lies in the hyperplane of linear combinations of the g vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_g$. Since $\mathbf{1} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_g$, the mean vector also lies in this hyperplane, and it is *always* perpendicular to the treatment vector. (See Exercise 6.10.) Thus, the mean vector has the freedom to lie anywhere along the one-dimensional equiangular line, and the treatment vector has the freedom to lie anywhere in the other $g - 1$ dimensions. The residual vector, $\hat{\mathbf{e}} = \mathbf{y} - (\bar{x}\mathbf{1}) - [(\bar{x}_1 - \bar{x})\mathbf{u}_1 + \cdots + (\bar{x}_g - \bar{x})\mathbf{u}_g]$ is perpendicular to both the mean vector and the treatment effect vector and has the freedom to lie anywhere in the subspace of dimension $n - (g - 1) - 1 = n - g$ that is perpendicular to their hyperplane.

To summarize, we attribute 1 d.f. to SS_{mean}, $g - 1$ d.f. to SS_{tr}, and $n - g = (n_1 + n_2 + \cdots + n_g) - g$ d.f. to SS_{res}. The total number of degrees of freedom is $n = n_1 + n_2 + \cdots + n_g$. Alternatively, by appealing to the univariate distribution theory, we find that these are the degrees of freedom for the chi-square distributions associated with the corresponding sums of squares.

The calculations of the sums of squares and the associated degrees of freedom are conveniently summarized by an ANOVA table.

ANOVA Table for Comparing Univariate Population Means

Source of variation	Sum of squares (SS)	Degrees of freedom (d.f.)
Treatments	$SS_{tr} = \sum_{\ell=1}^{g} n_\ell(\bar{x}_\ell - \bar{x})^2$	$g - 1$
Residual (error)	$SS_{res} = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (x_{\ell j} - \bar{x}_\ell)^2$	$\sum_{\ell=1}^{g} n_\ell - g$
Total (corrected for the mean)	$SS_{cor} = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (x_{\ell j} - \bar{x})^2$	$\sum_{\ell=1}^{g} n_\ell - 1$

The usual F-test rejects $H_0: \tau_1 = \tau_2 = \cdots = \tau_g = 0$ at level α if

$$F = \frac{SS_{tr}/(g-1)}{SS_{res}\Big/\left(\sum_{\ell=1}^{g} n_\ell - g\right)} > F_{g-1, \Sigma n_\ell - g}(\alpha)$$

where $F_{g-1,\Sigma n_\ell - g}(\alpha)$ is the upper (100α)th percentile of the F-distribution with $g - 1$ and $\Sigma n_\ell - g$ degrees of freedom. This is equivalent to rejecting H_0 for large values of SS_{tr}/SS_{res} or for large values of $1 + SS_{tr}/SS_{res}$. The statistic appropriate for a multivariate generalization rejects H_0 for *small* values of the reciprocal

$$\frac{1}{1 + SS_{tr}/SS_{res}} = \frac{SS_{res}}{SS_{res} + SS_{tr}} \tag{6-37}$$

Example 6.8 (A univariate ANOVA table and *F*-test for treatment effects) Using the information in Example 6.7, we have the following ANOVA table:

Source of variation	Sum of squares	Degrees of freedom
Treatments	$SS_{tr} = 78$	$g - 1 = 3 - 1 = 2$
Residual	$SS_{res} = 10$	$\sum_{\ell=1}^{g} n_\ell - g = (3 + 2 + 3) - 3 = 5$
Total (corrected)	$SS_{cor} = 88$	$\sum_{\ell=1}^{g} n_\ell - 1 = 7$

Consequently,

$$F = \frac{SS_{tr}/(g-1)}{SS_{res}/(\Sigma n_\ell - g)} = \frac{78/2}{10/5} = 19.5$$

Since $F = 19.5 > F_{2,5}(.01) = 13.27$, we reject $H_0: \tau_1 = \tau_2 = \tau_3 = 0$ (no treatment effect) at the 1% level of significance. ■

Multivariate Analysis of Variance (MANOVA)

Paralleling the univariate reparameterization, we specify the MANOVA model:

MANOVA Model For Comparing g Population Mean Vectors

$$\mathbf{X}_{\ell j} = \boldsymbol{\mu} + \boldsymbol{\tau}_\ell + \mathbf{e}_{\ell j}, \quad j = 1, 2, \ldots, n_\ell \quad \text{and} \quad \ell = 1, 2, \ldots, g \tag{6-38}$$

where the $\mathbf{e}_{\ell j}$ are independent $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ variables. Here the parameter vector $\boldsymbol{\mu}$ is an overall mean (level), and $\boldsymbol{\tau}_\ell$ represents the ℓth treatment effect with $\sum_{\ell=1}^{g} n_\ell \boldsymbol{\tau}_\ell = \mathbf{0}$.

According to the model in (6-38), *each component* of the observation vector $\mathbf{X}_{\ell j}$ satisfies the univariate model (6-33). The errors for the components of $\mathbf{X}_{\ell j}$ are correlated, but the covariance matrix $\boldsymbol{\Sigma}$ is the same for all populations.

