\mathbf{S}| + \ln|\mathbf{\Sigma}_0^{-1}|) \\ &= -\ln|\mathbf{S}\mathbf{\Sigma}_0^{-1}| \quad \text{[by (2.89)]}\end{aligned}$$

7.3
$$\begin{aligned}-\ln\left(\textstyle\prod_{i=1}^p \lambda_i\right) + \sum_{i=1}^p \lambda_i &= -\textstyle\sum_{i=1}^p \ln\lambda_i + \sum_{i=1}^p \lambda_i \\ &= \textstyle\sum_{i=1}^p (\lambda_i - \ln\lambda_i)\end{aligned}$$

7.4 As noted in Section 7.1, the likelihood ratio in this case involves the ratio of the determinants of the sample covariance matrices under H_0 and H_1. Under H_1, which is essentially unrestricted, the maximum likelihood estimate of $\mathbf{\Sigma}$ (corrected for bias) is given by (4.12) as $\mathbf{S}$. Under H_0 it is assumed that each of the p y_i's in $\mathbf{y}$ has variance σ^2 and that all y_i's are independent. Thus we estimate σ^2 (unbiasedly) in each of the p columns of the $\mathbf{Y}$ matrix [see (3.17) and (3.23)] and pool the p estimates to obtain

$$\hat{\sigma}^2 = \sum_{i=1}^n \sum_{j=1}^p \frac{(y_{ij} - \overline{y}_j)^2}{(n-1)p}.$$

Show that by (3.22) and (3.23) this is equal to

$$\hat{\sigma}^2 = \sum_{j=1}^p \frac{s_{jj}}{p} = \frac{\text{tr}(\mathbf{S})}{p}.$$

Thus the likelihood ratio is

$$\text{LR} = \left(\frac{|\mathbf{S}|}{|\hat{\sigma}^2\mathbf{I}|}\right)^{n/2} = \left(\frac{|\mathbf{S}|}{|\mathbf{I}\,\text{tr}(\mathbf{S})/p|}\right)^{n/2}.$$

Show that by (2.85) this becomes

$$\mathrm{LR} = \left(\frac{|\mathbf{S}|}{(\mathrm{tr}\mathbf{S}/p)^p} \right)^{n/2}.$$

7.5 If $\lambda_1 = \lambda_2 = \cdots = \lambda_p = \lambda$, say, then by (7.5),

$$u = \frac{p^p \prod_{i=1}^p \lambda_i}{(\sum_{i=1}^p \lambda_i)^p} = \frac{p^p \lambda^p}{(p\lambda)^p} = 1.$$

7.6 $$[(1-\rho)\mathbf{I} + \rho\mathbf{J}] = \begin{pmatrix} 1-\rho & 0 & \ldots & 0 \\ 0 & 1-\rho & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1-\rho \end{pmatrix} + \begin{pmatrix} \rho & \rho & \ldots & \rho \\ \rho & \rho & \ldots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \ldots & \rho \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \rho & \ldots & \rho \\ \rho & 1 & \ldots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \ldots & 1 \end{pmatrix}$$

7.7 **(a)** Substitute $\mathbf{J} = \mathbf{j}\mathbf{j}'$ and $\mathbf{x} = \mathbf{j}$ into $\mathbf{J}\mathbf{x} = \lambda\mathbf{x}$ to obtain $\mathbf{j}\mathbf{j}'\mathbf{j} = \lambda\mathbf{j}$ which gives $p\mathbf{j} = \lambda\mathbf{j}$.

(b) $\mathbf{S}_0 = s^2[(1-r)\mathbf{I} + r\mathbf{J}] = s^2(1-r)\left(\mathbf{I} + \frac{r}{1-r}\mathbf{J}\right)$

(c) By (2.85) and (2.108), we have

$$\begin{aligned} |\mathbf{S}_0| &= \left| s^2(1-r)\left(\mathbf{I} + \frac{r}{1-r}\mathbf{J}\right)\right| = (s^2)^p(1-r)^p \left|\mathbf{I} + \frac{r}{1-r}\mathbf{J}\right| \\ &= (s^2)^p(1-r)^p \textstyle\prod_{i=1}^p (1+\lambda_i) = (s^2)^p(1-r)^p \left(1 + \frac{rp}{1-r}\right) \\ &= (s^2)^p(1-r)^{p-1}(1-r+rp) = (s^2)^p(1-r)^{p-1}[1+(p-1)r] \end{aligned}$$

7.8 $$M = \frac{|\mathbf{S}_1|^{\nu_1/2}|\mathbf{S}_2|^{\nu_2/2}\cdots|\mathbf{S}_k|^{\nu_k/2}}{|\mathbf{S}|^{\sum_i \nu_i/2}} = \frac{|\mathbf{S}_1|^{\nu_1/2}|\mathbf{S}_2|^{\nu_2/2}\cdots|\mathbf{S}_k|^{\nu_k/2}}{|\mathbf{S}|^{\nu_1/2}|\mathbf{S}|^{\nu_2/2}\cdots|\mathbf{S}|^{\nu_k/2}}$$

7.9 **(a)** $M = .7015$ **(b)** $M = .0797$

7.10 $$\begin{aligned} \Lambda &= \frac{|\mathbf{S}|}{|\mathbf{S}_{yy}|\,|\mathbf{S}_{xx}|} = \frac{|\mathbf{S}_{xx}|\,|\mathbf{S}_{yy} - \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}|}{|\mathbf{S}_{yy}|\,|\mathbf{S}_{xx}|} \\ &= |\mathbf{S}_{yy}^{-1}|\,|\mathbf{S}_{yy} - \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}| \quad \text{[by (2.91)]} \\ &= |\mathbf{S}_{yy}^{-1}(\mathbf{S}_{yy} - \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy})| \quad \text{[by (2.89)]} \\ &= |\mathbf{I} - \mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy})| \end{aligned}$$

$= \prod_{i=1}^{s}(1 - r_i^2)$ [by (2.108)],

where the r_i^2's are the nonzero eigenvalues of $\mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}$. It was shown in Section 2.11.2 that $1-\lambda_i$ is an eigenvalue of $\mathbf{I}-\mathbf{A}$, where λ_i is an eigenvalue of $\mathbf{A}$.

7.11 When all $p_i = 1$, we have $k = p$, and the submatrices in the denominators of (7.33) and (7.34) reduce to $\mathbf{S}_{jj} = s_{jj}, j = 1, 2, \ldots, p$, and $\mathbf{R}_{jj} = 1, j = 1, 2, \ldots, p$.

7.12 When all $p_i = 1$, we have $k = p$ and

$$\begin{aligned}
a_2 &= p^2 - \textstyle\sum_{i=1}^{p} p_i^2 = p^2 - p, \quad a_3 = p^3 - p, \\
c &= 1 - \frac{1}{12f\nu}(2a_3 + 3a_2) \\
&= 1 - \frac{1}{6(p^2 - p)\nu}[2(p^3 - p) + 3(p^2 - p)] \\
&= 1 - \frac{1}{6(p - 1)\nu}[2(p^2 - 1) + 3(p - 1)] \\
&= 1 - \frac{1}{6(p - 1)\nu}[2(p - 1)(p + 1) + 3(p - 1)] \\
&= 1 - \frac{1}{6\nu}[2p + 5].
\end{aligned}$$

7.13 As noted below (7.6), the degrees of freedom for the χ^2-approximation is the total number of parameters minus the number estimated under H_0. The number of distinct parameters in $\mathbf{\Sigma}$ is $p + \binom{p}{2} = \frac{1}{2}p(p+1)$. The number of parameters estimated under H_0 is p. The difference is $\frac{1}{2}\,p(p+1) - p = \frac{1}{2}\,p(p-1)$.

7.14 By (7.1) and (7.2), $u = 11.094$ and $u' = 10.668$.

7.15 By (7.7), $u = .0000594$. By (7.9), $u' = 23.519$. For $H_0\colon \mathbf{C\Sigma C}' = \sigma^2\mathbf{I}, u = .471$ and $u' = 2.050$.

7.16 For $H_0\colon \mathbf{\Sigma} = \sigma^2\mathbf{I}, u = .00513$ and $u' = 131.922$. For $H_0\colon \mathbf{C\Sigma C}' = \sigma^2\mathbf{I}, u = .129$ and $u' = 36.278$.

7.17 For $H_0\colon \mathbf{\Sigma} = \sigma^2\mathbf{I}, u = .00471$ and $u' = 136.190$. For $H_0\colon \mathbf{C\Sigma C}' = \sigma^2\mathbf{I}, u = .747$ and $u' = 7.486$.

7.18 By (7.16), $u' = 6.3323$ with 13 degrees of freedom. The F-approximation is $F = .4802$ with 13 and 1147 degrees of freedom.

7.19 $u' = 21.488, F = 2.511$ with 8 and 217 degrees of freedom

7.20 $u' = 35.795, F = 4.466$ with 8 and 4905 degrees of freedom

7.21 $u = 8.7457, F = .8730$ with 10 and 6502 degrees of freedom

7.22 $|\mathbf{S}_1| = 2.620 \times 10^{14}, |\mathbf{S}_2| = 2.410 \times 10^{14}, |\mathbf{S}_{\text{pl}}| = 4.368 \times 10^{14}, u = 17.502, F = .829$

7.23 $\ln M = -85.965, u = 156.434, a_1 = 21, a_2 = 17{,}797, F = 7.4396$

7.24 $\ln M = -7.082, u = 10.565, a_1 = 10, a_2 = 1340, F = 1.046$

7.25 $\ln M = -8.6062, u = 14.222, a_1 = 20, a_2 = 3909, F = .707$

7.26 $\ln M = -28.917, u = 44.018, a_1 = 50, a_2 = 3238, F = .8625$

7.27 $\ln M = -142.435, u = 174.285, a_1 = 110, a_2 = 2084, F = 1.448$

7.28 $|\mathbf{S}| = 1{,}207, 109.5, |\mathbf{S}_{yy}| = 2385.1, |\mathbf{S}_{xx}| = 1341.9, \Lambda = .3772$

7.29 $|\mathbf{S}| = 4.237 \times 10^{13}, |\mathbf{S}_{yy}| = 484, 926.6, |\mathbf{S}_{xx}| = 131, 406, 938, \Lambda = .6650$

7.30 $|\mathbf{S}| = 9.676 \times 10^{-8}, |\mathbf{S}_{yy}| = .02097, |\mathbf{S}_{xx}| = 9.94 \times 10^{-6}, \Lambda = .4642$

7.31 $|\mathbf{S}| = 1.7148 \times 10^{16}, |\mathbf{S}_{11}| = 11, 284.967, |\mathbf{S}_{22}| = 11, 891.15, |\mathbf{S}_{33}| = 25, 951.605, s_{44} = 22, 227.158, s_{55} = 214.06, u = .00103, u' = 274.787, \nu = 46$

7.32 $|\mathbf{S}| = 459.96, s_{11} = 140.54, s_{22} = 72.25, s_{33} = .250, u = .1811, u' = 12.246, f = 3$

7.33 $u = .0001379, u' = 16.297$

7.34 $u = .0005176, u' = 127.367$

7.35 $u = .005071, u' = 131.226$

CHAPTER 8

8.1 Using $\mathbf{a} = \mathbf{S}_{\text{pl}}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)$, we obtain

$$
\begin{aligned}
\frac{[\mathbf{a}'(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)]^2}{\mathbf{a}'\mathbf{S}_{\text{pl}}\mathbf{a}} &= \frac{[(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)]^2}{(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1}\mathbf{S}_{\text{pl}}\mathbf{S}_{\text{pl}}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)} \\
&= \frac{[(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)]^2}{(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)}.
\end{aligned}
$$

8.2 You may wish to use the following steps:

(i) In Section 5.6.2 the grouping variable w is defined as $n_2/(n_1 + n_2)$ for each observation in group 1 and $-n_1/(n_1 + n_2)$ for group 2. Show that with this formulation, $\overline{w} = 0$.

(ii) Because $\overline{w} = 0$, there is no intercept and the fitted model becomes

$$\hat{w}_i = b_1(y_{i1} - \overline{y}_1) + b_2(y_{i2} - \overline{y}_2) + \cdots + b_p(y_{ip} - \overline{y}_p),$$
$$i = 1, 2, \ldots, n_1 + n_2.$$

Denote the resulting matrix of y-values corrected for their means as $\mathbf{Y}_c$ and the vector of w's as $\mathbf{w}$. Then the least squares estimate $\mathbf{b} = (b_1, b_2, \ldots, b_p)'$ is obtained as

$$\mathbf{b} = (\mathbf{Y}_c'\mathbf{Y}_c)^{-1}\mathbf{Y}_c'\mathbf{w}.$$

Using (2.51), show that

$$\begin{aligned}\mathbf{Y}_c'\mathbf{Y}_c &= \sum_{i=1}^{2}\sum_{j=1}^{n_i}(\mathbf{y}_{ij} - \overline{\mathbf{y}})(\mathbf{y}_{ij} - \overline{\mathbf{y}})' \\ &= \sum_{i=1}^{2}\sum_{j=1}^{n_i}(\mathbf{y}_{ij} - \overline{\mathbf{y}}_i)(\mathbf{y}_{ij} - \overline{\mathbf{y}}_i)' + \frac{n_1 n_2}{n_1 + n_2}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)',\end{aligned}$$

where $\overline{\mathbf{y}} = (n_1\overline{\mathbf{y}}_1 + n_2\overline{\mathbf{y}}_2)/(n_1 + n_2)$. It will be helpful to write the first sum above as

$$\sum_{j=1}^{n_1}(\mathbf{y}_{1j} - \overline{\mathbf{y}})(\mathbf{y}_{1j} - \overline{\mathbf{y}})' + \sum_{j=1}^{n_2}(\mathbf{y}_{2j} - \overline{\mathbf{y}})(\mathbf{y}_{2j} - \overline{\mathbf{y}})'$$

and add and subtract $\overline{\mathbf{y}}_1$ in the first term and $\overline{\mathbf{y}}_2$ in the second.

(iii) Show that

$$\mathbf{Y}_c'\mathbf{w} = \sum_{i=1}^{2}\sum_{j=1}^{n_i}(\mathbf{y}_{ij} - \overline{\mathbf{y}})w_{ij} = \frac{n_1 n_2}{n_2 + n_2}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2).$$

Again it will be helpful to sum separately over the two groups.

(iv) From (ii) and (iii) we have

$$\mathbf{b} = (\nu\mathbf{S} + k\overline{\mathbf{d}}\,\overline{\mathbf{d}}')^{-1}k\overline{\mathbf{d}},$$

where $\mathbf{S} = \sum_{ij}(\mathbf{y}_{ij} - \overline{\mathbf{y}}_i)(\mathbf{y}_{ij} - \overline{\mathbf{y}}_i)'/(n_1 + n_2 - 2)$, $\nu = n_1 + n_2 - 2$, $k = n_1 n_2/(n_1 + n_2)$, and $\overline{\mathbf{d}} = \overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2$. Use (2.77) for the inverse of a patterned matrix of the type $\nu\mathbf{S} + k\overline{\mathbf{d}}\,\overline{\mathbf{d}}'$ to obtain (8.4).

8.3 You may want to use the following steps:

(i) R^2 is defined as [see (10.30)]

$$R^2 = \frac{\mathbf{b}'\mathbf{Y}_c'\mathbf{w} - n\overline{w}^2}{\mathbf{w}'\mathbf{w} - n\overline{w}^2}.$$

In this case the expression simplifies because $\overline{w} = 0$. Using $\mathbf{Y}_c'\mathbf{w}$ in Problem 8.2(iii) above, show that $R^2 = \mathbf{b}'(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)$.

(ii) Show that

$$\mathbf{b}'(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2) = \frac{T^2}{n_1 + n_2 - 2 + T^2}.$$

8.4 $[\mathbf{a}'(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)]^2 = \mathbf{a}'(\overline{\mathbf{y}} - \overline{\mathbf{y}}_2)\mathbf{a}'(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2) = \mathbf{a}'(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{a}$

8.5 $\mathbf{Ha} - \lambda\mathbf{Ea} = \mathbf{0}$
$\mathbf{E}^{-1}(\mathbf{Ha} - \lambda\mathbf{Ea}) = \mathbf{E}^{-1}\mathbf{0}$
$\mathbf{E}^{-1}\mathbf{Ha} - \lambda\mathbf{E}^{-1}\mathbf{Ea} = \mathbf{0}$
$(\mathbf{E}^{-1}\mathbf{H} - \lambda\mathbf{I})\mathbf{a} = \mathbf{0}$

8.6 Substituting $a_r^* = s_r a_r, r = 1, 2, \ldots, p$, into (8.15), we obtain

$$\begin{aligned} z_{1i} &= s_1 a_1 \frac{y_{1i1} - \overline{y}_{11}}{s_1} + s_2 a_2 \frac{y_{1i2} - \overline{y}_{12}}{s_2} + \cdots + s_p a_p \frac{y_{1ip} - \overline{y}_{1p}}{s_p} \\ &= a_1 y_{1i1} + a_2 y_{1i2} + \cdots + a_p y_{1ip} - a_1 \overline{y}_{11} - a_2 \overline{y}_{12} - \cdots - a_p \overline{y}_{1p} \\ &= a_1 y_{1i1} + a_2 y_{1i2} + \cdots + a_p y_{1ip} - \mathbf{a}'\overline{\mathbf{y}}_1 \end{aligned}$$

8.7 **(a)** $\mathbf{a}^{*\prime} = (1.366, -.810, 2.525, -1.463)$

(b) $t_1 = 5.417, t_2 = 2.007, t_3 = 7.775, t_4 = .688$

(c) The standardized coefficients rank the variables in the order y_3, y_4, y_1, y_2. The t-tests rank them in the order y_3, y_1, y_2, y_4.

(d) The partial F's calculated by (8.26) are $F(y_1|y_2, y_3, y_4) = 7.844$, $F(y_2|y_1, y_3, y_4) = 2.612$, $F(y_3|y_1, y_2, y_4) = 40.513$, and $F(y_4|y_1, y_2, y_3) = 9.938$.

8.8 **(a)** $\mathbf{a}' = (.345, -.130, -.106, -.143)$

(b) $\mathbf{a}^{*\prime} = (4.137, -2.501, -1.158, -2.068)$

(c) $t_1 = 3.888, t_2 = -3.865, t_3 = -5.691, t_4 = -5.043$

(e) $F(y_1|y_2, y_3, y_4) = 35.934, F(y_2|y_1, y_3, y_4) = 5.799$
$F(y_3|y_1, y_2, y_4) = 1.775, F(y_4|y_1, y_2, y_3) = 8.259$

8.9 **(a)** $\mathbf{a}' = (-.145, .052, -.005, -.089, -.007, -.022)$

(b) $\mathbf{a}^{*\prime} = (-1.016, .147, -.542, -1.035, -.107, -1.200)$

(c) $t_1 = -4.655, t_2 = .592, t_3 = -4.354, t_4 = -5.257, t_5 = -4.032, t_6 = -6.439$

(e) $F(y_1|y_2, y_3, y_4, y_5, y_6) = 8.081, F(y_2|y_1, y_3, y_4, y_5, y_6) = .150$, $F(y_3|y_1, y_2, y_4, y_5, y_6) = .835, F(y_4|y_1, y_2, y_3, y_5, y_6) = 8.503$, $F(y_5|y_1, y_2, y_3, y_4, y_6) = .028, F(y_6|y_1, y_2, y_3, y_4, y_5) = 9.192$

8.10 **(a)** $\mathbf{a}' = (.057, .010, .242, .071)$

(b) $\mathbf{a}^{*\prime} = (1.390, .083, 1.025, .032)$

(c) $t_1 = -3.713, t_2 = .549, t_3 = -3.262, t_4 = -.724$

(e) $F(y_1|y_2, y_3, y_4) = 3.332, F(y_2|y_1, y_3, y_4) = .010$
$F(y_3|y_1, y_2, y_4) = 1.482, F(y_4|y_1, y_2, y_3) = .001$

8.11 **(a)** $\mathbf{a}_1' = (.021, .533, -.347, -.135), \mathbf{a}_2' = (-.317, .298, .243, -.026)$

(b) $\lambda_1/(\lambda_1 + \lambda_2) = .958, \lambda_2/(\lambda_1 + \lambda_2) = .042$. Using the methods of Section 8.6.2, we have two tests, the first for significance of λ_1 and λ_2 and the second for significance of λ_2:

Test	Λ	F	p-Value for F
1	.2245	8.3294	<.0001
2	.8871	1.3157	.2869

(c) $\mathbf{a}_1^{*\prime} = (.076, 1.553, -1.182, -.439), \mathbf{a}_2^{*\prime} = (-1.162, .869, .828, -.085)$

(d) $F(y_1|y_2, y_3, y_4) = 1.067, F(y_2|y_1, y_3, y_4) = 20.975$
$F(y_3|y_1, y_2, y_4) = 9.630, F(y_4|y_1, y_2, y_3) = 1.228$

(e) In the plot, the first discriminant function separates groups 1 and 2 from group 3, but the second is ineffective in separating group 1 from group 2.