A vector of observations may be decomposed as suggested by the model. Thus,

$$\underset{\text{(observation)}}{\mathbf{x}_{\ell j}} = \underset{\begin{pmatrix}\text{overall sample}\\ \text{mean } \hat{\boldsymbol{\mu}}\end{pmatrix}}{\bar{\mathbf{x}}} + \underset{\begin{pmatrix}\text{estimated}\\ \text{treatment}\\ \text{effect } \hat{\boldsymbol{\tau}}_\ell\end{pmatrix}}{(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})} + \underset{\begin{pmatrix}\text{residual}\\ \hat{\mathbf{e}}_{\ell j}\end{pmatrix}}{(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)} \tag{6-39}$$

The decomposition in (6-39) leads to the multivariate analog of the univariate sum of squares breakup in (6-35). First we note that the product

$$(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'$$

can be written as

$$(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})' = [(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell) + (\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})][(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell) + (\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})]'$$
$$= (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)' + (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})' + (\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)' + (\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'$$

The sum over j of the middle two expressions is the zero matrix, because $\sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell) = \mathbf{0}$. Hence, summing the cross product over ℓ and j yields

$$\underbrace{\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'}_{\text{total (corrected) sum of squares and cross products}} = \underbrace{\sum_{\ell=1}^{g} n_\ell(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'}_{\text{treatment (Between) sum of squares and cross products}} + \underbrace{\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'}_{\text{residual (Within) sum of squares and cross products}} \tag{6-40}$$

The *within* sum of squares and cross products matrix can be expressed as

$$\mathbf{W} = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)' = (n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2 + \cdots + (n_g - 1)\mathbf{S}_g \tag{6-41}$$

where $\mathbf{S}_\ell$ is the sample covariance matrix for the ℓth sample. This matrix is a generalization of the $(n_1 + n_2 - 2)\mathbf{S}_{\text{pooled}}$ matrix encountered in the two-sample case. It plays a dominant role in testing for the presence of treatment effects.

Analogous to the univariate result, the hypothesis of no treatment effects,

$$H_0: \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \cdots = \boldsymbol{\tau}_g = \mathbf{0}$$

is tested by considering the relative sizes of the treatment and residual sums of squares and cross products. Equivalently, we may consider the relative sizes of the residual and total (corrected) sum of squares and cross products. Formally, we summarize the calculations leading to the test statistic in a MANOVA table.

MANOVA Table for Comparing Population Mean Vectors

Source of variation	Matrix of sum of squares and cross products (SSP)	Degrees of freedom (d.f.)
Treatment	$\mathbf{B} = \sum_{\ell=1}^{g} n_\ell(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'$	$g - 1$
Residual (Error)	$\mathbf{W} = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'$	$\sum_{\ell=1}^{g} n_\ell - g$
Total (corrected for the mean)	$\mathbf{B} + \mathbf{W} = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'$	$\sum_{\ell=1}^{g} n_\ell - 1$

This table is exactly the same form, component by component, as the ANOVA table, except that squares of scalars are replaced by their vector counterparts. For example, $(\bar{x}_\ell - \bar{x})^2$ becomes $(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'$. The degrees of freedom correspond to the univariate geometry and also to some multivariate distribution theory involving Wishart densities. (See [1].)

One test of $H_0: \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \cdots = \boldsymbol{\tau}_g = \mathbf{0}$ involves generalized variances. We reject H_0 if the ratio of generalized variances

$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|} = \frac{\left|\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'\right|}{\left|\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'\right|} \tag{6-42}$$

is too small. The quantity $\Lambda^* = |\mathbf{W}|/|\mathbf{B} + \mathbf{W}|$, proposed originally by Wilks (see [25]), corresponds to the equivalent form (6-37) of the F-test of H_0: no treatment effects in the univariate case. *Wilks' lambda* has the virtue of being convenient and related to the likelihood ratio criterion.[2] The exact distribution of Λ^* can be derived for the special cases listed in Table 6.3. For other cases and large sample sizes, a modification of Λ^* due to Bartlett (see [4]) can be used to test H_0.

Table 6.3 Distribution of Wilks' Lambda, $\Lambda^* = |\mathbf{W}|/|\mathbf{B} + \mathbf{W}|$

No. of variables	No. of groups	Sampling distribution for multivariate normal data
$p = 1$	$g \geq 2$	$\left(\frac{\Sigma n_\ell - g}{g - 1}\right)\left(\frac{1 - \Lambda^*}{\Lambda^*}\right) \sim F_{g-1, \Sigma n_\ell - g}$
$p = 2$	$g \geq 2$	$\left(\frac{\Sigma n_\ell - g - 1}{g - 1}\right)\left(\frac{1 - \sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2(g-1), 2(\Sigma n_\ell - g - 1)}$
$p \geq 1$	$g = 2$	$\left(\frac{\Sigma n_\ell - p - 1}{p}\right)\left(\frac{1 - \Lambda^*}{\Lambda^*}\right) \sim F_{p, \Sigma n_\ell - p - 1}$
$p \geq 1$	$g = 3$	$\left(\frac{\Sigma n_\ell - p - 2}{p}\right)\left(\frac{1 - \sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2p, 2(\Sigma n_\ell - p - 2)}$

[2] Wilks' lambda can also be expressed as a function of the eigenvalues of $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_s$ of $\mathbf{W}^{-1}\mathbf{B}$ as

$$\Lambda^* = \prod_{i=1}^{s}\left(\frac{1}{1 + \hat{\lambda}_i}\right)$$

where $s = \min(p, g - 1)$, the rank of $\mathbf{B}$. Other statistics for checking the equality of several multivariate means, such as Pillai's statistic, the Lawley–Hotelling statistic, and Roy's largest root statistic can also be written as particular functions of the eigenvalues of $\mathbf{W}^{-1}\mathbf{B}$. For large samples, all of these statistics are, essentially equivalent. (See the additional discussion on page 336.)