8.12 **(a)**

λ_i	$\lambda_i/\sum_{j=1}^{4}\lambda_j$	Eigenvector
1.8757	.6421	$\mathbf{a}_1' = (.470, -.263, .653, -.074)$
.7907	.2707	$\mathbf{a}_2' = (.176, .188, -1.058, 1.778)$
.2290	.0784	$\mathbf{a}_3' = (-.155, .258, .470, -.850)$
.0260	.0089	$\mathbf{a}_4' = (-3.614, .475, .310, -.479)$

(b) Test of significance of each eigenvalue and those that follow it:

Test	Λ	Approximate F	p-Value for F
1	.1540	4.937	<.0001
2	.4429	3.188	.0006
3	.7931	1.680	.1363
4	.9747	.545	.5839

(c) $\mathbf{a}_1^{*\prime} = (.266, -.915, 1.353, -.097), \mathbf{a}_2^{*\prime} = (.100, .654, -2.291, 2.333)$,
$\mathbf{a}_3^{*\prime} = (-.087, .899, .973, -1.115), \mathbf{a}_4^{*\prime} = (-2.044, 1.654, .643, -.628)$

(d) $F(y_1|y_2, y_3, y_4) = .299, F(y_2|y_1, y_3, y_4) = 1.931$
$F(y_3|y_1, y_2, y_4) = 6.085, F(y_4|y_1, y_2, y_3) = 4.659$

(e) In the plot, the first discriminant function separates groups 1, 4, and 6 from groups 2, 3, and 5. The second function achieves some separation of group 6 from groups 1 and 4 and some separation of group 3 from groups 2 and 5.

8.13 Three variables entered the model in the stepwise selection. The summary table is as follows:

Step	Variable Entered	Overall Λ	p-Value	Partial Λ	Partial F	p-Value
1	y_4	.4086	<.0001	.4086	12.158	<.0001
2	y_3	.2655	<.0001	.6499	4.418	.0026
3	y_2	.1599	<.0001	6022	5.284	.0008

8.14 Summary table:

Step	Variable Entered	Overall Λ	p-Value	Partial Λ	Partial F	p-Value
1	y_4	.6392	<.0001	.6392	21.451	<.0001
2	y_3	.5430	<.0001	.8495	6.554	.0147
3	y_6	.4594	<.0001	.8461	6.549	.0148
4	y_2	.4063	<.0001	.8843	4.578	.0394
5	y_5	.3639	<.0001	.8957	3.959	.0547

In this case, the fifth variable to enter, y_5, would not ordinarily be included in the subset. The p-value of .0547 is large in this setting, where several tests are run at each step and the variable with smallest p-value is selected.

8.15 Summary table:

Step	Variable Entered	Overall Λ	p-Value	Partial Λ	Partial F	p-Value
1	y_2	.6347	.0006	.6347	9.495	.0006
2	y_3	.2606	<.0001	.4106	22.975	<.0001

CHAPTER 9

9.1 $\bar{z}_1 - \bar{z}_2 = \mathbf{a}'\bar{\mathbf{y}}_1 - \mathbf{a}'\bar{\mathbf{y}}_2 = \mathbf{a}'(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2) = (\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)$

9.2 $\frac{1}{2}(\bar{z}_1 + \bar{z}_2) = \frac{1}{2}(\mathbf{a}'\bar{\mathbf{y}}_1 + \mathbf{a}'\bar{\mathbf{y}}_2) = \frac{1}{2}\mathbf{a}'(\bar{\mathbf{y}}_1 + \bar{\mathbf{y}}_2)$
$= \frac{1}{2}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1}(\bar{\mathbf{y}}_1 + \bar{\mathbf{y}}_2)$

9.3 Write (9.8) in the form

$$\frac{f(\mathbf{y}|G_1)}{f(\mathbf{y}|G_2)} > \frac{p_2}{p_1}$$

and substitute $f(\mathbf{y}|G_i) = N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$ from (4.2) to obtain

$$\frac{f(\mathbf{y}|G_1)}{f(\mathbf{y}|G_2)} = e^{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)'\boldsymbol{\Sigma}^{-1}\mathbf{y}-(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1+\boldsymbol{\mu}_2)/2} > \frac{p_2}{p_1}.$$

Substitute estimates for $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, and $\boldsymbol{\Sigma}$, and take the logarithm of both sides to obtain (9.9). Note that if $a > b$, then $\ln a > \ln b$.

9.4 Maximizing $p_i f(\mathbf{y}, G_i)$ is equivalent to maximizing $\ln[p_i f(\mathbf{y}|G_i)]$. Use $f(\mathbf{y}|G_i) = N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$ from (4.2) and take the logarithm to obtain

$$\ln[p_i f(\mathbf{y}|G_i)] = \ln p_i - \tfrac{1}{2}p\ln(2\pi) - \tfrac{1}{2}|\boldsymbol{\Sigma}| - \tfrac{1}{2}(\mathbf{y} - \boldsymbol{\mu}_i)'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu}_i).$$

Expand the last term, delete terms common to all groups (terms that do not involve i), and substitute estimators of $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}$ to obtain (9.13).

9.5 Use $f(\mathbf{y}|G_i) = N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ in $\ln[p_i f(\mathbf{y}|G_i)]$, delete $-(p/2)\ln(2\pi)$, and substitute $\overline{\mathbf{y}}_i$ and $\mathbf{S}_i$ for $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$.

9.6 Substitute the constant c for each occurrence of p_i/n_i to obtain the result.

9.7 **(a)** $\mathbf{a}' = (\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1} = (.345, -.130, -.106, -.143)$, $\frac{1}{2}(\overline{z}_1 + \overline{z}_2) = -15.8054$

(b)

		Predicted Group	
Actual Group	Number of Observations	1	2
1	19	19	0
2	20	1	19

Error rate $= \frac{1}{39} = .0256$

(c) Using the k nearest neighbor method with $k = 5$, we obtain the same classification table as in part (b). With $k = 4$, two observations are misclassified, and the error rate becomes $2/39 = .0513$.

9.8 **(a)** $\mathbf{a}' = (\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1} = (-.145, .052, -.005, -.089, -.007, -.022)$, $\frac{1}{2}(\overline{z}_1 + \overline{z}_2) = -17.045$

(b)

		Predicted Group	
Actual Group	Number of Observations	1	2
1	39	37	20
2	34	8	26

Error rate $= (2 + 8)/73 = .1370$

(c) p_1 and p_2 Proportional to Sample Sizes

		Predicted Group	
Actual Group	Number of Observations	1	2
1	39	37	2
2	34	8	26

Error rate $= (2 + 8)/73 = .1370$

9.9 **(a)** $\mathbf{a}' = (\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{S}_{\text{pl}}^{-1} = (-.057, -.010, -.242, -.071)$, $\frac{1}{2}(\overline{z}_1 + \overline{z}_2) = -7.9686$

(b) Linear Classification

Actual Group	Number of Observations	Predicted Group 1	Predicted Group 2
1	9	8	1
2	10	1	9

Error rate $= \frac{2}{19} = .1053$

(c) Holdout Method

Actual Group	Number of Observations	Predicted Group 1	Predicted Group 2
1	9	6	3
2	10	3	7

Error rate $= (3+3)/19 = .3158$

(d) Kernel Density Estimator with $h = 2$

Actual Group	Number of Observations	Predicted Group 1	Predicted Group 2
1	9	9	0
2	10	1	9

Error rate $= \frac{1}{19} = .0526$

9.10 **(a)**

Actual Group	Number of Observations	Predicted Group 1	Predicted Group 2
1	20	18	2
2	20	2	18

Error rate $= (2+2)/40 = .100$.

(b) Four variables were selected by the stepwise discriminant analysis: y_2, y_3, y_4, and y_6 (see Problem 8.14). With these four variables we obtain the classification table in part (c).

(c)

Actual Group	Number of Observations	Predicted Group 1	Predicted Group 2
1	20	18	2
2	20	2	18

Error rate $= (2+2)/40 = .100$. The four variables classified the sample as well as did all six variables in part (a).

9.11 **(a)** By (9.11), $L_i(\mathbf{y}) = \overline{\mathbf{y}}_i'\mathbf{S}_{\text{pl}}^{-1}\mathbf{y} - \frac{1}{2}\overline{\mathbf{y}}_i'\mathbf{S}_{\text{pl}}^{-1}\overline{\mathbf{y}}_i = \mathbf{c}_i'\mathbf{y} + c_{0i}$. The vectors $\begin{pmatrix} c_{0i} \\ \mathbf{c}_i \end{pmatrix}$, $i = 1, 2, 3$, are

Group 1	Group 2	Group 3
−72.77	−65.18	−68.57
.81	2.12	.68
15.15	10.11	2.79
−1.03	−.24	6.54
10.02	11.06	13.09

(b) Linear Classification

Actual Group	Number of Observations	Predicted Group 1	Predicted Group 2	Predicted Group 3
1	12	9	3	0
2	12	3	7	2
3	12	0	1	11

Error rate $= (3 + 3 + 2 + 1)/36 = .250$

(c) Quadratic Classification

Actual Group	Number of Observations	Predicted Group 1	Predicted Group 2	Predicted Group 3
1	12	10	2	0
2	12	2	8	2
3	12	0	1	11

Error rate $= (2 + 2 + 2 + 1)/36 = .194$

(d) Linear Classification–Holdout Method

Actual Group	Number of Observations	Predicted Group 1	Predicted Group 2	Predicted Group 3
1	12	7	5	0
2	12	4	5	3
3	12	0	1	11

Error rate $= (5 + 4 + 3 + 1)/12 = .361$

(e) k Nearest Neighbor with $k = 5$

Actual Group	Number of Observations	Predicted Group 1	2	3
1	11	9	2	0
2	11	2	7	2
3	12	0	1	11

Error rate $= (2 + 2 + 2 + 1)/34 = .206$

9.12 **(a)** By (9.11), $L_i(\mathbf{y}) = \overline{\mathbf{y}}_i'\mathbf{S}_{\text{pl}}^{-1}\mathbf{y} - \frac{1}{2}\overline{\mathbf{y}}_i'\mathbf{S}_{\text{pl}}^{-1}\overline{\mathbf{y}}_i = \mathbf{c}_i'\mathbf{y} + c_{0i}$. The vectors $\begin{pmatrix} c_{0i} \\ \mathbf{c}_i \end{pmatrix}$, $i = 1, 2, \ldots, 6$, are

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
−300.0	−353.2	−328.5	−291.8	−347.5	−315.8
314.6	317.1	324.6	307.3	316.8	311.3
−59.4	−64.0	−65.2	−59.4	−65.8	−63.1
149.6	168.2	154.9	147.7	168.2	160.6
−161.2	−172.6	−150.4	−153.4	−172.9	−175.5

(b) Linear Classification

Actual Group	Number of Observations	Predicted Group 1	2	3	4	5	6
1	8	5	0	0	1	0	2
2	8	0	3	2	1	2	0
3	8	0	0	6	1	1	0
4	8	3	0	1	4	0	0
5	8	0	3	1	0	3	1
6	8	2	0	0	0	2	4

Correct classification rate $= (5 + 3 + 6 + 4 + 3 + 4)/48 = .521$

Error rate $= 1 - .521 = .479$

(c) Quadratic Classification

Actual Group	Number of Observations	Predicted Group 1	2	3	4	5	6
1	8	8	0	0	0	0	0
2	8	0	7	0	1	0	0
3	8	1	0	6	0	1	0
4	8	0	0	1	7	0	0
5	8	0	3	0	0	4	1
6	8	2	0	0	0	1	5

Correct classification rate $= (8 + 7 + 6 + 7 + 4 + 5)/48 = .771$

Error rate $= 1 - .771 = .229$

(d) k Nearest Neighbor with $k = 3$

Actual Group	Number of Observations	Predicted Group 1	2	3	4	5	6	Ties
1	8	5	0	0	2	0	0	1
2	8	0	4	0	0	1	0	3
3	8	1	0	6	0	1	0	0
4	8	0	0	0	5	0	0	3
5	8	0	1	0	0	6	1	0
6	8	2	0	0	0	0	5	1

Correct classification rate $= (5 + 4 + 6 + 5 + 6 + 5)/40 = .775$
Error rate $= 1 - .775 = .225$

(e) Normal Kernel with $h = 1$
(For this data set, larger values of h do much worse.)

Actual Group	Number of Observations	Predicted Group 1	2	3	4	5	6
1	8	8	0	0	0	0	0
2	8	0	8	0	0	0	0
3	8	1	0	6	0	1	0
4	8	1	0	0	7	0	0
5	8	0	0	0	0	7	1
6	8	2	0	0	0	0	6

Correct classification rate $= (8 + 8 + 6 + 7 + 7 + 6)/48 = .875$
Error rate $= 1 - .875 = .125$

CHAPTER 10

10.1 $$\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_n' \end{pmatrix} \hat{\boldsymbol{\beta}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} \mathbf{x}_1'\hat{\boldsymbol{\beta}} \\ \mathbf{x}_2'\hat{\boldsymbol{\beta}} \\ \vdots \\ \mathbf{x}_n'\hat{\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} y_1 - \mathbf{x}_1'\hat{\boldsymbol{\beta}} \\ y_2 - \mathbf{x}_2'\hat{\boldsymbol{\beta}} \\ \vdots \\ y_n - \mathbf{x}_n'\hat{\boldsymbol{\beta}} \end{pmatrix}$$

By (2.33), $\sum_{i=1}^n (y_i - \mathbf{x}_i'\hat{\boldsymbol{\beta}})^2 = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$.

10.2 $$\begin{aligned} \sum_{i=1}^n (y_i - \mu)^2 &= \sum_i (y_i - \overline{y} + \overline{y} - \mu)^2 \\ &= \sum_i (y_i - \overline{y})^2 + 2\sum_i (y_i - \overline{y})(\overline{y} - \mu) + \sum_i (\overline{y} - \mu)^2 \\ &= \sum_i (y_i - \overline{y})^2 + (\overline{y} - \mu)\sum_i (y_i - \overline{y}) + n(\overline{y} - \mu)^2 \\ &= \sum_i (y_i - \overline{y})^2 + n(\overline{y} - \mu)^2 \quad [\text{since } \textstyle\sum_i (y_i - \overline{y}) = 0] \end{aligned}$$

10.3 $\sum_{i=1}^n (x_{i2} - \overline{x}_2)\overline{y} = \overline{y}\sum_i (x_{i2} - \overline{x}_2) = \overline{y}(\sum_i x_{i2} - n\overline{x}_2) = \overline{y}(n\overline{x}_2 - n\overline{x}_2)$

10.4 $E[\hat{y}_i - E(y_i)]^2 = E[\hat{y}_i - E(\hat{y}_i) + E(\hat{y}_i) - E(y_i)]^2$
$= E[\hat{y}_i - E(\hat{y}_i)]^2 + 2E[\hat{y}_i - E(\hat{y}_i)][E(\hat{y}_i) - E(y_i)]$
$+ E[E(\hat{y}_i) - E(y_i)]^2$

The second term on the right vanishes because $[E(\hat{y}_i) - E(y_i)]$ is constant and $E[\hat{y}_i - E(\hat{y}_i)] = E(\hat{y}_i) - E(\hat{y}_i) = 0$. For the third term, we have $E[E(\hat{y}_i) - E(y_i)]^2 = [E(\hat{y}_i) - E(y_i)]^2$, because $[E(\hat{y}_i) - E(y_i)]^2$ is constant.

10.5 First show that $\text{cov}(\hat{\boldsymbol{\beta}}_p) = \sigma^2(\mathbf{X}_p'\mathbf{X}_p)^{-1}$. This can be done by noting that $\hat{\boldsymbol{\beta}}_p = (\mathbf{X}_p'\mathbf{X}_p)^{-1}\mathbf{X}_p'\mathbf{y} = \mathbf{A}\mathbf{y}$, say. Then, by (3.74), $\text{cov}(\mathbf{A}\mathbf{y}) = \mathbf{A}\text{cov}(\mathbf{y})\mathbf{A}' = \mathbf{A}(\sigma^2\mathbf{I})\mathbf{A}' = \sigma^2\mathbf{A}\mathbf{A}'$. By substituting $\mathbf{A} = (\mathbf{X}_p'\mathbf{X}_p)^{-1}\mathbf{X}_p'$, this becomes $\text{cov}(\hat{\boldsymbol{\beta}}_p) = \sigma^2(\mathbf{X}_p'\mathbf{X}_p)^{-1}$. Then by (3.70), $\text{var}(\mathbf{x}_{pi}'\hat{\boldsymbol{\beta}}_p) = \mathbf{x}_{pi}'\text{cov}(\hat{\boldsymbol{\beta}}_p)\mathbf{x}_{pi}$ and the remaining steps follow as indicated.

10.6 By (10.36), $s_p^2 = \text{SSE}_p/(n-p)$. Then by (10.44),

$$
\begin{aligned}
C_p &= p + (n-p)\frac{s_p^2 - s_k^2}{s_k^2} = p + (n-p)\left(\frac{s_p^2}{s_k^2} - 1\right) \\
&= p + (n-p)\frac{s_p^2}{s_k^2} - (n-p) = (n-p)\frac{\text{SSE}_p/s_k^2}{n-p} - n + 2p \\
&= \frac{\text{SSE}_p}{s_k^2} - (n-2p).
\end{aligned}
$$

10.7 $(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}) = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\hat{\mathbf{B}} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{X}\hat{\mathbf{B}}$. Transpose $\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ from (10.48) and substitute into $\hat{\mathbf{B}}'\mathbf{X}'\mathbf{X}\hat{\mathbf{B}}$.

10.8 $E[\hat{\mathbf{y}}_i - E(\mathbf{y}_i)][\hat{\mathbf{y}}_i - E(\mathbf{y}_i)]' = E[\hat{\mathbf{y}}_i - E(\hat{\mathbf{y}}_i) + E(\hat{\mathbf{y}}_i) - E(\mathbf{y}_i)][\hat{\mathbf{y}}_i - E(\hat{\mathbf{y}}_i)$
$+ E(\hat{\mathbf{y}}_i) - E(\mathbf{y}_i)]'$
$= E[\hat{\mathbf{y}}_i - E(\hat{\mathbf{y}}_i)][\hat{\mathbf{y}}_i - E(\hat{\mathbf{y}}_i)]'$
$+ E[\hat{\mathbf{y}}_i - E(\hat{\mathbf{y}}_i)][E(\hat{\mathbf{y}}_i) - E(\mathbf{y}_i)]'$
$+ E[E(\hat{\mathbf{y}}_i) - E(\mathbf{y}_i)][\hat{\mathbf{y}}_i - E(\hat{\mathbf{y}}_i)]'$
$+ E[E(\hat{\mathbf{y}}_i) - E(\mathbf{y}_i)][E(\hat{\mathbf{y}}_i) - E(\mathbf{y}_i)]'$

The second and third terms are equal to $\mathbf{O}$ because $[E(\hat{\mathbf{y}}_i) - E(\mathbf{y}_i)]$ is a constant vector and $E[\hat{\mathbf{y}}_i - E(\hat{\mathbf{y}}_i)] = E(\hat{\mathbf{y}}_i) - E(\hat{\mathbf{y}}_i) = \mathbf{0}$. The fourth term is a constant matrix and the first E can be deleted.

10.9 As in Problem 10.5, we have $\text{cov}(\hat{\boldsymbol{\beta}}_{p(j)}) = \sigma_{jj}(\mathbf{X}_p'\mathbf{X}_p)^{-1}$, where $\sigma_{jj} = \text{var}(y_j)$ is the jth diagonal element of $\boldsymbol{\Sigma} = \text{cov}(\mathbf{y})$. Similarly, $\text{cov}(\hat{\boldsymbol{\beta}}_{p(j)}, \hat{\boldsymbol{\beta}}_{p(k)}) = \sigma_{jk}(\mathbf{X}_p'\mathbf{X}_p)^{-1}$, where $\sigma_{jk} = \text{cov}(y_j, y_k)$ is the (jk)th element of $\boldsymbol{\Sigma}$. The notation $\text{cov}(\hat{\boldsymbol{\beta}}_{p(j)}, \hat{\boldsymbol{\beta}}_{p(k)})$ indicates a matrix containing the covariance of each element of $\hat{\boldsymbol{\beta}}_{p(j)}$ and each element of $\hat{\boldsymbol{\beta}}_{p(k)}$. Now for the covariance matrix, $\text{cov}(\hat{\mathbf{y}}_i') = \text{cov}(\mathbf{x}_{pi}'\hat{\boldsymbol{\beta}}_{p(1)}, \ldots, \mathbf{x}_{pi}'\hat{\boldsymbol{\beta}}_{p(m)})$, we need the variance of each of the m random variables and the covariance of each pair. By Problem 10.5

and (3.70), $\text{var}(\mathbf{x}'_{pi}\hat{\boldsymbol{\beta}}_{p(1)}) = \mathbf{x}'_{pi}\text{cov}(\hat{\boldsymbol{\beta}}_{p(1)})\mathbf{x}_{pi} = \sigma_{11}\mathbf{x}'_{pi}(\mathbf{X}'_p\mathbf{X}_p)^{-1}\mathbf{x}_{pi}$. Similarly, $\text{cov}(\mathbf{x}'_{pi}\hat{\boldsymbol{\beta}}_{p(1)}, \mathbf{x}'_{pi}\hat{\boldsymbol{\beta}}_{p(2)}) = \sigma_{12}\mathbf{x}'_{pi}(\mathbf{X}'_p\mathbf{X}_p)^{-1}\mathbf{x}_{pi}$. The other variances and covariances can be obtained in an analogous manner.