Bartlett (see [4]) has shown that if H_0 is true and $\Sigma n_\ell = n$ is large,

$$-\left(n - 1 - \frac{(p+g)}{2}\right)\ln \Lambda^* = -\left(n - 1 - \frac{(p+g)}{2}\right)\ln\left(\frac{|\mathbf{W}|}{|\mathbf{B}+\mathbf{W}|}\right) \qquad (6\text{-}43)$$

has approximately a chi-square distribution with $p(g-1)$ d.f. Consequently, for $\Sigma n_\ell = n$ large, we reject H_0 at significance level α if

$$-\left(n - 1 - \frac{(p+g)}{2}\right)\ln\left(\frac{|\mathbf{W}|}{|\mathbf{B}+\mathbf{W}|}\right) > \chi^2_{p(g-1)}(\alpha) \qquad (6\text{-}44)$$

where $\chi^2_{p(g-1)}(\alpha)$ is the upper (100α)th percentile of a chi-square distribution with $p(g-1)$ d.f.

Example 6.9 (A MANOVA table and Wilks' lambda for testing the equality of three mean vectors) Suppose an additional variable is observed along with the variable introduced in Example 6.7. The sample sizes are $n_1 = 3, n_2 = 2,$ and $n_3 = 3$. Arranging the observation pairs $\mathbf{x}_{\ell j}$ in rows, we obtain

$$\begin{pmatrix} \begin{bmatrix}9\\3\end{bmatrix} & \begin{bmatrix}6\\2\end{bmatrix} & \begin{bmatrix}9\\7\end{bmatrix} \\ \begin{bmatrix}0\\4\end{bmatrix} & \begin{bmatrix}2\\0\end{bmatrix} & \\ \begin{bmatrix}3\\8\end{bmatrix} & \begin{bmatrix}1\\9\end{bmatrix} & \begin{bmatrix}2\\7\end{bmatrix} \end{pmatrix} \quad \text{with } \bar{\mathbf{x}}_1 = \begin{bmatrix}8\\4\end{bmatrix},\ \bar{\mathbf{x}}_2 = \begin{bmatrix}1\\2\end{bmatrix},\ \bar{\mathbf{x}}_3 = \begin{bmatrix}2\\8\end{bmatrix}, \quad \text{and } \bar{\mathbf{x}} = \begin{bmatrix}4\\5\end{bmatrix}$$

We have already expressed the observations on the first variable as the sum of an overall mean, treatment effect, and residual in our discussion of univariate ANOVA. We found that

$$\underset{\text{(observation)}}{\begin{pmatrix}9&6&9\\0&2&\\3&1&2\end{pmatrix}} = \underset{\text{(mean)}}{\begin{pmatrix}4&4&4\\4&4&\\4&4&4\end{pmatrix}} + \underset{\text{(treatment effect)}}{\begin{pmatrix}4&4&4\\-3&-3&\\-2&-2&-2\end{pmatrix}} + \underset{\text{(residual)}}{\begin{pmatrix}1&-2&1\\-1&1&\\1&-1&0\end{pmatrix}}$$

and

$$SS_{obs} = SS_{mean} + SS_{tr} + SS_{res}$$
$$216 = 128 + 78 + 10$$
$$\text{Total SS (corrected)} = SS_{obs} - SS_{mean} = 216 - 128 = 88$$

Repeating this operation for the observations on the second variable, we have

$$\underset{\text{(observation)}}{\begin{pmatrix}3&2&7\\4&0&\\8&9&7\end{pmatrix}} = \underset{\text{(mean)}}{\begin{pmatrix}5&5&5\\5&5&\\5&5&5\end{pmatrix}} + \underset{\text{(treatment effect)}}{\begin{pmatrix}-1&-1&-1\\-3&-3&\\3&3&3\end{pmatrix}} + \underset{\text{(residual)}}{\begin{pmatrix}-1&-2&3\\2&-2&\\0&1&-1\end{pmatrix}}$$

and

$$SS_{obs} = SS_{mean} + SS_{tr} + SS_{res}$$
$$272 = 200 + 48 + 24$$
$$\text{Total SS (corrected)} = SS_{obs} - SS_{mean} = 272 - 200 = 72$$

These two single-component analyses must be augmented with the sum of entry-by-entry *cross products* in order to complete the entries in the MANOVA table. Proceeding row by row in the arrays for the two variables, we obtain the cross product contributions:

$$\text{Mean: } 4(5) + 4(5) + \cdots + 4(5) = 8(4)(5) = 160$$
$$\text{Treatment: } 3(4)(-1) + 2(-3)(-3) + 3(-2)(3) = -12$$
$$\text{Residual: } 1(-1) + (-2)(-2) + 1(3) + (-1)(2) + \cdots + 0(-1) = 1$$
$$\text{Total: } 9(3) + 6(2) + 9(7) + 0(4) + \cdots + 2(7) = 149$$

$$\text{Total (corrected) cross product} = \text{total cross product} - \text{mean cross product}$$
$$= 149 - 160 = -11$$