10.10 By (10.92), $\mathbf{S}_p = \mathbf{E}_p/(n-p)$. Then by (10.98),

$$\begin{aligned}\mathbf{C}_p &= p\mathbf{I} + (n-p)\mathbf{S}_k^{-1}(\mathbf{S}_p - \mathbf{S}_k) \\ &= p\mathbf{I} + (n-p)\mathbf{S}_k^{-1}\frac{\mathbf{E}_p}{n-p} - (n-p)\mathbf{I} \\ &= \mathbf{S}_k^{-1}\mathbf{E}_p + (2p-n)\mathbf{I}.\end{aligned}$$

10.11 $|\mathbf{E}_k^{-1}\mathbf{E}_p| = |\mathbf{E}_k^{-1}||\mathbf{E}_p| > 0$, because both $\mathbf{E}_k^{-1}$ and $\mathbf{E}_p$ are positive definite.

10.12 By (10.99), $\mathbf{C}_p = \mathbf{S}_k^{-1}\mathbf{E}_p + (2p-n)\mathbf{I}$. Using $\mathbf{S}_k = \mathbf{E}_k/(n-k)$, we obtain

$$\begin{aligned}\left(\frac{\mathbf{E}_k}{n-k}\right)^{-1}\mathbf{E}_p &= \mathbf{C}_p - (2p-n)\mathbf{I} \\ (n-k)\mathbf{E}_k^{-1}\mathbf{E}_p &= \mathbf{C}_p + (n-2p)\mathbf{I}.\end{aligned}$$

10.13 If $\mathbf{C}_p$ is replaced by $p\mathbf{I}$ in (10.101), we obtain

$$\mathbf{E}_k^{-1}\mathbf{E}_p = \frac{\mathbf{C}_p + (n-2p)\mathbf{I}}{n-k} = \frac{p\mathbf{I} + n\mathbf{I} - 2p\mathbf{I}}{n-k} = \frac{(n-p)\mathbf{I}}{n-k}.$$

10.14 **(a)** For this scenario, the Zyskind condition holds if $(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{X}^{\text{SUR}} = \mathbf{X}^{\text{SUR}}\mathbf{Q}$. Use property (2.121) to find the matrix $\mathbf{Q}$.

(b) For this scenario, the Zyskind condition holds if $(\boldsymbol{\Sigma}\otimes\mathbf{I})\mathbf{X}^{\text{SUR}} = \mathbf{X}^{\text{SUR}}\mathbf{Q}$. Use the diagonal property of $\boldsymbol{\Sigma}$ to find the matrix $\mathbf{Q}$.

10.15 **(a)** $\hat{\mathbf{B}} = \begin{pmatrix} .6264 & 83.243 \\ .0009 & .029 \\ -.0010 & -.013 \\ .0015 & -.004 \end{pmatrix}$

(b) $\Lambda = .724, V^{(s)} = .280, U^{(s)} = .375, \theta = .264$

(c) $\lambda_1 = .3594, \lambda_2 = .0160$. The essential rank of $\hat{\mathbf{B}}_1$ is 1, and the power ranking is $\theta > U^{(s)} > \Lambda > V^{(s)}$.

(d) The Wilks' Λ test of x_2 adjusted for x_1 and x_3, for example, is given by (10.74) as

$$\Lambda(x_2|x_1, x_3) = \frac{\Lambda(x_1, x_2, x_3)}{\Lambda(x_1, x_3)},$$

which is distributed as $\Lambda_{p,1,n-4}$ and has an exact F-transformation. The tests for x_1 and x_3 are similar. For the three tests we obtain the following:

	Λ	F	p-Value
$x_1 \mid x_2, x_3$	.931	1.519	.231
$x_2 \mid x_1, x_3$	.887	2.606	.086
$x_3 \mid x_1, x_2$	.762	6.417	.004

(e) Confidence intervals for the two elements of $E(\mathbf{y}_0)$ are: (0.86, 0.94) and (87.20, 93.39).

(f) Prediction intervals for the two elements of $\mathbf{y}_0$ are: (0.67, 1.13) and (72.63, 107.98).

10.16 **(a)** $\hat{\mathbf{B}} = \begin{pmatrix} 34.282 & 35.802 \\ .394 & .245 \\ .529 & .471 \end{pmatrix}$

(b) $\Lambda = .377, V^{(s)} = .625, U^{(s)} = 1.647, \theta = .622$

(c) $\lambda_1 = 1.644, \lambda_2 = .0029$. The essential rank of $\hat{\mathbf{B}}_1$ is 1, and the power ranking is $\theta > U^{(s)} > \Lambda > V^{(s)}$.

(d)

	Λ	F	p-Value
$x_1 \mid x_2$	.888	1.327	.287
$x_2 \mid x_1$	.875	1.506	.245

(e) Confidence intervals for the two elements of $E(\mathbf{y}_0)$ are: (170.31, 180.44) and (139.51, 147.25).

(f) Prediction intervals for the two elements of $\mathbf{y}_0$ are: (160.23, 190.52) and (131.82, 154.94).

10.17 **(a)** $\hat{\mathbf{B}} = \begin{pmatrix} 54.870 & 65.679 & 58.106 \\ .054 & -.048 & .018 \\ -.024 & .163 & .012 \\ .107 & -.036 & .125 \end{pmatrix}$

(b) $\Lambda = .665, V^{(s)} = .365, U^{(s)} = .458, \theta = .240$

(c) $\lambda_1 = .3159, \lambda_2 = .1385, \lambda_3 = .0037$. The essential rank of $\hat{\mathbf{B}}_1$ is 2, and the power ranking is $V^{(s)} > \Lambda > U^{(s)} > \theta$.

(d)

	Λ	F	p-Value
$x_1 \mid x_2, x_3$	.942	.903	.447
$x_2 \mid x_1, x_3$	.847	2.653	.060
$x_3 \mid x_1, x_2$	.829	3.020	.040

(e)

	Λ	F	p-Value
$y_1 \mid y_2, y_3$	.890	1.804	.160
$y_2 \mid y_1, y_3$	.833	2.932	.044
$y_3 \mid y_1, y_2$	.872	2.159	.106

(f) Confidence intervals for the three elements of $E(\mathbf{y}_0)$ are: (65.55, 71.62), (71.06, 76.27), and (70.96, 76.29).

(g) Prediction intervals for the three elements of $\mathbf{y}_0$ are: (49.07, 88.10), (56.92, 90.41), and (56.51, 90.74).

10.18 **(a)** $$\hat{\mathbf{B}} = \begin{pmatrix} -4.140 & 4.935 \\ 1.103 & -.955 \\ .231 & -.222 \\ 1.171 & 1.773 \\ .111 & .048 \\ .617 & -.058 \\ .267 & .485 \\ -.263 & -.209 \\ -.004 & -.004 \end{pmatrix}$$

Test of overall regression of (y_1, y_2) on $(x_1, x_2, \ldots, x_8)$: $\Lambda = .4642$ (with $p = 2$, exact $F = 1.169, p$-value $= .332$). Tests on subsets (the F's are exact because $p = 2$):

	Λ	F	p-Value
(b) $x_7, x_8 \mid x_1, x_2, \ldots, x_6$	.856	.808	.527
(c) $x_4, x_5, x_6 \mid x_1, x_2, x_3, x_7, x_8$	.674	1.457	.218
(d) $x_1, x_2, x_3 \mid x_4, x_5, \ldots, x_8$	.569	2.170	.066

10.19 **(a)** The overall test of (y_1, y_2) on $(x_1, x_2, \ldots, x_8)$ gives $\Lambda = .4642$, with (exact) $F = 1.169$ (p-value $= .332$). Even though this test result is not significant, we give the results of a backward elimination for illustrative purposes:

Partial Λ-Test on Each x_i Using (10.87)

Step	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	.723	.969	.817	.859	.821	.945	.924	.943
2	.741		.801	.851	.839	.948	.927	.940
3	.737		.837	.798	.757		.949	.938
4	.675		.852	.821	.794			.925
5	.680		.861	.835	.817			
6	.701			.805	.806			
7	.855			.930				
8	.891							

At each step, the variable deleted was not significant. In fact, the variable remaining at the last step, x_1, is not a significant predictor of y_1 and y_2.

(b) There were no significant x's, but to illustrate, we will use the three x's at step 6 above and test for each y:

	Λ	F	p-Value
$y_1 \mid y_2$	.701	3.548	.029
$y_2 \mid y_1$	.808	1.984	.142

10.20 (a) $\hat{\mathbf{B}} = \begin{pmatrix} 43.703 & 46.793 & 187.923 \\ .019 & -.098 & 1.016 \\ .139 & .185 & -4.953 \\ .204 & .107 & 1.606 \end{pmatrix}$

$\Lambda = .167, V^{(s)} = .883, U^{(s)} = 4.709, \theta = .823$

(b) $\hat{\mathbf{B}} = \begin{pmatrix} 99.817 & -29.120 & 121.595 \\ -.008 & -.224 & -.027 \\ .097 & 1.252 & 5.775 \\ -.049 & -.442 & -1.768 \\ -.022 & -.631 & -.488 \\ -.159 & 2.128 & 4.387 \\ .054 & -.037 & -.476 \end{pmatrix}$

$\Lambda = .110, V^{(s)} = 1.350, U^{(s)} = 4.319, \theta = .769$

(c) $\hat{\mathbf{B}} = \begin{pmatrix} 710.236 & 123.403 \\ -1.625 & .055 \\ 24.648 & .094 \\ -8.622 & -.334 \\ -8.224 & .462 \\ 23.626 & -.110 \\ 2.862 & .427 \\ -16.186 & -.267 \\ -.268 & .014 \\ -1.160 & -.336 \end{pmatrix}$

$\Lambda = .102, V^{(s)} = 1.236, U^{(s)} = 5.475, \theta = .827$

10.21 Using a backward elimination based on (10.87), we obtain the following partial Λ-values:

Step	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
1	.993	.962	.916	.958	.919	.879	.981	.999	.797
2	.994	.962	.916	.956	.909	.874	.980		.626
3		.951	.883	.954	.912	.873	.981		.626
4		.948	.884	.955	.861	.867			.561
5		.953	.862		.840	.803			.561
6			.830		.781	.783			.535

At step 6, we stop and retain all four x's because each Λ has a p-value less than .05.

CHAPTER 11

11.1 By (3.38), $\mathbf{S}_{yy} = \mathbf{D}_y\mathbf{R}_{yy}\mathbf{D}_y$ and $\mathbf{S}_{xx} = \mathbf{D}_x\mathbf{R}_{xx}\mathbf{D}_x$, where $\mathbf{D}_y$ and $\mathbf{D}_x$ are defined below (11.14). Similarly, $\mathbf{S}_{yx} = \mathbf{D}_y\mathbf{R}_{yx}\mathbf{D}_x$ and $\mathbf{S}_{xy} = \mathbf{D}_x\mathbf{R}_{xy}\mathbf{D}_y$. Substitute these into (11.7), replace $\mathbf{I}$ by $\mathbf{D}_y^{-1}\mathbf{D}_y$, and factor out $\mathbf{D}_y$ on the right.

11.2 Multiply (11.7) by $\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}$ on the left to obtain $(\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy} - r^2\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy})\mathbf{a} = \mathbf{0}$. Factor out $\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}$ on the right to write this in the form $(\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{S}_{yy}^{-1}\mathbf{S}_{yx} - r^2\mathbf{I})\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a} = \mathbf{0}$. Upon comparing this to (11.8), we see that $\mathbf{b} = \mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a}$.

11.3 When $p = 1, s$ is also 1, and there is only one canonical correlation, which is equal to R^2 from multiple regression [see comments between (11.28) and (11.29)]. Thus

$$\Lambda = \frac{1 - r_1^2}{1 - c_1^2} = \frac{1 - R_f^2}{1 - R_r^2}.$$

11.4
$$\begin{aligned} F &= \frac{(1 - \Lambda)(n - q - 1)}{\Lambda h} = \frac{[1 - (1 - R_f^2)/(1 - R_r^2)](n - q - 1)}{[(1 - R_f^2)/(1 - R_r^2)]h} \\ &= \frac{[1 - R_r^2 - (1 - R_f^2)](n - q - 1)}{(1 - R_f^2)h} \\ &= \frac{(R_f^2 - R_r^2)(n - q - 1)}{(1 - R_f^2)h} \end{aligned}$$

11.5 By (11.39),

$$r_i^2 = \frac{\lambda_i}{1 + \lambda_i}, \quad r_i^2 + r_i^2\lambda_i = \lambda_i, \quad \lambda_i(1 - r_i^2) = r_i^2.$$

11.6 Substitute $\mathbf{E} = (n - 1)(\mathbf{S}_{yy} - \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy})$ and $\mathbf{H} = (n - 1)\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}$ from (11.44) and (11.45) into (11.37):

$$\begin{aligned} \mathbf{Ha} &= \lambda\mathbf{Ea}, \\ (n - 1)\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a} &= (n - 1)\lambda(\mathbf{S}_{yy} - \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy})\mathbf{a}, \\ \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a} &= \lambda(\mathbf{S}_{yy} - \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy})\mathbf{a}. \end{aligned}$$

11.7 By (11.42), $\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a} = r^2\mathbf{S}_{yy}\mathbf{a}$. Subtracting $r^2\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a}$ from both sides gives

$$\begin{aligned} \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a} - r^2\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a} &= r^2\mathbf{S}_{yy}\mathbf{a} - r^2\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a}, \\ (1 - r^2)\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}\mathbf{a} &= r^2(\mathbf{S}_{yy} - \mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy})\mathbf{a}. \end{aligned}$$

11.8 **(a)** $r_1 = .5142, \quad r_2 = .1255$

(b)

	$\mathbf{c}_1$	$\mathbf{c}_2$
y_1	1.020	−.048
y_2	−.160	1.009

	$\mathbf{d}_1$	$\mathbf{d}_2$
x_1	.436	.823
x_2	−.704	−.455
x_3	1.081	−.401

(c)

k	Λ	Approximate F	p-Value
1	.7240	2.395	.035
2	.9843	.336	.716

11.9 **(a)** $r_1 = .7885$, $r_2 = .0537$

(b)

	$\mathbf{c}_1$	$\mathbf{c}_2$		$\mathbf{d}_1$	$\mathbf{d}_2$
y_1	.5522	-1.3664	x_1	.5044	−1.7686
y_2	.5215	1.3784	x_2	.5383	1.7586

(c)

k	Λ	Approximate F	p-Value
1	.3772	6.5972	.0003
2	.9971	0.0637	.8031

11.10 **(a)** $r_1 = .4900$, $r_2 = .3488$, $r_3 = .0609$

(b)

	$\mathbf{c}_1$	$\mathbf{c}_2$	$\mathbf{c}_3$		$\mathbf{d}_1$	$\mathbf{d}_2$	$\mathbf{d}_3$
y_1	.633	.091	.806	x_1	.482	−.262	1.054
y_2	−.624	816	.147	x_2	−.578	1.024	−.059
y_3	.643	.400	−.690	x_3	.865	.216	−.626

(c)

k	Λ	Approximate F	p-Value
1	.665	2.175	.029
2	.875	1.552	.194
3	.996	.171	.681

11.11 **(a)** $r_1 = .6251$, $r_2 = .4135$

(b)

	$\mathbf{c}_1$	$\mathbf{c}_2$		$\mathbf{d}_1$	$\mathbf{d}_2$
y_1	1.120	−.007	x_1	1.091	−.794
y_2	−.498	1.003	x_2	.184	−.288
			x_3	.842	1.807
			x_4	.944	.641
			x_5	1.040	−.154
			x_6	.215	1.256
			x_7	−.603	−.528
			x_8	−.641	−.588

(c)

k	Λ	Approximate F	p-Value
1	.4642	1.1692	.3321
2	.7553	.9718	.4766

11.12 **(b)** By (11.34),

$$\Lambda(x_7, x_8 | x_1, x_2, \ldots, x_6) = \frac{\prod_{i=1}^2 (1 - r_i^2)}{\prod_{i=1}^2 (1 - c_i^2)},$$

where r_1^2 and r_2^2 are the squared canonical correlations from the full model and c_1^2 and c_2^2 are the squared canonical correlations from the reduced model:

$$\Lambda(x_7, x_8 | x_1, x_2, \ldots, x_6) = \frac{(1 - .6208^2)(1 - .4947^2)}{(1 - .2650^2)(1 - .0886^2)} = \frac{.4643}{.9225} = .5033$$

(c) $$\Lambda(x_4, x_5, x_6 | x_1, x_2, x_3, x_7, x_8) = \frac{(1 - .6208^2)(1 - .4947^2)}{(1 - .3301^2)(1 - .1707^2)} = \frac{.4643}{.8651} = .5367$$

(d) $$\Lambda(x_1, x_2, x_3 | x_4, x_5, \ldots, x_8) = \frac{(1 - .6208^2)(1 - .4947^2)}{(1 - .4831^2)(1 - .2185^2)} = \frac{.4643}{.7300} = .6359$$

11.13 **(a)** $r_1 = .9279, r_2 = .5622, r_3 = .1660,$

k	Λ	Approximate F	p-Value
1	.0925	17.9776	<.0001
2	.6651	4.6366	.0020
3	.9725	1.1898	.2816

(b) $r_1 = .8770, r_2 = .6776, r_3 = .3488,$

k	Λ	Approximate F	p-Value
1	.1097	6.919	<.0001
2	.4751	3.427	.001
3	.8783	1.351	.269

(c) $r_1 = .9095, r_2 = .6395,$

k	Λ	Approximate F	p-Value
1	.1022	8.2757	<.0001
2	.5911	3.1129	.0089

(d) $r_1 = .9029, r_2 = .7797, r_3 = .3597, r_4 = .3233, r_5 = .0794,$

k	Λ	Approximate F	p-Value
1	.0561	4.992	<.0001
2	.3037	2.601	.0007
3	.7747	.829	.6210
4	.8898	.761	.6030
5	.9937	.124	.8840

CHAPTER 12

12.1 From $\lambda = \mathbf{a}'\mathbf{S}\mathbf{a}/\mathbf{a}'\mathbf{a}$ in (12.7), we obtain $\lambda\mathbf{a}'\mathbf{a} = \mathbf{a}\mathbf{S}\mathbf{a}$, which can be factored as $\mathbf{a}'(\mathbf{S}\mathbf{a} - \lambda\mathbf{a}) = 0$. Since $\mathbf{a} = \mathbf{0}$ is not a solution to $\lambda = \mathbf{a}'\mathbf{S}\mathbf{a}/\mathbf{a}'\mathbf{a}$, we have $\mathbf{S}\mathbf{a} - \lambda\mathbf{a} = \mathbf{0}$.

12.2 $|\mathbf{R} - \lambda\mathbf{I}| = 0,$ $\quad \begin{vmatrix} 1-\lambda & r \\ r & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - r^2 = 0,$

$(1 - \lambda + r)(1 - \lambda - r) = 0, \quad \lambda = 1 \pm r$

With $\lambda_1 = 1 + r$ in $(\mathbf{R} - \lambda_1\mathbf{I})\mathbf{a}_1 = \mathbf{0}$, we obtain

$$\begin{pmatrix} -r & r \\ r & -r \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which gives $a_{11} = a_{12}$ for any r. Normalizing to $\mathbf{a}_1'\mathbf{a}_1 = 1$, yields $a_{11} = 1/\sqrt{2}$.

12.3 **(a)** By (4.14) and a comment in Section 7.1, the likelihood ratio is given by LR = $(|\mathbf{S}|/|\mathbf{S}_0|)^{n/2}$, where $\mathbf{S}_0$ is the estimate of $\mathbf{\Sigma}$ under H_0. By (2.108) and (7.6), the test statistic is

$$-2\ln \mathrm{LR} = -2\ln\left(\frac{|\mathbf{S}|}{|\mathbf{S}_0|}\right)^{n/2} = -2\left(\frac{n}{2}\right)\ln\frac{\prod_{i=1}^{p-k}\lambda_i\prod_{i=p-k+1}^{p}\lambda_i}{\prod_{i=1}^{p-k}\lambda_i\prod_{i=p-k+1}^{p}\overline{\lambda}}$$

$$= -n\ln\frac{\prod_{i=p-k+1}^{p}\lambda_i}{\overline{\lambda}^k} = n\left(k\ln\overline{\lambda} - \sum_{i=p-k+1}^{p}\lambda_i\right).$$

In (12.15), the coefficient n is modified to give an improved chi-square approximation.