Thus, the MANOVA table takes the following form:

Source of variation	Matrix of sum of squares and cross products	Degrees of freedom
Treatment	$\begin{bmatrix}78 & -12\\-12 & 48\end{bmatrix}$	$3 - 1 = 2$
Residual	$\begin{bmatrix}10 & 1\\1 & 24\end{bmatrix}$	$3 + 2 + 3 - 3 = 5$
Total (corrected)	$\begin{bmatrix}88 & -11\\-11 & 72\end{bmatrix}$	7

Equation (6-40) is verified by noting that

$$\begin{bmatrix}88 & -11\\-11 & 72\end{bmatrix} = \begin{bmatrix}78 & -12\\-12 & 48\end{bmatrix} + \begin{bmatrix}10 & 1\\1 & 24\end{bmatrix}$$

Using (6-42), we get

$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B}+\mathbf{W}|} = \frac{\begin{vmatrix}10 & 1\\1 & 24\end{vmatrix}}{\begin{vmatrix}88 & -11\\-11 & 72\end{vmatrix}} = \frac{10(24) - (1)^2}{88(72) - (-11)^2} = \frac{239}{6215} = .0385$$

Since $p = 2$ and $g = 3$, Table 6.3 indicates that an exact test (assuming normality and equal group covariance matrices) of $H_0: \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \boldsymbol{\tau}_3 = \mathbf{0}$ (no treatment effects) versus H_1: at least one $\boldsymbol{\tau}_\ell \neq \mathbf{0}$ is available. To carry out the test, we compare the test statistic

$$\left(\frac{1 - \sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right)\frac{(\Sigma n_\ell - g - 1)}{(g - 1)} = \left(\frac{1 - \sqrt{.0385}}{\sqrt{.0385}}\right)\left(\frac{8 - 3 - 1}{3 - 1}\right) = 8.19$$

with a percentage point of an F-distribution having $\nu_1 = 2(g - 1) = 4$ and $\nu_2 = 2(\Sigma n_\ell - g - 1) = 8$ d.f. Since $8.19 > F_{4,8}(.01) = 7.01$, we reject H_0 at the $\alpha = .01$ level and conclude that treatment differences exist. ■

When the number of variables, p, is large, the MANOVA table is usually not constructed. Still, it is good practice to have the computer print the matrices **B** and **W** so that especially large entries can be located. Also, the residual vectors

$$\hat{\mathbf{e}}_{\ell j} = \mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell$$

should be examined for normality and the presence of outliers using the techniques discussed in Sections 4.6 and 4.7 of Chapter 4.

Example 6.10 (A multivariate analysis of Wisconsin nursing home data) The Wisconsin Department of Health and Social Services reimburses nursing homes in the state for the services provided. The department develops a set of formulas for rates for each facility, based on factors such as level of care, mean wage rate, and average wage rate in the state.

Nursing homes can be classified on the basis of ownership (private party, nonprofit organization, and government) and certification (skilled nursing facility, intermediate care facility, or a combination of the two).

One purpose of a recent study was to investigate the effects of ownership or certification (or both) on costs. Four costs, computed on a per-patient-day basis and measured in hours per patient day, were selected for analysis: X_1 = cost of nursing labor, X_2 = cost of dietary labor, X_3 = cost of plant operation and maintenance labor, and X_4 = cost of housekeeping and laundry labor. A total of $n = 516$ observations on each of the $p = 4$ cost variables were initially separated according to ownership. Summary statistics for each of the $g = 3$ groups are given in the following table.

Group	Number of observations	Sample mean vectors
$\ell = 1$ (private)	$n_1 = 271$	$\bar{\mathbf{x}}_1 = \begin{bmatrix} 2.066 \\ .480 \\ .082 \\ .360 \end{bmatrix}$
$\ell = 2$ (nonprofit)	$n_2 = 138$	$\bar{\mathbf{x}}_2 = \begin{bmatrix} 2.167 \\ .596 \\ .124 \\ .418 \end{bmatrix}$
$\ell = 3$ (government)	$n_3 = 107$	$\bar{\mathbf{x}}_3 = \begin{bmatrix} 2.273 \\ .521 \\ .125 \\ .383 \end{bmatrix}$
	$\sum_{\ell=1}^{3} n_\ell = 516$	

Sample covariance matrices

$$\mathbf{S}_1 = \begin{bmatrix} .291 & & & \\ -.001 & .011 & & \\ .002 & .000 & .001 & \\ .010 & .003 & .000 & .010 \end{bmatrix}; \quad \mathbf{S}_2 = \begin{bmatrix} .561 & & & \\ .011 & .025 & & \\ .001 & .004 & .005 & \\ .037 & .007 & .002 & .019 \end{bmatrix};$$

$$\mathbf{S}_3 = \begin{bmatrix} .261 & & & \\ .030 & .017 & & \\ .003 & -.000 & .004 & \\ .018 & .006 & .001 & .013 \end{bmatrix}$$

Source: Data courtesy of State of Wisconsin Department of Health and Social Services.

Since the $\mathbf{S}_\ell$'s seem to be reasonably compatible,[3] they were pooled [see (6-41)] to obtain

$$\mathbf{W} = (n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2 + (n_3 - 1)\mathbf{S}_3$$

$$= \begin{bmatrix} 182.962 & & & \\ 4.408 & 8.200 & & \\ 1.695 & .633 & 1.484 & \\ 9.581 & 2.428 & .394 & 6.538 \end{bmatrix}$$

Also,

$$\bar{\mathbf{x}} = \frac{n_1\bar{\mathbf{x}}_1 + n_2\bar{\mathbf{x}}_2 + n_3\bar{\mathbf{x}}_3}{n_1 + n_2 + n_3} = \begin{bmatrix} 2.136 \\ .519 \\ .102 \\ .380 \end{bmatrix}$$

and

$$\mathbf{B} = \sum_{\ell=1}^{3} n_\ell(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})' = \begin{bmatrix} 3.475 & & & \\ 1.111 & 1.225 & & \\ .821 & .453 & .235 & \\ .584 & .610 & .230 & .304 \end{bmatrix}$$

To test $H_0: \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \boldsymbol{\tau}_3$ (no ownership effects or, equivalently, no difference in average costs among the three types of owners—private, nonprofit, and government), we can use the result in Table 6.3 for $g = 3$.

Computer-based calculations give

$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|} = .7714$$

[3] However, a normal-theory test of $H_0: \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_3$ would reject H_0 at any reasonable significance level because of the large sample sizes (see Example 6.12).

and

$$\left(\frac{\Sigma n_\ell - p - 2}{p}\right)\left(\frac{1 - \sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) = \left(\frac{516 - 4 - 2}{4}\right)\left(\frac{1 - \sqrt{.7714}}{\sqrt{.7714}}\right) = 17.67$$

Let $\alpha = .01$, so that $F_{2(4),2(510)}(.01) \doteq \chi_8^2(.01)/8 = 2.51$. Since $17.67 > F_{8,1020}(.01) \doteq 2.51$, we reject H_0 at the 1% level and conclude that average costs differ, depending on type of ownership.

It is informative to compare the results based on this "exact" test with those obtained using the large-sample procedure summarized in (6-43) and (6-44). For the present example, $\Sigma n_\ell = n = 516$ is large, and H_0 can be tested at the $\alpha = .01$ level by comparing

$$-(n - 1 - (p + g)/2)\ln\left(\frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|}\right) = -511.5\ln(.7714) = 132.76$$

with $\chi^2_{p(g-1)}(.01) = \chi^2_8(.01) = 20.09$. Since $132.76 > \chi^2_8(.01) = 20.09$, we reject H_0 at the 1% level. This result is consistent with the result based on the foregoing F-statistic. ■

6.5 Simultaneous Confidence Intervals for Treatment Effects

When the hypothesis of equal treatment effects is rejected, those effects that led to the rejection of the hypothesis are of interest. For pairwise comparisons, the Bonferroni approach (see Section 5.4) can be used to construct simultaneous confidence intervals for the components of the differences $\boldsymbol{\tau}_k - \boldsymbol{\tau}_\ell$ (or $\boldsymbol{\mu}_k - \boldsymbol{\mu}_\ell$). These intervals are shorter than those obtained for all contrasts, and they require critical values only for the univariate t-statistic.

Let τ_{ki} be the ith component of $\boldsymbol{\tau}_k$. Since $\boldsymbol{\tau}_k$ is estimated by $\hat{\boldsymbol{\tau}}_k = \bar{\mathbf{x}}_k - \bar{\mathbf{x}}$

$$\hat{\tau}_{ki} = \bar{x}_{ki} - \bar{x}_i \tag{6-45}$$

and $\hat{\tau}_{ki} - \hat{\tau}_{\ell i} = \bar{x}_{ki} - \bar{x}_{\ell i}$ is the difference between two independent sample means. The two-sample t-based confidence interval is valid with an appropriately modified α. Notice that

$$\text{Var}(\hat{\tau}_{ki} - \hat{\tau}_{\ell i}) = \text{Var}(\bar{X}_{ki} - \bar{X}_{\ell i}) = \left(\frac{1}{n_k} + \frac{1}{n_\ell}\right)\sigma_{ii}$$

where σ_{ii} is the ith diagonal element of $\boldsymbol{\Sigma}$. As suggested by (6-41), $\text{Var}(\bar{X}_{ki} - \bar{X}_{\ell i})$ is estimated by dividing the corresponding element of $\mathbf{W}$ by its degrees of freedom. That is,

$$\widehat{\text{Var}}(\bar{X}_{ki} - \bar{X}_{\ell i}) = \left(\frac{1}{n_k} + \frac{1}{n_\ell}\right)\frac{w_{ii}}{n - g}$$

where w_{ii} is the ith diagonal element of $\mathbf{W}$ and $n = n_1 + \cdots + n_g$.

It remains to apportion the error rate over the numerous confidence statements. Relation (5-28) still applies. There are p variables and $g(g - 1)/2$ pairwise differences, so each two-sample t-interval will employ the critical value $t_{n-g}(\alpha/2m)$, where

$$m = pg(g - 1)/2 \tag{6-46}$$

is the number of simultaneous confidence statements.

Result 6.5. Let $n = \sum_{k=1}^{g} n_k$. For the model in (6-38), with confidence at least $(1 - \alpha)$,

$$\tau_{ki} - \tau_{\ell i} \quad \text{belongs to} \quad \bar{x}_{ki} - \bar{x}_{\ell i} \pm t_{n-g}\left(\frac{\alpha}{pg(g-1)}\right)\sqrt{\frac{w_{ii}}{n-g}\left(\frac{1}{n_k} + \frac{1}{n_\ell}\right)}$$

for all components $i = 1, \ldots, p$ and all differences $\ell < k = 1, \ldots, g$. Here w_{ii} is the ith diagonal element of $\mathbf{W}$.