12.4 If $\mathbf{S}$ is diagonal, then $\lambda_i = s_{ii}$, as in (12.17). Thus

$$\mathbf{S}\mathbf{a}_i = \lambda_i\mathbf{a}_i = s_{ii}\mathbf{a}_i$$

$$\begin{pmatrix} s_{11} & 0 & \cdots & 0 \\ 0 & s_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & s_{pp} \end{pmatrix}\begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{ip} \end{pmatrix} = \begin{pmatrix} s_{11}a_{i1} \\ s_{22}a_{i2} \\ \vdots \\ s_{pp}a_{ip} \end{pmatrix} = \begin{pmatrix} s_{ii}a_{i1} \\ s_{ii}a_{i2} \\ \vdots \\ s_{ii}a_{ip} \end{pmatrix}$$

From the first element, we obtain $s_{11}a_{i1} = s_{ii}a_{i1}$ or $(s_{11} - s_{ii})a_{i1} = 0$. Since $s_{11} - s_{ii} \neq 0$ (unless $i = 1$), we must have $a_{i1} = 0$. Thus, $\mathbf{a}_i = (0, \ldots, 0, a_{ii}, 0, \ldots, 0)'$, and normalizing $\mathbf{a}_i$ leads to $a_{ii} = 1$.

12.5 By (10.34) and (12.2),

$$R^2_{y_i|z_1,\ldots,z_k} = \frac{\mathbf{s}'_{y_iz}\mathbf{S}_{zz}^{-1}\mathbf{s}_{y_iz}}{s^2_{y_i}}$$

$$= \left(s_{y_iz_1}, s_{y_iz_2}, \ldots, s_{y_iz_k}\right)\begin{pmatrix} s^2_{z_1} & 0 & \cdots & 0 \\ 0 & s^2_{z_2} & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & s^2_{z_k} \end{pmatrix}^{-1}\begin{pmatrix} s_{y_iz_1} \\ s_{y_iz_2} \\ \vdots \\ s_{y_iz_k} \end{pmatrix}\Bigg/ s^2_{y_i}.$$

Show that this is equal to

$$R^2_{y_i|z_1,\ldots,z_k} = \sum_{j=1}^{k}\frac{s^2_{y_iz_j}}{s^2_{z_j}s^2_{y_i}} = \sum_{j=1}^{k} r^2_{y_iz_j}.$$

12.6 The variances of y_1, y_2, x_1, x_2, and x_3 on the diagonal of **S** are .016, 70.6, 1106.4, 2381.9, and 2136.4. The eigenvalues of **S** and **R** are as follows:

S			**R**		
λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative	λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative
3466.18	.608607	.60861	1.72	.34	.34
1264.47	.222021	.83063	1.23	.25	.59
895.27	.157195	.98782	.96	.19	.78
69.34	.012174	.99999	.79	.16	.94
.01	.000002	1.00000	.30	.06	1.00

Two principal components of **S** account for 83% of the variance, but it requires three principal components of **R** to reach 78%. For most purposes we would use two components of **S**, although with three we could account for 99% of the variance. However, we show all five eigenvectors below because of the interesting pattern they exhibit. The first principal component is largely a weighted average of the last two variables, x_2 and x_3, which have the largest variances. The second and third components represent contrasts in the last three variables and could be described as "shape" components. The fourth and fifth components are associated uniquely with y_2 and y_1, respectively. These components are "variable specific," as described in the discussion of method 1 in Section 12.6.

As expected, the principal components of **R** show an entirely different pattern. All five variables contribute to the first three components of **R**, whereas in **S**, y_1 and y_2 have small variances and contribute almost nothing to the first three components. The eigenvectors of **S** and **R** are as follows:

	S					**R**				
	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_4$	$\mathbf{a}_5$	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_4$	$\mathbf{a}_5$
y_1	.0004	−.0008	.0018	.0029	.9999	.42	.53	−.42	−.40	.46
y_2	−.0080	.0166	.0286	.9994	−.0029	.07	.68	.16	.70	−.10
x_1	.1547	.6382	.7535	−.0309	−.0008	.36	.20	.76	−.44	−.24
x_2	.7430	.4279	−.5145	.0136	.0009	.54	−.43	.25	.39	.56
x_3	.6511	−.6397	.4083	.0042	−.0015	.63	−.18	−.40	.10	−.64

12.7 $$\mathbf{S} = \begin{pmatrix} 65.1 & 33.6 & 47.6 & 36.8 & 25.4 \\ 33.6 & 46.1 & 28.9 & 40.3 & 28.4 \\ 47.6 & 28.9 & 60.7 & 37.4 & 41.1 \\ 36.8 & 40.3 & 37.4 & 62.8 & 31.7 \\ 25.4 & 28.4 & 41.1 & 31.7 & 58.2 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} 1.00 & .61 & .76 & .58 & .41 \\ .61 & 1.00 & .55 & .75 & .55 \\ .76 & .55 & 1.00 & .61 & .69 \\ .58 & .75 & .61 & 1.00 & .52 \\ .41 & .55 & .69 & .52 & 1.00 \end{pmatrix}$$

The eigenvalues of **S** and **R** are as follows:

S			R		
λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative	λ_i	$\lambda_i/\sum_j \lambda_i$	Cumulative
200.4	.684	.684	3.42	.683	.683
36.1	.123	.807	.61	.123	.806
34.1	.116	.924	.57	.114	.921
15.0	.051	.975	.27	.054	.975
7.4	.025	1.000	.13	.025	1.000

The first three eigenvectors of **S** and **R** are as follows:

S			R		
$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$
.47	−.58	−.42	.44	−.20	−.68
.39	−.11	.45	.45	−.43	.35
.49	.10	−.48	.47	.37	−.38
.47	−.12	.62	.45	−.39	.33
.41	.80	−.09	.41	.70	.41

The variances in **S** are nearly identical, and the covariances are likewise similar in magnitude. Consequently, the percent of variance explained by the eigenvalues of **S** and **R** are indistinguishable. The interpretation of the second principal component from **S** is slightly different from that of the second one from **R**, but otherwise there is little to choose between them.

12.8 The variances on the diagonal of **S** are 95.5, 73.2, 76.2, 808.6, 505.9, and 508.7. The eigenvalues of **S** and **R** are as follows:

S			R		
λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative	λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative
1152.0	.557	.557	2.17	.363	.363
394.1	.191	.748	1.08	.180	.543
310.8	.150	.898	.98	.163	.706
97.8	.047	.945	.87	.144	.850
68.8	.033	.978	.55	.092	.942
44.6	.022	1.000	.35	.058	1.000

We could keep either two or three components from **S**. The first three components of **S** account for a larger percent of variance than do those from **R**. The first three eigenvectors of **S** and **R** are as follows:

S			R		
$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$
.080	.092	−.069	.336	.176	.497
.034	−.018	.202	.258	.843	−.093
.076	.122	−.011	.370	.049	.466
.758	−.446	−.469	.475	−.329	−.358
.493	−.081	.844	.486	.079	−.567
.412	.878	−.147	.471	−.376	.278

As expected, the first three principal components from **S** are heavily influenced by the last three variables because of their relatively large variances.

12.9 The variances on the diagonal of **S** are .69; 5.4; 2,006, 682.4; 90.3; 56.4; 18.1. With the large variance of y_3, we would expect the first principal component from **S** to account for most of the variance, and y_3 would essentially constitute that single component. This is indeed the pattern that emerges in the eigenvalues and eigenvectors of **S**. The principal components from **R**, on the other hand, are not dominated by y_3. The eigenvalues of **S** and **R** are as follows:

S		R		
λ_i	$\lambda_i/\sum_j \lambda_j$	λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative
2,006,760	.999954	2.42	.404	.404
65	.000033	1.40	.234	.638
18	.000009	1.03	.171	.809
7	.000003	.92	.153	.963
3	.000001	.20	.033	.996
0	.000000	.02	.004	1.000

Most of the correlations in **R** are small (only three exceed .3), and its first three principal components account for only 72% of the variance. The first three eigenvectors of **S** and **R** are as follows:

S			R		
$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$
.00016	.005	−.0136	.424	−.561	−.150
.00051	.017	.0787	.446	−.528	.087
.99998	−.001	−.0002	.563	.387	−.051
.00529	.698	.0174	.454	.267	.166
.00322	−.716	.0195	.303	.425	−.296
.00020	.025	.9965	.073	.069	.923

12.10 Covariance matrix for males:

$$\mathbf{S}_M = \begin{pmatrix} 5.19 & 4.55 & 6.52 & 5.25 \\ 4.55 & 13.18 & 6.76 & 6.27 \\ 6.52 & 6.76 & 28.67 & 14.47 \\ 5.25 & 6.27 & 14.47 & 16.65 \end{pmatrix}$$

Covariance matrix for females:

$$\mathbf{S}_F = \begin{pmatrix} 9.14 & 7.55 & 4.86 & 4.15 \\ 7.55 & 18.60 & 10.22 & 5.45 \\ 4.86 & 10.22 & 30.04 & 13.49 \\ 4.15 & 5.45 & 13.49 & 28.00 \end{pmatrix}$$

The eigenvalues are as follows:

Males			Females		
λ_i	$\lambda_i / \sum_j \lambda_j$	Cumulative	λ_i	$\lambda_i / \sum_j \lambda_j$	Cumulative
43.56	.684	.684	48.96	.571	.571
11.14	.175	.858	18.46	.215	.786
6.47	.102	.960	13.54	.158	.944
2.52	.040	1.000	4.82	.056	1.000

The first two eigenvectors are as follows:

Males		Females	
$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_1$	$\mathbf{a}_2$
.24	.21	.22	.27
.31	.85	.39	.62
.76	−.48	.68	.17
.52	.09	.58	−.72

The variances in $\mathbf{S}_M$ have a slightly wider range (5.19 to 28.67) than those in $\mathbf{S}_F$ (9.14 to 30.04), and this is reflected in the eigenvalues. The first two components account for 86% of the variance from $\mathbf{S}_M$, whereas the first two account for 79% from $\mathbf{S}_F$.

12.11 Covariance matrix for species 1:

$$\mathbf{S}_1 = \begin{pmatrix} 187.6 & 176.9 & 48.4 & 113.6 \\ 176.9 & 345.4 & 76.0 & 118.8 \\ 48.4 & 76.0 & 66.4 & 16.2 \\ 113.6 & 118.8 & 16.2 & 239.9 \end{pmatrix}$$

Covariance matrix for species 2:

$$\mathbf{S}_2 = \begin{pmatrix} 101.8 & 128.1 & 37.0 & 32.6 \\ 128.1 & 389.0 & 165.4 & 94.4 \\ 37.0 & 165.4 & 167.5 & 66.5 \\ 32.6 & 94.4 & 66.5 & 177.9 \end{pmatrix}$$

The eigenvalues are as follows:

Species 1			Species 2		
λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative	λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative
561.3	.669	.669	555.7	.664	.664
169.0	.201	.870	145.4	.174	.838
65.3	.078	.948	93.5	.112	.950
43.7	.057	1.000	41.7	.050	1.000

The first two eigenvectors are as follows:

Species 1		Species 2	
$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_1$	$\mathbf{a}_2$
.50	.01	.28	−.20
.72	−.48	.81	−.34
.17	−.22	.42	.14
.45	.85	.30	.91

The variances in $\mathbf{S}_1$ have a wider range than those in $\mathbf{S}_2$, and the first two components of $\mathbf{S}_1$ account for a higher percent of variance.

12.12 The variances on the diagonal of $\mathbf{S}$ in each case are:

(a) Pooled: 536.0, 59.9, 116.0, 896.4, 248.1, 862.0

(b) Unpooled: 528.2, 68.9, 145.2, 1366.4, 264.4, 1069.1

The eigenvalues are as follows:

Pooled			Unpooled		
λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative	λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative
1050.6	.386	.386	1722.0	.500	.500
858.3	.316	.702	878.4	.255	.755
398.9	.147	.849	401.4	.117	.872
259.2	.095	.944	261.1	.076	.948
108.1	.040	.984	128.9	.037	.985
43.4	.016	1.000	50.4	.015	1.000

The first three eigenvectors are as follows:

Pooled			Unpooled		
$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$
.441	−.190	.864	.212	.389	.888
.041	−.038	.082	−.039	.064	.096
−.039	.031	.143	.080	−.066	.081
.450	.892	−.033	.776	−.608	.081
−.019	−.001	−.054	−.096	.010	.015
.774	−.407	−.471	.580	.686	−.434

(c) The pattern in both eigenvalues and eigenvectors is similar for the pooled and unpooled cases. The first three principal components account for 87.2% of the variance in the unpooled case compared to 84.9% for the pooled case.

12.13 The variances on the diagonal of **S** in each case are:

(a) Pooled: 49.1, 8.1, 12140.8, 136.2, 210.8, 2983.9

(b) Unpooled: 63.2, 8.0, 15168.9, 186.6, 255.4, 4660.7

The eigenvalues are as follows:

Pooled			Unpooled		
λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative	λ_i	$\lambda_i/\sum_j \lambda_j$	Cumulative
12,809.0	.8249	.8249	17,087.0	.8400	.8400
2,455.9	.1582	.9830	2,958.0	.1454	.9854
137.1	.0088	.9918	168.6	.0083	.9937
77.2	.0050	.9968	77.1	.0038	.9974
42.2	.0027	.9995	44.7	.0022	.9996
7.4	.0005	1.0000	7.3	.0004	1.0000

The eigenvectors are as follows:

Pooled		Unpooled	
$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_1$	$\mathbf{a}_2$
−.004	−.000	.013	.027
−.005	.004	−.004	.004
.968	−.233	.931	−.355
−.002	.023	.028	.069
.103	.041	.103	.021
.228	.971	.350	.932

12.14 The variances on the diagonal of **S** are all less than 1 except $s_{x_4}^2 = 5.02$ and $s_{x_8}^2 = 1541.08$. We therefore expect the last variable, x_8, to dominate the principal components of **S**. This is the case for **S** but not for **R**. The eigenvalues of **S** and **R** are as follows:

S		R		
λ_i	$\lambda_i / \sum_j \lambda_j$	λ_i	$\lambda_i / \sum_j \lambda_j$	Cumulative
1541.55	.996273	3.174	.317	.317
4.83	.003123	2.565	.256	.574
.44	.000286	1.432	.143	.717
.27	.000174	1.277	.128	.845
.10	.000066	.542	.054	.899
.07	.000043	.473	.047	.946
.02	.000014	.251	.025	.971
.02	.000011	.118	.012	.983
.01	.000005	.104	.010	.994
.00	.000003	.064	.006	1.000

The eigenvectors of **S** and **R** are as follows:

S		R			
$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_4$
.0009	−.005	.12	.19	.69	.10
.0007	−.034	.06	.32	.54	.26
.0029	−.007	.46	−.06	.07	−.38
.0014	.004	.29	.17	−.18	.49
.0059	−.009	.52	.14	−.04	−.01
−.0150	.982	−.09	−.42	.07	.55
−.0028	−.092	−.31	.45	−.01	−.14
−.0022	−.158	−.23	.54	−.14	−.10
.0044	−.011	.09	.36	−.38	.44
.9998	.014	.50	.11	−.13	−.09

12.15 The variances in the diagonal of **S** are: 55.7, 10.9, 402.7, 25.7, 13.4, 438.3, 1.5, 106.2, 885.6, 22227.2, 214.1

The eigenvalues of **S** and **R** are as follows:

S			R		
λ_i	$\lambda_i / \sum_j \lambda_j$	Cumulative	λ_i	$\lambda_i / \sum_j \lambda_j$	Cumulative
22,303.5	.91479	.91479	6.020	.54730	.54730
1590.7	.06524	.98003	2.119	.19267	.73996
358.0	.01469	.99471	1.130	.10275	.84272
63.4	.00260	.99731	.760	.06909	.91181
29.3	.00120	.99852	.355	.03231	.94411
17.1	.00070	.99922	.259	.02358	.96769
12.7	.00052	.99974	.122	.01110	.97879
2.8	.00012	.99986	.110	.01004	.98883
1.9	.00008	.99994	.060	.00544	.99427
.9	.00004	.99997	.042	.00384	.99810
.7	.00003	1.00000	.021	.00190	1.00000

The eigenvectors of $\mathbf{S}$ and $\mathbf{R}$ are as follows:

	S		R			
	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_4$
y_1	−.0097	.1331	.3304	−.0787	.0880	−.2807
y_2	.0006	.0608	.3542	.1928	.1071	−.2301
y_3	−.0141	.4397	.3923	.0518	.1105	−.1413
y_4	−.0033	.1078	.3820	.0474	.1334	−.0104
y_5	.0101	.0398	.2323	.5303	.0154	−.0710
y_6	.0167	.4290	.3621	.2361	.1198	.1350
y_7	−.0012	−.0072	−.0884	.0213	.7946	.5414
y_8	.0275	−.1844	−.2501	.5023	.0826	−.1506
y_9	.0456	−.6657	−.3111	.3595	.2136	−.2278
y_{10}	.9982	.0346	−.0243	.4685	−.4669	.5001
y_{11}	.0034	.3311	.3357	−.1153	−.1853	.4550

For most purposes, one or two principal components would suffice for $\mathbf{S}$, with 91% or 98% of the variance explained. For $\mathbf{R}$, on the other hand, three components are required to explain 84% of the variance, and seven components are necessary to reach 98%. The reduction to one or two components for $\mathbf{S}$ is due in part to the relatively large variances of y_3, y_6, y_9, and y_{10}. In the eigenvectors of $\mathbf{S}$, we see that these four variables figure prominently in the first two principal components.

CHAPTER 13

13.1
$$\begin{aligned}\text{var}(y_i) &= \text{var}(y_i - \mu_i) = \text{var}(\lambda_{i1}f_1 + \lambda_{i2}f_2 + \cdots + \lambda_{im}f_m + \varepsilon_i)\\ &= \textstyle\sum_{j=1}^m \lambda_{ij}^2 \text{var}(f_j) + \text{var}(\varepsilon_i) + \sum_{j\neq k}\lambda_{ij}\lambda_{ik}\text{cov}(f_j, f_k)\\ &\quad + \textstyle\sum_{j=1}^m \lambda_{ij}\text{cov}(f_j, \varepsilon_i)\\ &= \textstyle\sum_{j=1}^m \lambda_{ij}^2 + \psi_i.\end{aligned}$$

The last equality follows by the assumptions $\text{var}(f_j) = 1, \text{var}(\varepsilon_i) = \psi_i$, $\text{cov}(f_j, f_k) = 0$, and $\text{cov}(f_j, \varepsilon_i) = 0$.

13.2
$$\begin{aligned}\text{cov}(\mathbf{y}, \mathbf{f}) &= \text{cov}(\mathbf{\Lambda f} + \boldsymbol{\varepsilon}, \mathbf{f}) && \text{[by (13.3)]}\\ &= \text{cov}(\mathbf{\Lambda f}, \mathbf{f}) && \text{[by (13.10)]}\\ &= E[\mathbf{\Lambda f} - E(\mathbf{\Lambda f})][\mathbf{f} - E(\mathbf{f})]' && \text{[by analogy to (3.31)]}\\ &= E[\mathbf{\Lambda f} - \mathbf{\Lambda}E(\mathbf{f})][\mathbf{f} - E(\mathbf{f})]'\\ &= \mathbf{\Lambda}E[\mathbf{f} - E(\mathbf{f})][\mathbf{f} - E(\mathbf{f})]'\\ &= \mathbf{\Lambda}\text{cov}(\mathbf{f}) = \mathbf{\Lambda} && \text{[by (13.7)]}\end{aligned}$$

13.3 $E(\mathbf{f}^*) = E(\mathbf{T}'\mathbf{f}) = \mathbf{T}'E(\mathbf{f}) = \mathbf{T}'\mathbf{0} = \mathbf{0}$,
$\text{cov}(\mathbf{f}^*) = \text{cov}(\mathbf{T}'\mathbf{f}) = \mathbf{T}'\text{cov}(\mathbf{f})\mathbf{T} = \mathbf{T}'\mathbf{IT} = \mathbf{I}$

13.4 Let $\mathbf{E} = \mathbf{S} - (\hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}' + \hat{\mathbf{\Psi}})$. Then by (2.98), $\text{tr}(\mathbf{E}'\mathbf{E}) = \sum_{ij} e_{ij}^2$. By (13.26), $\hat{\mathbf{\Psi}} = \text{diag}(\mathbf{S} - \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}')$, and $\mathbf{E}$ has zeros on the diagonal. This gives the inequality

$\sum_{ij} e_{ij}^2 \leq$ sum of squared elements of $\mathbf{S} - \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}'$.

By (2.98),

$$\text{sum of squared elements of } \mathbf{S} - \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}' = \text{tr}(\mathbf{S} - \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}')'(\mathbf{S} - \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}').$$

Since $\mathbf{S} - \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}'$ is symmetric, we have by (13.20), (13.23), and (13.24),

$$\begin{aligned}\mathbf{S} - \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}' &= \mathbf{CDC}' - \mathbf{C}_1\mathbf{D}_1^{1/2}\mathbf{D}_1^{1/2}\mathbf{C}_1' \\ &= \mathbf{CDC}' - \mathbf{C}_1\mathbf{D}_1\mathbf{C}_1',\end{aligned}$$

where $\mathbf{C} = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_p)$ contains normalized eigenvectors of $\mathbf{S}, \mathbf{D} = \text{diag}(\theta_1, \theta_2, \ldots, \theta_p)$ contains eigenvalues of $\mathbf{S}, \mathbf{C}_1 = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m)$, and $\mathbf{D}_1 = \text{diag}(\theta_1, \theta_2, \ldots, \theta_m)$.