We shall illustrate the construction of simultaneous interval estimates for the pairwise differences in treatment means using the nursing-home data introduced in Example 6.10.

Example 6.11 (Simultaneous intervals for treatment differences—nursing homes) We saw in Example 6.10 that average costs for nursing homes differ, depending on the type of ownership. We can use Result 6.5 to estimate the magnitudes of the differences. A comparison of the variable X_3, costs of plant operation and maintenance labor, between privately owned nursing homes and government-owned nursing homes can be made by estimating $\tau_{13} - \tau_{33}$. Using (6-39) and the information in Example 6.10, we have

$$\hat{\boldsymbol{\tau}}_1 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}) = \begin{bmatrix} -.070 \\ -.039 \\ -.020 \\ -.020 \end{bmatrix}, \qquad \hat{\boldsymbol{\tau}}_3 = (\bar{\mathbf{x}}_3 - \bar{\mathbf{x}}) = \begin{bmatrix} .137 \\ .002 \\ .023 \\ .003 \end{bmatrix}$$

$$\mathbf{W} = \begin{bmatrix} 182.962 & & & \\ 4.408 & 8.200 & & \\ 1.695 & .633 & 1.484 & \\ 9.581 & 2.428 & .394 & 6.538 \end{bmatrix}$$

Consequently,

$$\hat{\tau}_{13} - \hat{\tau}_{33} = -.020 - .023 = -.043$$

and $n = 271 + 138 + 107 = 516$, so that

$$\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_3}\right)\frac{w_{33}}{n - g}} = \sqrt{\left(\frac{1}{271} + \frac{1}{107}\right)\frac{1.484}{516 - 3}} = .00614$$

Since $p = 4$ and $g = 3$, for 95% simultaneous confidence statements we require that $t_{513}(.05/4(3)2) \doteq 2.87$. (See Appendix, Table 1.) The 95% simultaneous confidence statement is

$$\tau_{13} - \tau_{33} \text{ belongs to } \hat{\tau}_{13} - \hat{\tau}_{33} \pm t_{513}(.00208)\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_3}\right)\frac{w_{33}}{n - g}}$$
$$= -.043 \pm 2.87(.00614)$$
$$= -.043 \pm .018, \text{ or } (-.061, -.025)$$

We conclude that the average maintenance and labor cost for government-owned nursing homes is higher by .025 to .061 hour per patient day than for privately owned nursing homes. With the same 95% confidence, we can say that

$$\tau_{13} - \tau_{23} \text{ belongs to the interval } (-.058, -.026)$$

and

$$\tau_{23} - \tau_{33} \text{ belongs to the interval } (-.021, .019)$$

Thus, a difference in this cost exists between private and nonprofit nursing homes, but no difference is observed between nonprofit and government nursing homes. ■

6.6 Testing for Equality of Covariance Matrices

One of the assumptions made when comparing two or more multivariate mean vectors is that the covariance matrices of the potentially different populations are the same. (This assumption will appear again in Chapter 11 when we discuss discrimination and classification.) Before pooling the variation across samples to form a pooled covariance matrix when comparing mean vectors, it can be worthwhile to test the equality of the population covariance matrices. One commonly employed test for equal covariance matrices is Box's M-test ([8], [9]).

With g populations, the null hypothesis is

$$H_0: \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \cdots = \mathbf{\Sigma}_g = \mathbf{\Sigma} \tag{6-47}$$

where $\mathbf{\Sigma}_\ell$ is the covariance matrix for the ℓth population, $\ell = 1, 2, \ldots, g$, and $\mathbf{\Sigma}$ is the presumed common covariance matrix. The alternative hypothesis is that at least two of the covariance matrices are not equal.

Assuming multivariate normal populations, a likelihood ratio statistic for testing (6-47) is given by (see [1])

$$\Lambda = \prod_\ell \left(\frac{|\mathbf{S}_\ell|}{|\mathbf{S}_{\text{pooled}}|}\right)^{(n_\ell - 1)/2} \tag{6-48}$$

Here n_ℓ is the sample size for the ℓth group, $\mathbf{S}_\ell$ is the ℓth group sample covariance matrix and $\mathbf{S}_{\text{pooled}}$ is the pooled sample covariance matrix given by

$$\mathbf{S}_{\text{pooled}} = \frac{1}{\sum_\ell (n_\ell - 1)}\left\{(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2 + \cdots + (n_g - 1)\mathbf{S}_g\right\} \tag{6-49}$$

Box's test is based on his χ^2 approximation to the sampling distribution of $-2 \ln \Lambda$ (see Result 5.2). Setting $-2 \ln \Lambda = M$ (Box's M statistic) gives

$$M = \left[\sum_\ell (n_\ell - 1)\right]\ln|\mathbf{S}_{\text{pooled}}| - \sum_\ell [(n_\ell - 1)\ln|\mathbf{S}_\ell|] \tag{6-50}$$