Using the partitioned forms $\mathbf{C} = (\mathbf{C}_1, \mathbf{C}_2)$ and $\mathbf{D} = \begin{pmatrix} \mathbf{D}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{D}_2 \end{pmatrix}$, show that $\mathbf{C}_1'\mathbf{C}_1 = \mathbf{I}_m, \mathbf{C}_1'\mathbf{C}_2 = \mathbf{O}, \mathbf{C}'\mathbf{C}_1 = \begin{pmatrix} \mathbf{I}_m \\ \mathbf{O} \end{pmatrix}, \mathbf{D}\begin{pmatrix} \mathbf{I}_m \\ \mathbf{O} \end{pmatrix} = \begin{pmatrix} \mathbf{D}_1 \\ \mathbf{O} \end{pmatrix}, \mathbf{C}\begin{pmatrix} \mathbf{D}_1 \\ \mathbf{O} \end{pmatrix} = \mathbf{C}_1\mathbf{D}_1$, and $\mathbf{CDC}'\mathbf{C}_1\mathbf{D}_1\mathbf{C}_1' = \mathbf{C}_1\mathbf{D}_1^2\mathbf{C}_1'$. Show similarly that $\mathbf{C}_1\mathbf{D}_1\mathbf{C}_1'\mathbf{CDC}' = \mathbf{C}_1\mathbf{D}_1^2\mathbf{C}_1'$ and $\mathbf{C}_1\mathbf{D}_1\mathbf{C}_1'\mathbf{C}_1\mathbf{D}_1\mathbf{C}_1' = \mathbf{C}_1\mathbf{D}_1^2\mathbf{C}_1'$. Now by (2.97) $\text{tr}(\mathbf{CD}^2\mathbf{C}') = \text{tr}(\mathbf{C}'\mathbf{CD}^2) = \text{tr}(\mathbf{D}^2) = \sum_{i=1}^p \theta_i^2$. Similarly, $\text{tr}(\mathbf{C}_1\mathbf{D}_1^2\mathbf{C}_1') = \sum_{i=1}^m \theta_i^2$. Then

$$\begin{aligned}\text{tr}(\mathbf{S} - \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}')'(\mathbf{S} - \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}') &= \text{tr}(\mathbf{CDC}' - \mathbf{C}_1\mathbf{D}_1\mathbf{C}_1')(\mathbf{CDC}' - \mathbf{C}_1\mathbf{D}_1\mathbf{C}_1') \\ &= \text{tr}(\mathbf{CDC}'\mathbf{CDC}' - \mathbf{CDC}'\mathbf{C}_1\mathbf{D}_1\mathbf{C}_1' \\ &\quad - \mathbf{C}_1\mathbf{D}_1\mathbf{C}_1'\mathbf{CDC}' + \mathbf{C}_1\mathbf{D}_1\mathbf{C}_1'\mathbf{C}_1\mathbf{D}_1\mathbf{C}_1') \\ &= \textstyle\sum_{i=1}^p \theta_i^2 - \sum_{i=1}^m \theta_i^2 - \sum_{i=1}^m \theta_i^2 + \sum_{i=1}^m \theta_i^2 \\ &= \textstyle\sum_{i=m+1}^p \theta_i^2.\end{aligned}$$

13.5 $\sum_{i=1}^p \sum_{j=1}^m \hat{\lambda}_{ij}^2 = \sum_{i=1}^p \left[\sum_{j=1}^m \hat{\lambda}_{ij}^2\right] = \sum_{i=1}^p \hat{h}_i^2$ [by (13.28)]

By interchanging the order of summation, we have

$$\textstyle\sum_{i=1}^p \sum_{j=1}^m \hat{\lambda}_{ij}^2 = \sum_{j=1}^m \sum_{i=1}^p \hat{\lambda}_{ij}^2 = \sum_{j=1}^m \theta_j \quad \text{[by(13.29)]}.$$

13.6 We use the covariance matrix to avoid working with standardized variables. The eigenvalues of $\mathbf{S}$ are 39.16, 8.78, .66, .30, and 0. The eigenvector corresponding to $\lambda_5 = 0$ is

$$\mathbf{a}_5' = (-.75, -.25, .25, .50, .25).$$

As noted in Section 12.7, $s_{z_5}^2 = 0$ implies $z_5 = 0$. Thus

$$\begin{aligned}z_5 = \mathbf{a}_5'\mathbf{y} = -.75y_1 - .25y_2 + .25y_3 + .50y_4 + .25y_5 = 0, \\ 3y_1 + y_2 = y_3 + 2y_4 + y_5.\end{aligned}$$

13.7 Words data of Table 5.9:

	Principal Component Loadings f_1	f_2	Varimax Rotated Loadings f_1	f_2	Communalities, $\hat{h}_i^2$
Variables					
Informal words	.802	−.535	**.956**	.129	.930
Informal verbs	.856	−.326	**.858**	.321	.839
Formal words	.883	.270	.484	**.786**	.853
Formal verbs	.714	.658	.101	**.966**	.943
Variance	2.666	.899	1.894	1.671	3.565
Proportion	.666	.225	.474	.418	.891

The orthogonal matrix $\mathbf{T}$ for the varimax rotation as given by (13.49) is

$$\mathbf{T} = \begin{pmatrix} .750 & .661 \\ -.661 & .750 \end{pmatrix}.$$

Thus $\sin\phi = -.661$ and $\phi = -41.4°$. A graphical rotation of $-40°$ would produce results very close to the varimax rotation.

13.8 Ramus bone data of Table 3.7:

	Principal Component Loadings f_1	f_2	Varimax Rotated Loadings f_1	f_2	Communalities, $\hat{h}_i^2$	Orthoblique Pattern Loadings f_1	f_2
Variables							
8 years	.949	−.295	**.884**	.455	.988	−.108	**1.087**
$8\frac{1}{2}$ years	.974	−.193	**.830**	.545	.986	.106	**.900**
9 years	.978	.171	.578	**.808**	.986	**.825**	.188
$9\frac{1}{2}$ years	.943	.319	.449	**.888**	.991	**1.099**	−.121
Variance	3.695	.255	2.005	1.946	3.951		
Proportion	.924	.064	.501	.486	.988		

The Harris–Kaiser orthoblique rotation produced loadings for which the variables have a complexity of 1. These oblique loadings provide a much cleaner simple structure than the varimax loadings. For interpretation, we see that one factor represents variables 1 and 2, and the other factor represents variables 3 and 4. This same clustering of variables can be deduced from the varimax loadings if we simply use the larger of the two loadings for each variable.

The correlation between the two oblique factors is .87. The angle between the oblique axes is $\cos^{-1}(.87) = 29.5°$. With such a small angle between the axes and a large correlation between the factors, it is clear that a single factor would better represent the variables. This is also borne out by the eigenvalues of the correlation matrix: 3.695, .255, .033, and .017. The first accounts for 92% of the variance and the second for only 6%.

13.9 Rootstock data of Table 6.2:

	Principal Component Loadings		Varimax Rotated Loadings		Communalities,
	f_1	f_2	f_1	f_2	$\hat{h}_i^2$
Variables					
Trunk 4 years	.787	.575	.167	**.960**	.949
Extension 4 years	.849	.467	.287	**.925**	.939
Trunk 15 years	.875	−.455	**.946**	.280	.973
Weight 15 years	.824	−.547	**.973**	.179	.978
Variance	2.785	1.054	1.951	1.888	3.839
Proportion	.696	.264	.488	.472	.960

The rotation was successful in producing variables with a complexity of 1, that is, partitioning the variables into two groups, each with two variables.

13.10 Fish data of Table 6.17:

	Principal Component Loadings		Varimax Rotated Loadings		Communalities,
	f_1	f_2	f_1	f_2	$\hat{h}_i^2$
Variables					
y_1	.830	−.403	**.874**	.294	.851
y_2	.783	−.504	**.911**	.189	.866
y_3	.803	.432	.270	**.871**	.831
y_4	.769	.497	.200	**.893**	.838
Variance	2.537	.850	1.709	1.678	3.386
Proportion	.634	.213	.427	.420	.847

(b) The loadings for y_1 and y_2 are similar. In **R** below we see some indication of the reason for this; y_1 and y_2 are more highly correlated than any other pair of variables and their correlations with y_3 and y_4 are similar:

$$\mathbf{R} = \begin{pmatrix} 1.00 & .71 & .51 & .40 \\ .71 & 1.00 & .38 & .40 \\ .51 & .38 & 1.00 & .67 \\ .40 & .40 & .67 & 1.00 \end{pmatrix}.$$

(c) By (13.58), the factor score coefficient matrix is

$$\hat{\mathbf{B}}_1 = \mathbf{R}^{-1}\hat{\mathbf{\Lambda}} = \begin{pmatrix} .566 & -.109 \\ .636 & -.207 \\ -.130 & .584 \\ -.194 & .630 \end{pmatrix},$$

where $\hat{\mathbf{\Lambda}}$ is the matrix of rotated factor loadings given above in part (a). The factor scores are given by (13.59) as follows:

Method 1		Method 2		Method 3	
$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_1$	$\hat{f}_2$
.544	1.151	−.254	.309	−1.156	2.104
1.250	−.254	−.309	−1.534	−.321	.878
1.017	1.120	−1.865	−1.558	−.671	.947
−.147	−1.583	−.999	−.690	.067	1.130
.219	−.103	.520	−.343	−1.610	−.458
1.007	.679	.919	−.111	.557	.491
1.413	−.186	−.443	−.018	−.454	1.157
−.666	−2.279	−.265	.676	−.961	.063
1.057	−1.870	1.449	−.295	−.230	1.721
.388	−.440	1.371	.295	−1.309	.054
1.328	−.298	1.260	−.027	−1.766	−.111
.694	−.033	−.000	−1.452	−1.636	−.048

(d) A one-way MANOVA on the two factor scores comparing the three methods yielded the following values for $\mathbf{E}$ and $\mathbf{H}$:

$$\mathbf{E} = \begin{pmatrix} 21.8606 & 10.3073 \\ 10.3073 & 25.2081 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 13.1394 & -10.3073 \\ -10.3073 & 9.7919 \end{pmatrix}$$

The four MANOVA test statistics are $\Lambda = .3631, V^{(s)} = .6552, U^{(s)} = 1.7035$, and $\theta = .6259$. All are highly significant.

13.11 **(a)** For the flea data of Table 5.5, the eigenvalues of $\mathbf{R}$ are 2.273, 1.081, .450, and .196. There is a noticeable gap between 1,081 and .450, and the first two factors account for 83.9% of the variance. Thus $m = 2$ factors seem to be indicated for this set of data.

(b)

	Principal Component Loadings		Varimax Rotated Loadings		Communalities,	Orthoblique Pattern Loadings	
	f_1	f_2	f_1	f_2	$\hat{h}_i^2$	f_1	f_2
Variables							
y_1	−.038	.989	−.025	**.990**	.980	−.003	**.990**
y_2	.889	.269	**.892**	.256	.862	**.898**	.253
y_3	.893	−.157	**.891**	−.170	.823	**.887**	−.173
y_4	.827	−.073	**.823**	−.084	.689	**.824**	−.087
Variance	2.273	1.081	2.273	1.081	3.354		
Proportion	.568	.270	.568	.270	.839		

(The variance explained by the varimax rotated factors remains the same as for the initial factors when rounded to three decimal places.)

(c) In this case, neither of the rotations changes the initial loadings appreciably. The reason for this unusual outcome can be seen in the correlation matrix:

$$\mathbf{R} = \begin{pmatrix} 1.00 & .18 & -.17 & -.07 \\ .18 & 1.00 & .73 & .59 \\ -.17 & .73 & 1.00 & .59 \\ -.07 & .59 & .59 & 1.00 \end{pmatrix}.$$

There are clearly two clusters of variables: $\{y_1\}$ and $\{y_2, y_3, y_4\}$. We would expect two factors corresponding to these groupings to emerge after rotation. That the same pattern surfaces in the initial factor loadings (based on eigenvectors) is due to their affiliation with principal components. As noted in Section 12.8.1, if a variable has small correlations with all other variables, the variable itself will essentially constitute a principal component. In this case, y_1 has this property and makes up most of the second principal component. The first component is comprised of the other three variables.

13.12 (a) For the engineer data of Table 5.6, the number of eigenvalues greater than 1 is three, but the three account for only 70% of the variance. It requires four eigenvalues to reach 84%. The scree plot also indicates four eigenvalues.

(b)

	Principal Component Loadings			Varimax Rotated Loadings			Communalities,
	f_1	f_2	f_3	f_1	f_2	f_3	$\hat{h}_i^2$
Variables							
y_1	.536	.461	.478	−.063	**.834**	.170	.729
y_2	−.129	.870	−.182	−.357	.100	**.818**	.806
y_3	.514	−.254	−.448	**.724**	−.026	.068	.529
y_4	.724	−.366	−.110	**.739**	.295	−.193	.670
y_5	−.416	−.414	.649	−.484	−.013	**−.729**	.766
y_6	.715	.124	.420	.239	**.800**	−.069	.702
Variance	1.775	1.354	1.073	1.493	1.435	1.275	4.202
Proportion	.296	.226	.179	.249	.239	.212	.700

(c) The initial communality estimates for the six variables are given by (13.36) as .215, .225, .113, .255, .161, .248. With these substituted for the diagonal of **R**, the eigenvalues of $\mathbf{R} - \hat{\mathbf{\Psi}}$ are

Eigenvalue	.994	.569	.255	−.025	−.237	−.339
Proportion	.816	.468	.209	−.020	−.195	−.278
Cumulative	.816	1.284	1.493	1.473	1.278	1.000

The principal factor loadings and varimax rotation are as follows:

	Principal Component Loadings			Varimax Rotated Loadings			Communalities,
	f_1	f_2	f_3	f_1	f_2	f_3	$\hat{h}_i^2$
Variables							
y_1	.403	.312	.227	.030	**.536**	.151	.311
y_2	−.106	.569	−.100	−.288	.083	**.505**	.345
y_3	.343	−.139	−.197	**.413**	.060	.037	.176
y_4	.559	−.247	−.090	**.564**	.233	−.094	.381
y_5	−.286	−.246	.328	−.262	−.088	**−.417**	.250
y_6	.556	.089	.197	.258	**.537**	.003	.356

(d) The pattern of loadings is similar in parts (b) and (c), and the interpretation of the three factors would be the same.

13.13 Probe word data of Table 3.6:

	Principal Component Loadings		Varimax Rotated Loadings		Communalities,	Orthoblique Pattern Loadings	
	f_1	f_2	f_1	f_2	$\hat{h}_i^2$	f_1	f_2
Variables							
y_1	.817	−.157	**.732**	.395	.692	**.737**	.131
y_2	.838	−.336	**.861**	.271	.815	**.963**	−.092
y_3	.874	.288	.494	**.776**	.847	.248	**.734**
y_4	.838	−.308	**.844**	.292	.798	**.931**	−.057
y_5	.762	.547	.244	**.905**	.879	−.134	**1.023**
Variance	3.416	.614	2.294	1.736	4.031		
Proportion	.683	.123	.459	.347	.806		

The loadings for y_2 are similar to those for y_4 in all three sets of loadings. The reason for this can be seen in the correlation matrix

$$\mathbf{R} = \begin{pmatrix} 1.00 & .61 & .76 & .58 & .41 \\ .61 & 1.00 & .55 & .75 & .55 \\ .76 & .55 & 1.00 & .61 & .69 \\ .58 & .75 & .61 & 1.00 & .52 \\ .41 & .55 & .69 & .52 & 1.00 \end{pmatrix}.$$

The correlations of y_2 with y_1, y_3, and y_5 are very similar to the correlations of y_4 with y_1, y_3, and y_5.

CHAPTER 14

14.1 Multiply both sides of the inequality by 2 to obtain:

$$\begin{aligned} p^2 + p &\geq 2pm - 2m^2 + m^2 + m + 2p \\ \Rightarrow p^2 - 2pm + m^2 &\geq p + m \\ \Rightarrow (p-m)^2 &\geq p + m \end{aligned}$$

14.2 **(a)** Identified

(b) Identified

(c) Not identified

(d) Identified

14.3 **(a)** Identified

(b) 0

(c) Find $\text{var}(y_1), \text{var}(y_2), \text{var}(y_3), \text{cov}(y_1, y_2), \text{cov}(y_1, y_3)$, and $\text{cov}(y_2, y_3)$. For example, $\text{var}(y_1) = \phi + \psi_1$ and $\text{cov}(y_1, y_2) = \lambda_2\phi$.

(d) $\hat{\lambda}_2 = s_{23}/s_{13}$, $\hat{\lambda}_3 = s_{23}/s_{12}$, $\hat{\phi} = s_{12}s_{13}/s_{23}$, $\hat{\psi}_{11} = s_{11} - s_{12}s_{13}/s_{23}$, $\hat{\psi}_{22} = s_{22} - s_{12}s_{23}/s_{13}$, and $\hat{\psi}_{33} = s_{33} - s_{13}s_{23}/s_{12}$.

14.4 **(a)** $\chi^2 = 6.5811$, df $= 5$, p-value $= 0.2537$
Bentler's CFI = 0.9280
RMSEA = 0.1778
SRMR = 0.0711

(b) Using (14.25), $z_{21} = 2.5483$, $z_{31} = 2.8623$, $z_{41} = 2.5851$, and $z_{51} = 2.2674$. Since all exceed $z_{\alpha/2} = 1.96$ in absolute value, all loadings are significant and the model cannot be simplified.

(c) Same as part (a).

14.5 **(a)** $\chi^2 = 14.8718$, df $= 12$ (due to one error variance being constrained to equal zero), p-value $= 0.2485$
Bentler's CFI = 0.9824
RMSEA = 0.0908
SRMR = 0.0308

14.6 **(a)** $\chi^2 = 17.7127$, df $= 5$ (due to one error variance being constrained to equal zero), p-value $= 0.0033$
Bentler's CFI = 0.9401
RMSEA = 0.1690
SRMR = 0.0630

(b) $\chi^2 = 34.3303$, df $= 9$ (due to one error variance being constrained to equal zero), p-value < 0.0001
Bentler's CFI = 0.8806
RMSEA = 0.1778
SRMR = 0.0790

14.7 **(a)** $\chi^2 = 4.1224$, df $= 4$, p-value $= 0.3897$
Bentler's CFI = 0.9895
RMSEA = 0.0280
SRMR = 0.0608

(b)

Variable	Variance explained by the factors
Intelligence	20.5%
Form relations	undefined
Dynamometer	4.4%
Dotting	18.8%
Sensory motor coordination	2.3%
Perseveration	53.8%

14.8 $\chi^2 = 101.3717$, df $= 2$, p-value < 0.0001
Bentler's CFI = 0.9895

RMSEA = 1.0282
SRMR = 0.2237

CHAPTER 15

15.1 Adding and subtracting $\overline{x}$ and $\overline{y}$ in (15.2) (squared), we obtain

$$\begin{aligned} d^2(\mathbf{x},\mathbf{y}) &= \sum_{j=1}^{p}[(x_j-\overline{x})-(y_j-\overline{y})+(\overline{x}-\overline{y})]^2 \\ &= \sum_{j=1}^{p}(x_j-\overline{x})^2+\sum_{j=1}^{p}(y_i-\overline{y})^2+p(\overline{x}-\overline{y})^2 \\ &\quad -2\sum_{j=1}^{p}(x_j-\overline{x})(y_j-\overline{y}). \end{aligned}$$

The other two terms vanish because $\sum_j(x_j-\overline{x})=\sum_j(y_j-\overline{y})=0$. Substituting $v_x^2=\sum_{j=1}^{p}(x_j-\overline{x})^2$ and $v_y^2=\sum_{j=1}^{p}(y_j-\overline{y})^2$ and adding and subracting $-2\sqrt{v_x^2v_y^2}=-2v_xv_y$, we obtain

$$\begin{aligned} d^2(\mathbf{x},\mathbf{y}) &= v_x^2+v_y^2-2\sqrt{v_x^2v_y^2}+p(\overline{x}-\overline{y})^2+2v_xv_y \\ &\quad -2\sqrt{v_x^2v_y^2}\frac{\sum_{j=1}^{p}(x_j-\overline{x})(y_j-\overline{y})}{\sqrt{v_x^2v_y^2}} \\ &= (v_x-v_y)^2+p(\overline{x}-\overline{y})^2+2v_xv_y(1-r_{xy}) \end{aligned}$$

15.2 **(a)** Since $\overline{\mathbf{y}}_{AB}=\sum_{i=1}^{n_{AB}}\mathbf{y}_i/n_{AB}$, we have by (15.16),

$$\begin{aligned} \text{SSE}_{AB} &= \sum_{i=1}^{n_{AB}}(\mathbf{y}_i-\overline{\mathbf{y}}_{AB})'(\mathbf{y}_i-\overline{\mathbf{y}}_{AB}) \\ &= \sum_{i=1}^{n_{AB}}\mathbf{y}_i'\mathbf{y}_i-\sum_{i=1}^{n_{AB}}\mathbf{y}_i'\overline{\mathbf{y}}_{AB}-\sum_{i=1}^{n_{AB}}\overline{\mathbf{y}}_{AB}'\mathbf{y}_i \\ &\quad +\sum_{i=1}^{n_{AB}}\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} \\ &= \sum_{i=1}^{n_{AB}}\mathbf{y}_i'\mathbf{y}_i-n_{AB}\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB}-n_{AB}\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} \\ &\quad +n_{AB}\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} \\ &= \sum_{i=1}^{n_{AB}}\mathbf{y}_i'\mathbf{y}_i-n_{AB}\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB}. \end{aligned}$$

Similarly, $\text{SSE}_A = \sum_{i=1}^{n_A} \mathbf{y}_i'\mathbf{y}_i - n_A\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_A$ and $\text{SSE}_B = \sum_{i=1}^{n_B} \mathbf{y}_i'\mathbf{y}_i - n_B\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_B$. Now

$$\begin{aligned} n_{AB}\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} &= (n_A + n_B)\frac{(n_A\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B)'}{n_A + n_B}\frac{(n_A\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B)}{n_A + n_B} \\ &= \frac{n_A^2\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_A + n_An_B\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_B + n_An_B\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_B + n_B^2\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_B}{n_A + n_B}. \end{aligned}$$

Thus

$$\begin{aligned} \text{SSE}_{AB} - (\text{SSE}_A + \text{SSE}_B) &= \sum_{i=1}^{n_{AB}} \mathbf{y}_i'\mathbf{y}_i - \sum_{i=1}^{n_A} \mathbf{y}_i'\mathbf{y}_i - \sum_{i=1}^{n_B} \mathbf{y}_i'\mathbf{y}_i \\ &\quad + n_A\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_B - n_{AB}\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} \\ &= n_A\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_B - n_{AB}\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB}. \end{aligned}$$

Show that when the right side of (15.16) is expanded, it reduces to this same expression [see Problem 15.3(b) below].

(b) Multiplying out the right side of (15.16), we have

$$\begin{aligned} & n_A\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_A - n_A\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_{AB} - n_A\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_A + n_A\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} + n_B\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_B \\ &\quad - n_B\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_{AB} - n_B\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_B + n_B\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} \\ &= n_A\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_B - 2(n_A\overline{\mathbf{y}}_A' + n_B\overline{\mathbf{y}}_B')\overline{\mathbf{y}}_{AB} + (n_A + n_B)\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} \\ &= n_A\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_B - 2(n_A + n_B)\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} + (n_A + n_B)\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB} \\ &= n_A\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_B - (n_A + n_B)\overline{\mathbf{y}}_{AB}'\overline{\mathbf{y}}_{AB}. \end{aligned}$$

Substitute $\overline{\mathbf{y}}_{AB} = (n_A\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B)/(n_A + n_B)$.

15.3 **(a)** Complete linkage. From Table 15.2, we have

$$D(C, AB) = \tfrac{1}{2}D(C, A) + \tfrac{1}{2}D(C, B) + \tfrac{1}{2}|D(C, A) - D(C, B)| \quad (1)$$

If $D(C, A) > D(C, B)$, then $|D(C, A) - D(C, B)| = D(C, A) - D(C, B)$, and (1) becomes $D(C, AB) = D(C, A)$. If $D(C, A) > D(C, B)$, then $|D(C, A) - D(C, B)| = D(C, B) - D(C, A)$ and equation (1) becomes $D(C, AB) = D(C, B)$. Thus equation (1) can be written as $D(C, AB) = \max[D(C, A), D(C, B)]$, which is equivalent to (15.9), the definition of distance for the complete linkage method.

(b) Average linkage. From Table 15.2, we have

$$D(C, AB) = \frac{n_A}{n_A + n_B}D(C, A) + \frac{n_B}{n_A + n_B}D(C, B). \quad (2)$$

By (15.10) equation (2) can be written as

$$\begin{aligned} D(C, AB) &= \frac{n_A}{n_A + n_B} \cdot \frac{1}{n_C n_A} \sum_{i=1}^{n_C} \sum_{j=1}^{n_A} d(\mathbf{y}_i, \mathbf{y}_j) \\ &\quad + \frac{n_B}{n_A + n_B} \cdot \frac{1}{n_C n_B} \sum_{i=1}^{n_C} \sum_{j=1}^{n_B} d(\mathbf{y}_i, \mathbf{y}_j) \\ &= \frac{1}{n_C(n_A + n_B)} \sum_{i=1}^{n_C} \left[\sum_{j=1}^{n_A} d(\mathbf{y}_i, \mathbf{y}_j) + \sum_{j=1}^{n_B} d(\mathbf{y}_i, \mathbf{y}_j) \right] \\ &= \frac{1}{n_C n_{AB}} \sum_{i=1}^{n_C} \sum_{j=1}^{n_{AB}} d(\mathbf{y}_i, \mathbf{y}_j), \end{aligned}$$

which, by (15.10), is the definition of distance for the average linkage method.

(c) Substitute $\overline{\mathbf{y}}_{AB} = (n_A\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B)/(n_A + n_B)$ in the left side of (15.40) in the statement of Problem 15.3(c) and multiply to obtain

$$\begin{aligned} &\overline{\mathbf{y}}_C'\overline{\mathbf{y}}_C - \frac{2n_A\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_C}{n_A + n_B} + \frac{2n_An_B\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_B}{(n_A + n_B)^2} - \frac{2n_B\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_C}{n_A + n_B} \\ &\quad + \frac{n_A^2\overline{\mathbf{y}}_A'\overline{\mathbf{y}}_A}{(n_A + n_B)^2} + \frac{n_B^2\overline{\mathbf{y}}_B'\overline{\mathbf{y}}_B}{(n_A + n_B)^2}. \end{aligned}$$

Similarly, multiply on the right side of (15.40) to obtain the same result.

(d) Using $n_A = n_B$ in $\overline{\mathbf{y}}_{AB} = (n_A\overline{\mathbf{y}}_A + n_B\overline{\mathbf{y}}_B)/(n_A + n_B)$ in (15.12), we obtain $\mathbf{m}_{AB} = \frac{1}{2}(\overline{\mathbf{y}}_A + \overline{\mathbf{y}}_B)$ in (15.13). Then (15.40) [see part (c)] becomes

$$\begin{aligned} (\overline{\mathbf{y}}_C - \mathbf{m}_{AB})'(\overline{\mathbf{y}}_C - \mathbf{m}_{AB}) &= \tfrac{1}{2}(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_A)'(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_A) \\ &\quad + \tfrac{1}{2}(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_B)'(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_B) - \tfrac{1}{4}(\overline{\mathbf{y}}_A - \overline{\mathbf{y}}_B)'(\overline{\mathbf{y}}_A - \overline{\mathbf{y}}_B), \end{aligned}$$

which matches the parameter values for the median method in Table 15.2.

(e) By (15.19),

$$(\overline{\mathbf{y}}_A - \overline{\mathbf{y}}_B)'(\overline{\mathbf{y}}_A - \overline{\mathbf{y}}_B) = \frac{n_A + n_B}{n_An_B} I_{AB},$$

and we have analogous expressions for $(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_{AB})'(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_{AB})$, $(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_A)'(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_A)$, and $(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_B)'(\overline{\mathbf{y}}_C - \overline{\mathbf{y}}_B)$. Then (15.40) in part (c) becomes

$$\begin{aligned}
\frac{n_C + n_{AB}}{n_C n_{AB}} I_{C(AB)} &= \left(\frac{n_A}{n_A + n_B}\right)\left(\frac{n_C + n_A}{n_C n_A}\right) I_{CA} \\
&\quad + \left(\frac{n_B}{n_A + n_B}\right)\left(\frac{n_c + n_B}{n_C n_B}\right) I_{CB} \\
&\quad - \left[\frac{n_A n_B}{(n_A + n_B)^2}\right]\left(\frac{n_A + n_B}{n_A n_B}\right) I_{AB} \\
&= \frac{n_A + n_B}{n_C n_{AB}} I_{AC} + \frac{n_B + n_C}{n_C n_{AB}} I_{CB} - \frac{1}{n_{AB}} I_{AB}.
\end{aligned}$$

Solve for $I_{C(AB)}$.

15.4 If $\gamma = 0$, then (15.20) becomes

$$D(C, AB) = \alpha_A D(C, A) + \alpha_B D(C, B) + \beta D(A, B). \qquad (1)$$

By (15.25), we have $D(A, C) > D(A, B)$ and $D(B, C) > D(A, B)$. Thus, replacing $D(C, A)$ and $D(C, B)$ in equation (1) by $D(A, B)$, we obtain

$$D(C, AB) > \alpha_A D(A, B) + \alpha_B D(A, B) + \beta D(A, B),$$

which is equivalent to (15.26).

15.5 **(a)** $\overline{\mathbf{v}}_{..} = \frac{1}{gn} \sum_{i=1}^{g} \sum_{j=1}^{n} \mathbf{v}_{ij} = \frac{1}{gn} \sum_{i=1}^{g} \sum_{j=1}^{n} (\mathbf{A}\mathbf{y}_{ij} + \mathbf{b})$

$$= \frac{1}{gn} \left(\mathbf{A} \sum_{ij} \mathbf{y}_{ij} + gn\mathbf{b}\right) = \mathbf{A}\left(\frac{1}{gn} \sum_{ij} \mathbf{y}_{ij}\right) + \mathbf{b} = \mathbf{A}\overline{\mathbf{y}}_{..} + \mathbf{b}$$

Show similarly that $\overline{\mathbf{v}}_{i.} = \mathbf{A}\overline{\mathbf{y}}_{i.} + \mathbf{b}$. Then by (6.9), we have

$$\begin{aligned}
\mathbf{H}_v &= n \sum_{i=1}^{g} (\overline{\mathbf{v}}_{i.} - \overline{\mathbf{v}}_{..})(\overline{\mathbf{v}}_{i.} - \overline{\mathbf{v}}_{..})' \\
&= n \sum_{i} [\mathbf{A}\overline{\mathbf{y}}_{i.} + \mathbf{b} - (\mathbf{A}\overline{\mathbf{y}}_{..} + \mathbf{b})][\mathbf{A}\overline{\mathbf{y}}_{i.} + \mathbf{b} - (\mathbf{A}\overline{\mathbf{y}}_{..} + \mathbf{b})]' \\
&= n \sum_{i} (\mathbf{A}\overline{\mathbf{y}}_{i.} - \mathbf{A}\overline{\mathbf{y}}_{..})(\mathbf{A}\overline{\mathbf{y}}_{i.} - \mathbf{A}\overline{\mathbf{y}}_{..})' \\
&= n \sum_{i} \mathbf{A}(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})'\mathbf{A}' \qquad \text{[by (2.27)]} \\
&= n\mathbf{A} \left[\sum_{i} (\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})(\overline{\mathbf{y}}_{i.} - \overline{\mathbf{y}}_{..})'\right] \mathbf{A}' \qquad \text{[by (2.45)]} \\
&= \mathbf{A}\mathbf{H}_y\mathbf{A}'.
\end{aligned}$$

Show similarly that $\mathbf{E}_v = \mathbf{A}\mathbf{E}_y\mathbf{A}'$.

(b) $\text{tr}(\mathbf{E}_v) = \text{tr}(\mathbf{A}\mathbf{E}_y\mathbf{A}') = \text{tr}(\mathbf{A}'\mathbf{A}\mathbf{E}_y) \neq \text{tr}(\mathbf{E}_y)$

(c) $|\mathbf{E}_v| = |\mathbf{A}\mathbf{E}_y\mathbf{A}'| = |\mathbf{A}|\,|\mathbf{E}_y|\,|\mathbf{A}'| = |\mathbf{A}|^2|\mathbf{E}_y| = c|\mathbf{E}_y|$, where $c > 0$. Thus minimizing $|\mathbf{E}_v|$ is equivalent to minimizing $|\mathbf{E}_y|$.

(d)
$$\begin{aligned}\text{tr}(\mathbf{E}_v^{-1}\mathbf{H}_v) &= \text{tr}[(\mathbf{A}\mathbf{E}_y\mathbf{A}')^{-1}(\mathbf{A}\mathbf{H}_y\mathbf{A}')] \\ &= \text{tr}[(\mathbf{A}')^{-1}\mathbf{E}_y^{-1}\mathbf{A}^{-1}\mathbf{A}\mathbf{H}_y\mathbf{A}')] \\ &= \text{tr}[(\mathbf{A}')^{-1}\mathbf{E}_y^{-1}\mathbf{H}_y\mathbf{A}'] \\ &= \text{tr}[\mathbf{A}'(\mathbf{A}')^{-1}\mathbf{E}_y^{-1}\mathbf{H}_y] \\ &= \text{tr}(\mathbf{E}_y^{-1}\mathbf{H}_y)\end{aligned}$$

15.6 There are p parameters in each $\boldsymbol{\mu}_i$, $\frac{1}{2}p(p+1)$ unique parameters in each $\boldsymbol{\Sigma}_i$, and $g-1$ unique parameters α_i. Thus the total number is

$$\begin{aligned}gp + g[\tfrac{1}{2}p(p+1)] + g - 1 &= g[p + \tfrac{1}{2}p(p+1) + 1] - 1 \\ &= \tfrac{1}{2}g[2p + p^2 + p + 2] - 1 \\ &= \tfrac{1}{2}g(3p + p^2 + 2) - 1 \\ &= \tfrac{1}{2}g(p+1)(p+2) - 1.\end{aligned}$$

15.7 **(a)** The two-cluster solution from single linkage puts boy number 20 in one cluster and the other 19 boys in the other cluster.

(b)-(d) Based on the change in distance, average linkage and the other cluster solutions in parts (c) and (d) clearly indicate two clusters. These solutions generally agree and also correspond to a division into two groups seen in the first principal component in Figure 12.5. The separation of the three apparent outliers from the other 17 observations is less pronounced in the cluster analyses than in Figure 12.5. Note that the scale of the second component in Figure 12.5 is much larger than that of the first component, so the separation of points 9, 12, and 20 from the rest is not as large as it appears in the figure. Of the methods in parts (b), (c), and (d), only flexible beta with $\beta = -.50$ and $-.75$ place points 9, 12, and 20 together in one cluster. All others place 9 and 12 in one of the clusters and 20 in the other.

15.8 **(a)** The distance between centroids of the two clusters is $\sqrt{2994.9} = 54.7$.

(b) From the dendrogram produced by the average linkage method, the largest change in distance corresponds to a two-cluster solution.

(c) The discriminant function completely separates the two clusters, with no overlap.

15.9 **(a)** Observation 22 seems to be an outlier, because it forms its own cluster in both the single linkage and average linkage methods. The cluster consisting of observations 2, 21, 24, 26 and 30 is the same in all six methods.

(b) The discriminant function completely separates the two clusters, with no overlap.

15.10 **(a)** The following five clusters were found using as seeds the five observations that are mutually farthest apart.

Cluster	1	2	3	4	5
Observation(s)	9,15,16, 18,19	1,2,3, 4,5,17	6,7, 8,20	10,11, 12,13	14

In the plot of the first two discriminant functions, observation 14 is relatively far removed from the rest. Clusters 1, 2, and 3 are somewhat closer to each other.

(b) The following five clusters were found using as seeds the first five observations.

Cluster	1	2	3	4	5
Observation(s)	1,3,4	2	5,17, 18,19	6,7,8, 15,16,20	9,10,11, 12,13,14

The plot of the first two discriminant functions shows a pattern different from that in part (a).

(c) The following five clusters were found using as seeds the centroids of the five-cluster solution resulting from Ward's method.

Cluster	1	2	3	4	5
Observation(s)	6,7,8,15, 16,20	5,9,17, 18,19	10,11, 12,13	1,2, 3,4	14

The plot of the first two discriminant functions shows a pattern similar to that found in part (a), with observation 14 isolated.The dendrogram shows that Ward's method gives the same five-cluster solution as the k-means result.

(d) The following five clusters were found using the k-means method with seeds equal to the centroids of the five clusters from average linkage.

Cluster	1	2	3	4	5
Observation(s)	6,7,8,15, 16,20	1,2,3,4,5, 17,18,19	10,11, 12,13	9	14

The plot of the first two discriminant functions shows a pattern somewhat similar to that in part (a). In the dendrogram for average linkage, observations 9 and 14 are isolated clusters in the five-cluster solution, which is identical to the five-cluster solution using k-means clustering with these seeds.

(e) Observation 14 does not appear as an outlier in the plot of the first two principal compoments, but it does show up as an outlier in the plot of the second and third components. The solutions found in parts (a) and (c) seem to agree most with the principal component plots. This suggests that a different number of initial cluster seeds be used.

(f) The two clustering solutions are identical. The results are given below.

Cluster	1	2	3
Observation(s)	6,7,8,15,16,20	9,10,11,12,13,14	1,2,3,4,5,17,18,19

(g) The clustering solution is identical to that found in part (e), which indicates that the three-cluster solution is appropriate.

15.11 The numbers of clusters obtained from the indicated combinations of k and r are shown in the table below. Note that for each pair of values of k and r, the value of r was increased if necessary for each point until k points were included in the sphere.

k/r	.2	.4	.6	.8	1.0	1.2	1.4	1.6	1.8	2.0
2	10	10	10	10	8	6	4	3	3	2
3	5	5	5	5	5	3	2	2	2	2
4	2	2	2	2	2	2	2	2	2	2
5	1	1	1	1	1	1	1	1	1	1

The maximum value of k that yields a two-cluster solution is 4.

15.12 **(a)** The numbers of clusters obtained from the initial combinations of k and r are shown in the table below. The value of r was variable, as noted in Problem 15.11 above.

k/r	.2	.4	.6	.8	1.0	1.2	1.4	1.6	1.8	2.0
2	3	3	3	3	3	3	3	3	3	3
3	2	2	2	2	2	2	2	2	2	2
4	1	1	1	1	1	1	1	1	1	1

(b) The plot of the first two discriminant functions for $k = 2$ and $r = 1$ shows the three clusters to be well separated.

(c) The plot of the first two principal components shows the same groupings as in the plot in part (b).

(d) The plot of the discriminant function shows wide separation of the two clusters. The clusters do not overlap. The three-cluster solution found in part (b) is given below

Cluster 1	Cluster 2	Cluster 3
Harpers	Rosemaund	Cambridge
Morley	Terrington	Cockle Park
Myerscough	Headley	
Sparsholt	Seale-Hayne	
Sutton Bonington		
Wye		

The two-cluster solution found in part (d) merges clusters two and three of part (b).

CHAPTER 16

16.1 $$\mathbf{B} = \left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{A}\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)$$

$$= \mathbf{A} - \frac{1}{n}\mathbf{AJ} - \frac{1}{n}\mathbf{JA} + \frac{1}{n^2}\mathbf{JAJ} \tag{1}$$

By (2.38),

$$\frac{1}{n}\mathbf{Aj} = \frac{1}{n}\begin{pmatrix} \sum_j a_{1j} \\ \sum_j a_{2j} \\ \vdots \\ \sum_j a_{nj} \end{pmatrix} = \begin{pmatrix} \overline{a}_{1.} \\ \overline{a}_{2.} \\ \vdots \\ \overline{a}_{n.} \end{pmatrix}. \tag{2}$$

Hence,

$$\begin{aligned} \tfrac{1}{n}\mathbf{AJ} &= \tfrac{1}{n}\mathbf{A}(\mathbf{j},\mathbf{j},\ldots,\mathbf{j}) = (\tfrac{1}{n}\mathbf{Aj},\ldots,\tfrac{1}{n}\mathbf{Aj}) \\ &= \begin{pmatrix} \overline{a}_{1.} & \cdots & \overline{a}_{1.} \\ \overline{a}_{2.} & \cdots & \overline{a}_{2.} \\ \vdots & & \vdots \\ \overline{a}_{n.} & \cdots & \overline{a}_{n.} \end{pmatrix}. \end{aligned}$$

Show that

$$\frac{1}{n}\mathbf{JA} = \begin{pmatrix} \overline{a}_{.1} & \overline{a}_{.2} & \cdots & \overline{a}_{.n} \\ \overline{a}_{.1} & \overline{a}_{.2} & \cdots & \overline{a}_{.n} \\ \vdots & \vdots & & \vdots \\ \overline{a}_{.1} & \overline{a}_{.2} & \cdots & \overline{a}_{.n} \end{pmatrix}.$$

Using equation (2), we obtain

$$\frac{1}{n^2}\mathbf{j}'\mathbf{Aj} = \frac{1}{n^2}(1,1,\ldots,1)\begin{pmatrix} \sum_j a_{1j} \\ \sum_j a_{2j} \\ \vdots \\ \sum_j a_{nj} \end{pmatrix} = \frac{1}{n^2}\sum_{ij} a_{ij} = \overline{a}_{..}.$$

By (3.63),

$$\frac{1}{n^2}\mathbf{JAJ} = \frac{1}{n^2}\begin{pmatrix} \mathbf{j}'\mathbf{Aj} & \cdots & \mathbf{j}'\mathbf{Aj} \\ \vdots & & \vdots \\ \mathbf{j}'\mathbf{Aj} & \cdots & \mathbf{j}'\mathbf{Aj} \end{pmatrix} = \begin{pmatrix} \overline{a}_{..} & \cdots & \overline{a}_{..} \\ \vdots & & \vdots \\ \overline{a}_{..} & \cdots & \overline{a}_{..} \end{pmatrix}.$$

Hence the ijth element of equation (1) is $b_{ij} = a_{ij} - \overline{a}_{i.} - \overline{a}_{.j} + \overline{a}_{..}$.