If the null hypothesis is true, the individual sample covariance matrices are not expected to differ too much and, consequently, do not differ too much from the pooled covariance matrix. In this case, the ratio of the determinants in (6-48) will all be close to 1, Λ will be near 1 and Box's M statistic will be small. If the null hypothesis is false, the sample covariance matrices can differ more and the differences in their determinants will be more pronounced. In this case Λ will be small and M will be relatively large. To illustrate, note that the determinant of the pooled covariance matrix, $|\mathbf{S}_{\text{pooled}}|$, will lie somewhere near the "middle" of the determinants $|\mathbf{S}_\ell|$'s of the individual group covariance matrices. As the latter quantities become more disparate, the product of the ratios in (6-44) will get closer to 0. In fact, as the $|\mathbf{S}_\ell|$'s increase in spread, $|\mathbf{S}_{(1)}|/|\mathbf{S}_{\text{pooled}}|$ reduces the product proportionally more than $|\mathbf{S}_{(g)}|/|\mathbf{S}_{\text{pooled}}|$ increases it, where $|\mathbf{S}_{(1)}|$ and $|\mathbf{S}_{(g)}|$ are the minimum and maximum determinant values, respectively.

Box's Test for Equality of Covariance Matrices

Set

$$u = \left[\sum_\ell \frac{1}{(n_\ell - 1)} - \frac{1}{\sum_\ell (n_\ell - 1)}\right]\left[\frac{2p^2 + 3p - 1}{6(p + 1)(g - 1)}\right] \tag{6-51}$$

where p is the number of variables and g is the number of groups. Then

$$C = (1 - u)M = (1 - u)\left\{\left[\sum_\ell (n_\ell - 1)\right]\ln|\mathbf{S}_{\text{pooled}}| - \sum_\ell [(n_\ell - 1)\ln|\mathbf{S}_\ell|]\right\} \tag{6-52}$$

has an approximate χ^2 distribution with

$$v = g\frac{1}{2}p(p + 1) - \frac{1}{2}p(p + 1) = \frac{1}{2}p(p + 1)(g - 1) \tag{6-53}$$

degrees of freedom. At significance level α, reject H_0 if $C > \chi^2_{p(p+1)(g-1)/2}(\alpha)$.

Box's χ^2 approximation works well if each n_ℓ exceeds 20 and if p and g do not exceed 5. In situations where these conditions do not hold, Box ([7], [8]) has provided a more precise F approximation to the sampling distribution of M.

Example 6.12 (Testing equality of covariance matrices—nursing homes) We introduced the Wisconsin nursing home data in Example 6.10. In that example the sample covariance matrices for $p = 4$ cost variables associated with $g = 3$ groups of nursing homes are displayed. Assuming multivariate normal data, we test the hypothesis $H_0: \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}_3 = \mathbf{\Sigma}$.

Using the information in Example 6.10, we have $n_1 = 271$, $n_2 = 138$, $n_3 = 107$ and $|\mathbf{S}_1| = 2.783 \times 10^{-8}$, $|\mathbf{S}_2| = 89.539 \times 10^{-8}$, $|\mathbf{S}_3| = 14.579 \times 10^{-8}$, and $|\mathbf{S}_{\text{pooled}}| = 17.398 \times 10^{-8}$. Taking the natural logarithms of the determinants gives $\ln|\mathbf{S}_1| = -17.397$, $\ln|\mathbf{S}_2| = -13.926$, $\ln|\mathbf{S}_3| = -15.741$ and $\ln|\mathbf{S}_{\text{pooled}}| = -15.564$. We calculate

$$u = \left[\frac{1}{270} + \frac{1}{137} + \frac{1}{106} - \frac{1}{270 + 137 + 106}\right]\left[\frac{2(4^2) + 3(4) - 1}{6(4 + 1)(3 - 1)}\right] = .0133$$

$$M = [270 + 137 + 106](-15.564) - [270(-17.397) + 137(-13.926) + 106(-15.741)]$$
$$= 289.3$$

and $C = (1 - .0133)289.3 = 285.5$. Referring C to a χ^2 table with $\nu = 4(4 + 1)(3 - 1)/2 = 20$ degrees of freedom, it is clear that H_0 is rejected at any reasonable level of significance. We conclude that the covariance matrices of the cost variables associated with the three populations of nursing homes are not the same. ■

Box's M-test is routinely calculated in many statistical computer packages that do MANOVA and other procedures requiring equal covariance matrices. It is known that the M-test is sensitive to some forms of non-normality. More broadly, in the presence of non-normality, normal theory tests on covariances are influenced by the kurtosis of the parent populations (see [16]). However, with reasonably large samples, the MANOVA tests of means or treatment effects are rather robust to nonnormality. Thus the M-test may reject H_0 in some non-normal cases where it is not damaging to the MANOVA tests. Moreover, with equal sample sizes, some differences in covariance matrices have little effect on the MANOVA tests. To summarize, we may decide to continue with the usual MANOVA tests even though the M-test leads to rejection of H_0.

6.7 Two-Way Multivariate Analysis of Variance

Following our approach to the one-way MANOVA, we shall briefly review the analysis for a *univariate* two-way fixed-effects model and then simply generalize to the multivariate case by analogy.

Univariate Two-Way Fixed-Effects Model with Interaction

We assume that measurements are recorded at various levels of two factors. In some cases, these experimental conditions represent levels of a single treatment arranged within several blocks. The particular experimental design employed will not concern us in this book. (See [10] and [17] for discussions of experimental design.) We shall, however, assume that observations at different combinations of experimental conditions are independent of one another.