16.2 **(a)** (Seber 1984, pp. 236-237) The elements of $\mathbf{B} = (b_{ij})$ are defined as $b_{ij} = a_{ij} - \overline{a}_{i.} - \overline{a}_{.j} + \overline{a}_{..}$, where $a_{ij} = -\frac{1}{2}\delta_{ij}^2$. Thus

$$\begin{aligned} -2a_{ij} &= \delta_{ij}^2 = (\mathbf{z}_i - \mathbf{z}_j)'(\mathbf{z}_i - \mathbf{z}_j) \\ &= \mathbf{z}_i'\mathbf{z}_i + \mathbf{z}_j'\mathbf{z}_j - 2\mathbf{z}_i'\mathbf{z}_j. \end{aligned}$$

Then

$$\begin{aligned} -2\overline{a}_{i.} &= \frac{1}{n}\sum_{j=1}^{n}(-2a_{ij}) = \frac{1}{n}\sum_j(\mathbf{z}_i'\mathbf{z}_i + \mathbf{z}_j'\mathbf{z}_j - 2\mathbf{z}_i'\mathbf{z}_j) \\ &= \mathbf{z}_i'\mathbf{z}_i + \frac{1}{n}\sum_j \mathbf{z}_j'\mathbf{z}_j - \frac{2}{n}\mathbf{z}_i'\sum_j \mathbf{z}_j \\ &= \mathbf{z}_i'\mathbf{z}_i + \frac{1}{n}\sum_j \mathbf{z}_j'\mathbf{z}_j - 2\mathbf{z}_i'\overline{\mathbf{z}}. \end{aligned}$$

Similarly, show that

$$\begin{aligned} -2\overline{a}_{.j} &= \mathbf{z}_j'\mathbf{z}_j + \frac{1}{n}\sum_i \mathbf{z}_i'\mathbf{z}_i - 2\overline{\mathbf{z}}'\mathbf{z}_j, \\ -2\overline{a}_{..} &= \frac{2}{n}\sum_i \mathbf{z}_i'\mathbf{z}_i - 2\overline{\mathbf{z}}'\overline{\mathbf{z}}. \end{aligned}$$

Solve for $a_{ij}, \overline{a}_{i.}, \overline{a}_{.j}$, and $\overline{a}_{..}$ and substitute into $b_{ij} = a_{ij} - \overline{a}_{i.} - \overline{a}_{.j} + \overline{a}_{..}$ to obtain $b_{ij} = \mathbf{z}_i'\mathbf{z}_j - \mathbf{z}_i'\overline{\mathbf{z}} - \overline{\mathbf{z}}'\mathbf{z}_j + \overline{\mathbf{z}}'\overline{\mathbf{z}}$, which can be factored as $b_{ij} = (\mathbf{z}_i - \overline{\mathbf{z}})'(\mathbf{z}_j - \overline{\mathbf{z}})$. Hence

$$\begin{aligned} \mathbf{B} &= \begin{pmatrix} (\mathbf{z}_1 - \overline{\mathbf{z}})'(\mathbf{z}_1 - \overline{\mathbf{z}}) & \cdots & (\mathbf{z}_1 - \overline{\mathbf{z}})'(\mathbf{z}_n - \overline{\mathbf{z}}) \\ \vdots & & \vdots \\ (\mathbf{z}_n - \overline{\mathbf{z}})'(\mathbf{z}_1 - \overline{\mathbf{z}}) & \cdots & (\mathbf{z}_n - \overline{\mathbf{z}})'(\mathbf{z}_n - \overline{\mathbf{z}}) \end{pmatrix} \\ &= \begin{pmatrix} (\mathbf{z}_1 - \overline{\mathbf{z}})' \\ \vdots \\ (\mathbf{z}_n - \overline{\mathbf{z}})' \end{pmatrix}(\mathbf{z}_1 - \overline{\mathbf{z}}, \ldots, \mathbf{z}_n - \overline{\mathbf{z}}) \\ &= \mathbf{Z}_c\mathbf{Z}_c' \quad \text{[see (10.13)]}. \end{aligned}$$

Thus $\mathbf{B}$ is positive semidefinite (see Section 2.7).

(b) If $\mathbf{B}$ is positive semidefinite of rank q, then by (2.109) and Section 2.11.4, $\mathbf{B}$ can be expressed in the form $\mathbf{B} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}'$, where $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ is an orthogonal matrix of eigenvectors of $\mathbf{B}$, and $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues, q of which are positive, with the rest equal to zero. Letting $\mathbf{\Lambda}_1$ be the $q \times q$ upper left hand block of $\mathbf{\Lambda}$ with positive eigenvalues and $\mathbf{V}_1 = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_q)$ be the $n \times q$ matrix with the corresponding eigenvectors, we can write $\mathbf{B} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}'$ as

$$\begin{aligned}
\mathbf{B} &= (\mathbf{V}_1, \mathbf{V}_2)\begin{pmatrix} \mathbf{\Lambda}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}\begin{pmatrix} \mathbf{V}_1' \\ \mathbf{V}_2' \end{pmatrix} \\
&= \mathbf{V}_1\mathbf{\Lambda}_1\mathbf{V}_1' = \mathbf{V}_1\mathbf{\Lambda}_1^{1/2}\mathbf{\Lambda}_1^{1/2}\mathbf{V}_1' \\
&= \mathbf{Z}\mathbf{Z}',
\end{aligned} \tag{1}$$

where the $n \times q$ matrix $\mathbf{Z}$ is

$$\begin{aligned}
\mathbf{Z} &= \mathbf{V}_1\mathbf{\Lambda}_1^{1/2} = (\sqrt{\lambda_1}\mathbf{v}_1, \sqrt{\lambda_2}\mathbf{v}_2, \ldots, \sqrt{\lambda_q}\mathbf{v}_q) \\
&= \begin{pmatrix} \mathbf{z}_1' \\ \mathbf{z}_2' \\ \vdots \\ \mathbf{z}_n' \end{pmatrix}.
\end{aligned}$$

To show that $(\mathbf{z}_i - \mathbf{z}_j)'(\mathbf{z}_i - \mathbf{z}_j)$ is equal to δ_{ij}^2, we can proceed as follows:

$$(\mathbf{z}_i - \mathbf{z}_j)'(\mathbf{z}_i - \mathbf{z}_j) = \mathbf{z}_i'\mathbf{z}_i + \mathbf{z}_j'\mathbf{z}_j - 2\mathbf{z}_i'\mathbf{z}_j. \tag{2}$$

By equation (1), we have

$$\begin{aligned}
\mathbf{B} = \mathbf{Z}\mathbf{Z}' &= \begin{pmatrix} \mathbf{z}_1' \\ \mathbf{z}_2' \\ \vdots \\ \mathbf{z}_n' \end{pmatrix}(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n) \\
&= \begin{pmatrix} \mathbf{z}_1'\mathbf{z}_1 & \mathbf{z}_1'\mathbf{z}_2 & \cdots & \mathbf{z}_1'\mathbf{z}_n \\ \mathbf{z}_2'\mathbf{z}_1 & \mathbf{z}_2'\mathbf{z}_2 & \cdots & \mathbf{z}_2'\mathbf{z}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{z}_n'\mathbf{z}_1 & \mathbf{z}_n'\mathbf{z}_2 & \cdots & \mathbf{z}_n'\mathbf{z}_n \end{pmatrix}.
\end{aligned}$$

Hence equation (2) becomes

$$\begin{aligned}
(\mathbf{z}_i - \mathbf{z}_j)'(\mathbf{z}_i - \mathbf{z}_j) &= \mathbf{z}_i'\mathbf{z}_i + \mathbf{z}_j'\mathbf{z}_j - 2\mathbf{z}_i'\mathbf{z}_j \\
&= b_{ii} + b_{jj} - 2b_{ij}.
\end{aligned} \tag{3}$$

Show that substituting $b_{ij} = a_{ij} - \overline{a}_{i.} - \overline{a}_{.j} + \overline{a}_{..}$ into equation (3) leads to

$$(\mathbf{z}_i - \mathbf{z}_j)'(\mathbf{z}_i - \mathbf{z}_j) = a_{ii} + a_{jj} - 2a_{ij} + \overline{a}_{i.} - \overline{a}_{.i} + \overline{a}_{.j} - \overline{a}_{j.}.$$

Show that the symmetry of $\mathbf{A}$ implies $\bar{a}_{i.} = \bar{a}_{.i}$ and $\bar{a}_{.j} - \bar{a}_{j.}$. Hence,

$$(\mathbf{z}_i - \mathbf{z}_j)'(\mathbf{z}_i - \mathbf{z}_j) = a_{ii} + a_{jj} - 2a_{ij} = -2a_{ij} = \delta_{ij}^2,$$

since $a_{ii} = -\frac{1}{2}\delta_{ii}^2 = 0$ and $-2a_{ij} = \delta_{ij}^2$.

16.3 (a)
$$\begin{aligned}\mathbf{r} &= \sum_{j=1}^{b} p_{.j}\mathbf{c}_j = \sum_{j=1}^{b} p_{.j}\left(\frac{p_{1j}}{p_{.j}}, \frac{p_{2j}}{p_{.j}}, \ldots, \frac{p_{aj}}{p_{.j}}\right)' \\ &= \textstyle\sum_{j=1}^{b}(p_{1j}, p_{2j}, \ldots, p_{aj})' \quad \text{[by (2.61)]} \\ &= (\textstyle\sum_j p_{1j}, \sum_j p_{2j}, \ldots, \sum_j p_{aj})' \\ &= (p_{1.}, p_{2.}, \ldots, p_{a.})'\end{aligned}$$

(b)
$$\begin{aligned}\mathbf{c}' &= \sum_{i=1}^{a} p_{i.}\mathbf{r}_i' = \sum_{i=1}^{a} p_{i.}\left(\frac{p_{i1}}{p_{i.}}, \frac{p_{i2}}{p_{i.}}, \ldots, \frac{p_{ib}}{p_{i.}}\right) \\ &= \sum_{i=1}^{a}(p_{i1}, p_{i2}, \ldots, p_{ib}) \quad \text{[by (2.61)]} \\ &= (\textstyle\sum_i p_{i1}, \sum_i p_{i2}, \ldots, \sum_i p_{ib}) \\ &= (p_{.1}, p_{.2}, \ldots, p_{.b})\end{aligned}$$

16.4 $\mathbf{j}'\mathbf{r} = \sum_{i=1}^{a} p_{i.} = \sum_{i=1}^{a}\sum_{j=1}^{b} p_{ij} = \sum_{ij} n_{ij}/n = n/n = 1,$
$\mathbf{c}'\mathbf{j} = \sum_{j=1}^{b} p_{.j} = \sum_j n_{.j}/n = \sum_j \sum_i n_{ij}/n = n/n$

16.5 By (16.8), (16.9), and (16.10), $p_{ij} = n_{ij}/n, p_{i.} = n_{i.}/n$, and $p_{.j} = n_{.j}/n$. Substituting these into (16.25), we obtain

$$\begin{aligned}\chi^2 &= \sum_{ij} \frac{n\left(\frac{n_{ij}}{n} - \frac{n_{i.}n_{.j}}{n^2}\right)^2}{\frac{n_{i.}n_{.j}}{n^2}} = \sum_{ij} \frac{n\left[\frac{1}{n}(n_{ij} - \frac{n_{i.}n_{.j}}{n})\right]^2}{\frac{n_{i.}n_{.j}}{n^2}} \\ &= \sum_{ij} \frac{\frac{n}{n^2}(n_{ij} - \frac{n_{i.}n_{.j}}{n})^2}{\frac{n_{i.}n_{.j}}{n^2}} = \sum_{ij} \frac{(n_{ij} - \frac{n_{i.}n_{.j}}{n})^2}{\frac{n_{i.}n_{.j}}{n}}.\end{aligned}$$

16.6 (a) Multiplying numerator and denominator of (16.25) by $p_{i.}$, we obtain

$$\begin{aligned}\chi^2 &= \sum_i n \sum_j \frac{p_{i.}}{p_{i.}^2 p_{.j}}(p_{ij} - p_{i.}p_{.j})^2 \\ &= \sum_i n\,p_{i.} \sum_j \frac{1}{p_{.j}}\left[\frac{1}{p_{i.}}(p_{ij} - p_{i.}p_{.j})\right]^2 \\ &= \sum_i np_{i.} \sum_j \left(\frac{p_{ij}}{p_{i.}} - p_{.j}\right)^2 / p_{.j}.\end{aligned}$$

16.7 (a) By (16.29), (16.10), (16.12), and (16.18), we obtain

$$\chi^2 = \sum_i n p_{i.}(\mathbf{r}_i - \mathbf{c})'\mathbf{D}_c^{-1}(\mathbf{r}_i - \mathbf{c})$$

$$= \sum_i n p_{i.}\left(\frac{p_{i1}}{p_{i.}} - p_{.1}, \ldots, \frac{p_{ib}}{p_{i.}} - p_{.b}\right)\begin{pmatrix} p_{.1} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & p_{.b} \end{pmatrix}^{-1}\begin{pmatrix} \frac{p_{i1}}{p_{i.}} - p_{.1} \\ \vdots \\ \frac{p_{ib}}{p_{i.}} - p_{.b} \end{pmatrix}$$

$$= \sum_i n p_{i.}\left(\frac{\frac{p_{i1}}{p_{i.}} - p_{.1}}{p_{.1}}, \ldots, \frac{\frac{p_{ib}}{p_{i.}} - p_{.b}}{p_{.b}}\right)\begin{pmatrix} \frac{p_{i1}}{p_{i.}} - p_{.1} \\ \vdots \\ \frac{p_{ib}}{p_{i.}} - p_{.b} \end{pmatrix}.$$

16.8 (a) By (16.9) $\mathbf{r} = \mathbf{Pj}$. Then $\mathbf{D}_r^{-1}\mathbf{r} = \mathbf{D}_r^{-1}\mathbf{Pj} = \mathbf{Rj}$ by (16.15). By (16.13) $\mathbf{r}_i'\mathbf{j} = 1$, and therefore $\mathbf{Rj} = \mathbf{j}$. Now

$$\mathbf{D}_r^{-1}(\mathbf{P} - \mathbf{rc}') = \mathbf{D}_r^{-1}\mathbf{P} - \mathbf{D}_r^{-1}\mathbf{rc}' = \mathbf{R} - \mathbf{Rjc}' = \mathbf{R} - \mathbf{jc}'.$$

16.9 By (16.49), $\mathbf{z}_i' = \mathbf{y}_i'\mathbf{A}$ (ignoring the centering on $\mathbf{y}_i$). Thus the squared Euclidean distance can be written as

$$\begin{aligned}(\mathbf{z}_i - \mathbf{z}_k)'(\mathbf{z}_i - \mathbf{z}_k) &= (\mathbf{z}_i' - \mathbf{z}_k')(\mathbf{z}_i - \mathbf{z}_k) \\ &= (\mathbf{y}_i'\mathbf{A} - \mathbf{y}_k'\mathbf{A})(\mathbf{A}'\mathbf{y}_i - \mathbf{A}'\mathbf{y}_k) \\ &= (\mathbf{y}_i' - \mathbf{y}_k')\mathbf{AA}'(\mathbf{y}_i - \mathbf{y}_k) \\ &= (\mathbf{y}_i - \mathbf{y}_k)'(\mathbf{y}_i - \mathbf{y}_k),\end{aligned}$$

since $\mathbf{A}$ is orthogonal.

16.10 (a) From $\mathbf{Y}_c\mathbf{V} = \mathbf{U}\boldsymbol{\Lambda}$ in (16.55), we have $\mathbf{Y}_c\mathbf{V}\boldsymbol{\Lambda}^{-1} = \mathbf{U}$. Then

$$\begin{aligned}\mathbf{UU}' &= \mathbf{Y}_c\mathbf{V}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{V}'\mathbf{Y}_c' \\ &= \mathbf{Y}_c\mathbf{V}(\boldsymbol{\Lambda}^{-1})^2\mathbf{V}'\mathbf{Y}_c'. \end{aligned} \tag{1}$$

Since $(\boldsymbol{\Lambda}^{-1})^2 = \text{ diag }(1/\lambda_1^2, 1/\lambda_2^2, \ldots, 1/\lambda_p^2)$, where the λ_i^2's are eivenvalues of $\mathbf{Y}_c'\mathbf{Y}_c$, the matrix $(\boldsymbol{\Lambda}^{-1})^2$ contains eigenvalues of $(\mathbf{Y}_c'\mathbf{Y}_c)^{-1} = [(n-1)\mathbf{S}]^{-1} = \mathbf{S}^{-1}/(n-1)$ [see (2.115) and (2.116)]. The matrix $\mathbf{V}$ contains eigenvectors of $\mathbf{Y}_c'\mathbf{Y}_c$ and thereby of $(\mathbf{Y}_c'\mathbf{Y}_c)^{-1}$ (see Section 2.11.9). Hence we recognize $\mathbf{V}(\boldsymbol{\Lambda}^{-1})^2\mathbf{V}'$ as the spectral decomposition of $(\mathbf{Y}_c'\mathbf{Y}_c)^{-1}$ [see (2.109), (2.115), and (2.116)]. Therefore, equation (1) can be written as

$$\begin{aligned}\mathbf{UU}' &= \mathbf{Y}_c\mathbf{V}(\boldsymbol{\Lambda}^{-1})^2\mathbf{V}'\mathbf{Y}_c' = \mathbf{Y}_c(\mathbf{Y}_c'\mathbf{Y}_c)^{-1}\mathbf{Y}_c' \\ &= \mathbf{Y}_c\mathbf{S}^{-1}\mathbf{Y}_c'/(n-1).\end{aligned}$$

(b) If $\mathbf{H} = \mathbf{V}\boldsymbol{\Lambda}$, then $\mathbf{H}\mathbf{H}' = \mathbf{V}\boldsymbol{\Lambda}\boldsymbol{\Lambda}\mathbf{V}' = \mathbf{V}\boldsymbol{\Lambda}^2\mathbf{V}'$. The diagonal matrix $\boldsymbol{\Lambda}^2$ contains the eigenvalues λ_i^2 of the matrix $\mathbf{Y}_c'\mathbf{Y}_c$. Thus by (2.115), $\mathbf{V}\boldsymbol{\Lambda}^2\mathbf{V}'$ is the spectral decomposition of $\mathbf{Y}_c'\mathbf{Y}_c$, and

$$\mathbf{H}\mathbf{H}' = \mathbf{V}\boldsymbol{\Lambda}^2\mathbf{V}' = \mathbf{Y}_c'\mathbf{Y}_c = (n-1)\mathbf{S}.$$

16.11 By (16.64), (3.63), and (3.64) (ignoring $n-1$ and assuming the $\mathbf{y}_i$'s are centered),

$$\begin{aligned}(\mathbf{u}_i - \mathbf{u}_k)'(\mathbf{u}_i - \mathbf{u}_k) &= \mathbf{u}_i'\mathbf{u}_i + \mathbf{u}_k'\mathbf{u}_k - 2\mathbf{u}_i'\mathbf{u}_k \\ &= \mathbf{y}_i'\mathbf{S}^{-1}\mathbf{y}_i + \mathbf{y}_k'\mathbf{S}^{-1}\mathbf{y}_k - 2\mathbf{y}_i'\mathbf{S}^{-1}\mathbf{y}_k \\ &= (\mathbf{y}_i - \mathbf{y}_k)'\mathbf{S}^{-1}(\mathbf{y}_i - \mathbf{y}_k).\end{aligned}$$

16.12 **(a)** The first ten rows and columns of the matrix $\mathbf{B}$ are as follows:

$$\begin{pmatrix} 129849 & -26801 & -88750 & -53847 & -59118 & 43583 & -73877 & 81571 & 112101 & 80909 \\ -26801 & 2310 & 17029 & 11125 & 14394 & -11076 & 18149 & -18662 & -21852 & -16306 \\ -88750 & 17029 & 65973 & 32378 & 31044 & -31085 & 68156 & -56671 & -79481 & -54135 \\ -53847 & 11125 & 32378 & 27683 & 31808 & -18550 & 30003 & -34154 & -46882 & -32096 \\ -59118 & 14394 & 31044 & 31808 & 38141 & -19161 & 27269 & -37147 & -51673 & -34687 \\ 45383 & -11076 & -31085 & -18550 & -19161 & 14741 & -33620 & 27054 & 45347 & 29211 \\ -73877 & 18149 & 68156 & 30003 & 27269 & -33620 & 76423 & -45782 & -86804 & -58650 \\ 81571 & -18662 & -56671 & -34154 & -37147 & 27054 & -45782 & 49169 & 75557 & 50169 \\ 112101 & -21852 & -79481 & -46882 & -51673 & 45347 & -86804 & 75557 & 119634 & 81258 \\ 80909 & -16306 & -54135 & -32096 & -34687 & 29211 & -58650 & 50169 & 81258 & 53286 \end{pmatrix}$$

(b) The first two columns of the matrix $\mathbf{Z}$ are given by:

City	z_1	z_2	City	z_1	z_2
A	354.1	−10.2	M	391.6	47.5
B	−77.1	25.0	N	21.0	−44.7
C	−238.2	−75.7	O	9.8	30.9
D	−154.9	65.9	P	−173.8	−78.5
E	−163.2	72.2	Q	6.3	17.1
F	126.0	24.9	R	117.0	−48.0
G	−228.8	−149.4	S	−102.3	−170.2
H	223.9	1.5	T	−53.2	−27.3
I	337.7	44.8	U	−315.2	190.9
J	226.7	34.7	V	−255.7	140.2
K	−33.4	22.3	W	−19.3	−34.3
L	1.1	−79.4			

(c) The metric multidimensional scaling plot shows the relative positions of the cities.

16.13 **(a)** The multidimensional scaling plot shows two clusters, one for positive values of the first dimension, and one for negative values. The two clusters can be interpreted as comfort (positive values) and discomfort (negative values). Hence, the axis of the first dimension can be interpreted as the level of comfort.

(b) The dendrogram for Ward's method clearly shows two clusters, the same as in part (a).

16.14 **(a)** The initial configuration of points will vary. One example is as follows:

y_1	y_2	y_3	y_4	y_5	y_6
1.458	.769	−1.350	.456	−1.610	1.827
−.598	−1.069	−2.667	.458	.416	1.094
−1.777	−.409	.369	.655	−.058	1.177
.071	.361	1.157	−.154	.343	−.417
−.060	1.361	.743	1.436	.332	−.894
−.757	−.432	−.545	.233	.646	−.102
−1.971	−.492	−.461	.078	1.441	.039
−1.560	−.173	.657	-.528	1.001	1.030
−.597	.814	−.898	.283	−.355	−1.115
1.449	-.942	.867	−.922	.833	1.196
−1.809	−.093	−1.762	−.533	−1.136	−.226
1.067	.199	.978	.884	−1.060	−.800

(b) Answers will vary. For the seeds given in part (a), STRESS = 0.0266.

(c) Answers will vary. The plot of STRESS versus k for one solution showed that two dimensions should be retained. The nonmetric MDS plot showed that Franco, Mussolini, and Hitler lie closely together, as well as Churchill and De Gaulle, and Eisenhower and Truman.

(d) Answers will vary. One solution gave results similar to part (c).

(e) Answers will vary. One solution showed three dimensions. A plot of two dimensions showed Mussolini and Franco together in the center with the others forming a circle around them almost equally spaced.

(f) Answers will vary. One solution was similar to that in part (c).

16.15 **(a)** The correspondence matrix **P** is found by dividing each element of Table 16.16 by $n = 1281$ to obtain the following:

Death Birth	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Jan	.007	.011	.009	.011	.007	.009	.008	.012	.007	.009	.009	.010	.108
Feb	.010	.005	.005	.006	.007	.004	.003	.004	.005	.009	.001	.010	.069
Mar	.009	.011	.007	.005	.013	.008	.007	.008	.007	.002	.010	.007	.094
Apr	.005	.009	.008	.005	.007	.009	.003	.009	.003	.007	.006	.009	.080
May	.006	.005	.009	.005	.003	.009	.007	.007	.009	.005	.007	.003	.075
Jun	.011	.004	.004	.005	.010	.004	.005	.003	.006	.007	.005	.004	.069
Jul	.009	.008	.010	.003	.004	.009	.005	.005	.003	.008	.003	.006	.073
Aug	.005	.005	.009	.010	.008	.007	.002	.006	.006	.006	.006	.009	.081
Sep	.005	.009	.009	.008	.008	.009	.003	.006	.009	.005	.006	.005	.083
Oct	.012	.006	.009	.007	.005	.008	.009	.006	.007	.006	.005	.005	.087
Nov	.005	.007	.012	.008	.009	.008	.005	.008	.005	.008	.007	.005	.087
Dec	.005	.014	.007	.009	.011	.006	.007	.007	.008	.005	.008	.006	.092
Total	.092	.094	.096	.084	.092	.088	.066	.080	.077	.075	.074	.081	1.000

(b) The $\mathbf{R}$ matrix is given by

$$\mathbf{R} = \begin{pmatrix}
.07 & .11 & .12 & .11 & .07 & .04 & .11 & .10 & .09 & .08 & .09 & .04 \\
.13 & .08 & .12 & .07 & .07 & .03 & .09 & .11 & .10 & .08 & .08 & .08 \\
.09 & .08 & .07 & .15 & .05 & .08 & .07 & .08 & .12 & .08 & .05 & .08 \\
.09 & .06 & .15 & .08 & .15 & .04 & .06 & .07 & .10 & .01 & .12 & .08 \\
.10 & .11 & .09 & .10 & .07 & .07 & .08 & .09 & .07 & .08 & .08 & .07 \\
.04 & .06 & .09 & .11 & .13 & .07 & .12 & .14 & .05 & .04 & .11 & .04 \\
.08 & .04 & .06 & .06 & .16 & .08 & .06 & .06 & .15 & .08 & .10 & .09 \\
.06 & .08 & .07 & .12 & .10 & .07 & .08 & .07 & .14 & .11 & .02 & .07 \\
.07 & .09 & .04 & .06 & .08 & .09 & .13 & .11 & .04 & .09 & .06 & .11 \\
.09 & .09 & .05 & .08 & .06 & .06 & .09 & .14 & .10 & .08 & .09 & .06 \\
.08 & .07 & .06 & .07 & .14 & .11 & .09 & .10 & .06 & .06 & .07 & .08 \\
.09 & .08 & .07 & .11 & .07 & .04 & .10 & .10 & .09 & .08 & .06 & .11
\end{pmatrix},$$

and the $\mathbf{C}$ matrix is given by

$$\mathbf{C} = \begin{pmatrix}
.07 & .11 & .12 & .09 & .06 & .05 & .10 & .08 & .08 & .08 & .09 & .04 \\
.12 & .08 & .12 & .06 & .04 & .08 & .09 & .08 & .08 & .08 & .08 & .08 \\
.10 & .09 & .08 & .15 & .05 & .11 & .07 & .07 & .12 & .11 & .06 & .10 \\
.07 & .05 & .13 & .06 & .11 & .05 & .04 & .05 & .08 & .01 & .11 & .07 \\
.13 & .15 & .13 & .12 & .08 & .12 & .10 & .10 & .08 & .12 & .11 & .09 \\
.04 & .06 & .08 & .08 & .10 & .08 & .10 & .11 & .04 & .04 & .10 & .04 \\
.07 & .04 & .05 & .04 & .12 & .08 & .04 & .04 & .11 & .07 & .09 & .08 \\
.07 & .10 & .09 & .12 & .10 & .11 & .09 & .07 & .14 & .14 & .02 & .09 \\
.07 & .09 & .04 & .05 & .07 & .11 & .11 & .09 & .03 & .09 & .10 & .07 \\
.08 & .08 & .07 & .07 & .14 & .14 & .09 & .09 & .06 & .07 & .08 & .09 \\
.09 & .08 & .07 & .10 & .06 & .05 & .10 & .09 & .08 & .08 & .06 & .12
\end{pmatrix}.$$

(c) The chi-square statistic is 117.7742 with 121 degrees of freedom, which gives a p-value of .5660. The two variables appear to be independent.

(d) In the correspondence plot, the following associations are seen:

{November births, June deaths}, {March deaths, April deaths, January births}, {September births, February deaths}, {August births, April births}, {May deaths, September deaths, May births}.

16.16 **(a)** The correspondence matrix **P** is found by dividing each element of Table 16.17 by 8193 to obtain the following:

Part of Country	Burglary	Fraud	Vandalism	Total
Oslo area	.048	.300	.215	.563
Mid-Norway	.018	.019	.112	.148
North Norway	.085	.040	.164	.289
Total	.151	.358	.491	1.000

(b) The **R** matrix is given by

Part of Country	Burglary	Fraud	Vandalism
Oslo area	.086	.533	.381
Mid-Norway	.121	.126	.753
North Norway	.293	.138	.569

and the **C** matrix is given by

Part of Country	Burglary	Fraud	Vandalism
Oslo area	.320	.837	.437
Mid-Norway	.119	.052	.228
North Norway	.561	.111	.335

(c) The chi-square statistic is 1662.6 with 4 degrees of freedom, which gives a p-value less than .0001. The two variables are dependent.

(d) In the correspondence plot, North Norway is associated with burglaries, Oslo is associated with fraud, and mid-Norway is associated with vandalism.

16.17 **(a)** The Burt matrix is given below.

No	5254	0	564	3408	1282	1830	3424	2466	2788	2190	3064	686	2666	1902
Yes	0	165	105	42	18	73	92	37	128	4	125	26	63	76
High dust	564	105	669	0	0	402	267	62	607	218	451	87	359	223
Low dust	3408	42	0	3450	0	1056	2394	1642	1808	1446	2004	480	1684	1286
Medium dust	1282	18	0	0	1300	445	855	799	501	566	734	145	686	469
Race–other	1830	73	402	1056	445	1930	0	932	971	799	1104	108	1658	137
White	3424	92	267	2394	855	0	3516	1571	1945	1431	2085	604	1071	1841
Female	2466	37	62	1642	799	932	1571	2503	0	1373	1130	266	1421	816
Male	2788	128	607	1808	501	971	1945	0	2916	857	2059	446	1308	1162
Nonsmoker	2190	40	218	1446	566	799	1431	1373	857	2230	0	231	1142	857
Smoker	3064	125	451	2004	734	1104	2085	1130	2059	0	3189	481	1587	1121
10–20	686	26	87	480	145	108	604	266	446	231	481	712	0	0
≤ 10	2666	63	359	1684	686	1658	1071	1421	1308	1142	1587	0	2729	0
≥ 20	1902	76	223	1286	469	137	1841	816	1162	857	1121	0	0	1978

(b) The column coordinates for the plot are given by

Variables	y_1	y_2
No	−.032	−.087
Yes	1.013	2.761
High dust	1.072	1.648
Low dust	−.209	−.107
Medium dust	.003	−.564
Race–other	1.184	−.153
White	−.641	.083
Female	.007	−.791
Male	−.006	.679
Nonsmoker	−.036	−.592
Smoker	.025	.414
10–20	−.605	.535
≤ 10	.789	−.300
≥ 20	−.871	.221

(c) Some associations seen in the plot are {byssinosis-yes, high dust}, {female, nonsmoker, medium dust}, {smoker, male}.

16.18 **(a)** The two-dimensional coordinates of the observation points and variable points are given below.

Observation Points

Name	Coordinate 1	Coordinate 2
Albania	14.102	−1.322
Austria	−5.461	1.548
Belgium	−6.077	−1.479
Bulgaria	26.116	3.319
Czech.	3.317	−2.092
Denmark	−13.861	1.374
E. Germany	−4.902	−8.360
Finland	−12.262	11.290
France	−6.345	.672
Greece	9.036	3.033
Hungary	10.805	−2.363
Ireland	−11.857	5.312
Italy	6.309	−1.314
Netherlands	−11.809	2.133
Norway	−11.005	−.077
Poland	2.526	2.999
Portugal	.784	−16.753
Romania	19.067	2.591
Spain	1.923	−10.483
Sweden	−14.842	.726
Switzerland	−9.068	4.000
UK	−9.311	.698
USSR	10.586	4.355
W. Germany	−13.514	−3.353
Yugoslavia	25.742	3.548

Variable Points

Name	Coordinate 1	Coordinate 2
R_MEAT	−.151	.133
W_MEAT	−.129	.043
EGGS	−.067	.021
MILK	−.425	.831
FISH	−.127	−.292
CEREALS	.861	.406
STARCHY	−.067	−.076
NUTS	.114	−.070
FRUIT_VEG	.020	−.169

In the biplot, the arrows for variables are too short to pass through the points for observations.

(b) The two-dimensional coordinates of the observation points and variable points are given below.

Observation Points

Name	Coordinate 1	Coordinate 2
Albania	.231	−.049
Austria	−.089	.057
Belgium	−.100	−.055
Bulgaria	.428	.122
Czech.	.054	−.077
Denmark	−.227	.051
E. Germany	−.080	−.308
Finland	−.201	.416
France	−.104	.025
Greece	.148	.112
Hungary	.177	−.087
Ireland	−.194	.196
Italy	.103	−.048
Netherlands	−.193	.079
Norway	−.180	−.003
Poland	.041	.110
Portugal	.013	−.617
Romania	.312	.095
Spain	.032	−.386
Sweden	−.243	.027
Switzerland	−.149	.147
UK	−.153	.026
USSR	.173	.160
W. Germany	−.221	−.124
Yugoslavia	.422	.131

Variable Points

Name	Coordinate 1	Coordinate 2
R_MEAT	−9.196	3.602
W_MEAT	−7.904	1.179
EGGS	−4.106	.569
MILK	−25.964	22.552
FISH	−7.750	−7.934
CEREALS	52.545	11.025
STARCHY	−4.080	−2.064
NUTS	6.953	−1.902
FRUIT_VEG	1.235	−4.593

In the biplot, the observation points are tightly clustered around the point (0,0), making them difficult to distinguish, whereas variable points are easily discerned. Red meats, white meats, and milk are highly positively correlated. These three variables are negatively correlated with nuts and fruit_veg.

(c) The two-dimensional coordinates of the observation points and variable points are given below.

Observation Points

Name	Coordinate 1	Coordinate 2
Albania	1.805	−.254
Austria	−.699	.297
Belgium	−.778	−.284
Bulgaria	3.343	.637
Czech.	.425	−.402
Denmark	−1.774	.264
E. Germany	−.627	−1.605
Finland	−1.570	2.167
France	−.812	.129
Greece	1.157	.582
Hungary	1.383	−.454
Ireland	−1.518	1.020
Italy	.808	−.252
Netherlands	−1.511	.409
Norway	−1.409	−.015
Poland	.323	.576
Portugal	.100	−3.216
Romania	2.441	.497
Spain	.246	−2.012
Sweden	−1.900	.139
Switzerland	−1.161	.768
UK	−1.192	.134
USSR	1.355	.836
W. Germany	−1.730	−.644
Yugoslavia	3.295	.681

Variable Points

Name	Coordinate 1	Coordinate 2
R_MEAT	−1.177	.691
W_MEAT	−1.012	.226
EGGS	−.526	.109
MILK	−3.323	4.329
FISH	−.992	−1.523
CEREALS	6.726	2.116
STARCHY	−.522	−.396
NUTS	.890	−.365
FRUIT_VEG	.158	−.882

In the biplot, the variable points and observation points are both well spaced. Finland scored high on the milk variable. Yugoslavia and Bulgaria scored high on the cereal variable. Spain and Portugal scored highest on the fish and frut_veg variables.

(d) The biplot from part (c) seems better because the scales on the variables and points are more evenly matched.

16.19 **(a)** The two-dimensional coordinates of the observation points and variable points are as follows.

Observation Points

Name	Coordinate 1	Coordinate 2
FSM1	−9.535	−4.752
Sister	2.705	.796
FSM2	4.043	−.584
Father	4.392	.614
Teacher	−8.708	5.008
MSM	3.409	.701
FSM3	3.694	−1.782

Variable Points

Name	Coordinate 1	Coordinate 2
KIND	.610	−.054
INTEL	.085	.413
HAPPY	.407	−.456
LIKE	.621	−.039
JUST	.264	.785

In the biplot, the arrows for the variables are too short to pass through the points for observations.

(b) The two-dimensional coordinates of the observation points and variable points are given below.

Observation Points

Name	Coordinate 1	Coordinate 2
FSM1	−.622	−.655
Sister	.176	.110
FSM2	.264	−.080
Father	.287	.085
Teacher	−.568	.690
MSM	.222	.097
FSM3	.241	−.246

Variable Points

Name	Coordinate 1	Coordinate 2
KIND	9.345	−.391
INTEL	1.298	2.997
HAPPY	6.235	−3.313
LIKE	9.521	−.282
JUST	4.054	5.700

In the biplot, the observation points are tightly clustered around the point (0,0) making them difficult to distinguish, whereas variable points are well spaced. Just and intelligent are highly positively correlated, as are kind and likeable.

(c) The two-dimensional coordinates of the observation points and variable points are given below.

Observation Points

Name	Coordinate 1	Coordinate 2
FSM1	−2.435	−1.764
Sister	.691	.295
FSM2	1.033	−.217
Father	1.122	.228
Teacher	−2.224	1.859
MSM	.871	.260
FSM3	.943	−.662

Variable Points

Name	Coordinate 1	Coordinate 2
KIND	2.387	−.145
INTEL	.331	1.113
HAPPY	1.593	−1.230
LIKE	2.432	−.105
JUST	1.036	2.116

In the biplot, the variable points and observation points are both well spaced. Father, sister, FSM2 and FSM3 all scored high on the kind, likeable, and happy variables, whereas teacher and FSM1 scored negatively on those variables.

(d) The biplot from part (c) seems better because the scales on the variables and points are more evenly matched.

16.20 **(a)** The two-dimensional coordinates of the observation points and variable points are given below.

Observation Points

Name	Coordinate 1	Coordinate 2
1	49.410	−5.832
2	25.407	−7.658
3	21.600	−2.340
4	−23.545	−6.367
5	−28.477	−4.773
6	−33.341	2.315
7	−28.176	7.992
8	−25.786	12.655
9	−29.703	9.275
10	−33.868	−3.776
11	−33.529	−1.977
12	28.186	−16.031
13	10.804	−6.608
14	.566	3.021
15	77.970	.109
16	12.859	16.294
17	41.960	5.103
18	46.930	19.064
19	34.958	−1.018
20	−16.477	1.148
21	−23.634	−1.055
22	−34.036	−2.424
23	20.632	−5.882
24	−15.873	−6.731
25	−23.023	.745
26	−15.183	−1.942
27	−11.903	−6.917
28	5.273	3.610

Variable Points

Name	Coordinate 1	Coordinate 2
North	.526	.225
East	.429	.752
South	.579	−.379
West	.452	−.490

In the biplot, the variable points are tightly grouped and the corresponding arrows do not pass through the points for observations.

(b) The two-dimensional coordinates of the observation points and variable points are given below.

Observation Points

Name	Coordinate 1	Coordinate 2
1	.303	−.145
2	.156	−.191
3	.132	−.058
4	−.144	−.158
5	−.175	−.119
6	−.205	.058
7	−.173	.199
8	−.158	.315
9	−.182	.231
10	−.208	−.094
11	−.206	−.049
12	.173	−.399
13	.066	−.164
14	.003	.075
15	.478	.003
16	.079	.406
17	.257	.127
18	.288	.474
19	.214	−.025
20	−.101	.029
21	−.145	−.026
22	−.209	−.060
23	.127	−.146
24	−.097	−.168
25	−.141	.019
26	−.093	−.048
27	−.073	−.172
28	.032	.090

Variable Points

Name	Coordinate 1	Coordinate 2
North	85.779	9.026
East	69.899	30.223
South	94.377	−15.213
West	73.682	−19.694

In the biplot, the observation points are tightly clustered around the point (0,0), making them difficult to distinguish, whereas variable points are well spaced. All the variables are positively correlated, with south and west showing the closest relationship.

(c) The two-dimensional coordinates of the observation points and variable points are given below.

Observation Points

Name	Coordinate 1	Coordinate 2
1	3.870	−.920
2	1.990	−1.208
3	1.692	−.369
4	−1.844	−1.004
5	−2.230	−.753
6	−2.611	.365
7	−2.207	1.261
8	−2.020	1.996
9	−2.326	1.463
10	−2.652	−.596
11	−2.626	−.312
12	2.207	−2.529
13	.846	−1.042
14	.044	.477
15	6.106	.017
16	1.007	2.571
17	3.286	.805
18	3.675	3.008
19	2.738	−.161
20	−1.290	.181
21	−1.851	−.166
22	−2.666	−.382
23	1.616	−.928
24	−1.243	−1.062
25	−1.803	.118
26	−1.189	−.306
27	−.932	−1.091
28	.413	.569

Variable Points

Name	Coordinate 1	Coordinate 2
North	6.718	1.424
East	5.474	4.768
South	7.391	−2.400
West	5.771	−3.107

In the biplot, the variable points and observation points are both well spaced. Tree 18 is associated with east, 17 with north, 1 and 3 with south, and 2 and 23 with west.

(d) The biplot from part (c) seems better because the scales on the variables and points are more evenly matched.