Let the two sets of experimental conditions be the levels of, for instance, factor 1 and factor 2, respectively.[4] Suppose there are g levels of factor 1 and b levels of factor 2, and that n independent observations can be observed at each of the gb combinations of levels. Denoting the rth observation at level ℓ of factor 1 and level k of factor 2 by $X_{\ell kr}$, we specify the univariate two-way model as

$$X_{\ell kr} = \mu + \tau_\ell + \beta_k + \gamma_{\ell k} + e_{\ell kr}$$
$$\ell = 1, 2, \ldots, g$$
$$k = 1, 2, \ldots, b \qquad (6\text{-}54)$$
$$r = 1, 2, \ldots, n$$

where $\sum_{\ell=1}^{g} \tau_\ell = \sum_{k=1}^{b} \beta_k = \sum_{\ell=1}^{g} \gamma_{\ell k} = \sum_{k=1}^{b} \gamma_{\ell k} = 0$ and the $e_{\ell kr}$ are independent $N(0, \sigma^2)$ random variables. Here μ represents an overall level, τ_ℓ represents the fixed effect of factor 1, β_k represents the fixed effect of factor 2, and $\gamma_{\ell k}$ is the interaction between factor 1 and factor 2. The expected response at the ℓth level of factor 1 and the kth level of factor 2 is thus

$$E(X_{\ell kr}) = \mu + \tau_\ell + \beta_k + \gamma_{\ell k}$$
$$\begin{pmatrix}\text{mean}\\\text{response}\end{pmatrix} = \begin{pmatrix}\text{overall}\\\text{level}\end{pmatrix} + \begin{pmatrix}\text{effect of}\\\text{factor 1}\end{pmatrix} + \begin{pmatrix}\text{effect of}\\\text{factor 2}\end{pmatrix} + \begin{pmatrix}\text{factor 1-factor 2}\\\text{interaction}\end{pmatrix}$$
$$\ell = 1, 2, \ldots, g, \quad k = 1, 2, \ldots, b \qquad (6\text{-}55)$$

The presence of interaction, $\gamma_{\ell k}$, implies that the factor effects are not additive and complicates the interpretation of the results. Figures 6.3(a) and (b) show

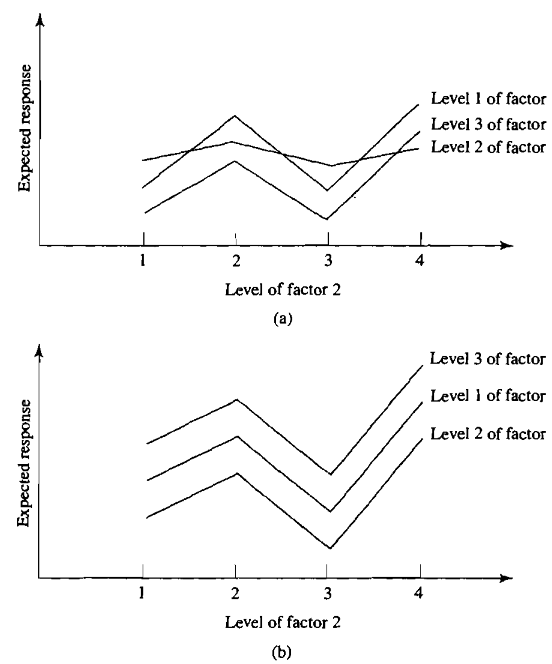

Figure 6.3 Curves for expected responses (a) with interaction and (b) without interaction.

[4] The use of the term "factor" to indicate an experimental condition is convenient. The factors discussed here should not be confused with the unobservable factors considered in Chapter 9 in the context of factor analysis.

expected responses as a function of the factor levels with and without interaction, respectively. The absense of interaction means $\gamma_{\ell k} = 0$ for all ℓ and k.

In a manner analogous to (6-55), each observation can be decomposed as

$$x_{\ell kr} = \bar{x} + (\bar{x}_{\ell\cdot} - \bar{x}) + (\bar{x}_{\cdot k} - \bar{x}) + (\bar{x}_{\ell k} - \bar{x}_{\ell\cdot} - \bar{x}_{\cdot k} + \bar{x}) + (x_{\ell kr} - \bar{x}_{\ell k}) \qquad (6\text{-}56)$$

where $\bar{x}$ is the overall average, $\bar{x}_{\ell\cdot}$ is the average for the ℓth level of factor 1, $\bar{x}_{\cdot k}$ is the average for the kth level of factor 2, and $\bar{x}_{\ell k}$ is the average for the ℓth level of factor 1 *and* the kth level of factor 2. Squaring and summing the deviations $(x_{\ell kr} - \bar{x})$ gives

$$\sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(x_{\ell kr} - \bar{x})^2 = \sum_{\ell=1}^{g} bn(\bar{x}_{\ell\cdot} - \bar{x})^2 + \sum_{k=1}^{b} gn(\bar{x}_{\cdot k} - \bar{x})^2 + \sum_{\ell=1}^{g}\sum_{k=1}^{b} n(\bar{x}_{\ell k} - \bar{x}_{\ell\cdot} - \bar{x}_{\cdot k} + \bar{x})^2 + \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(x_{\ell kr} - \bar{x}_{\ell k})^2 \qquad (6\text{-}57)$$

or

$$SS_{cor} = SS_{fac1} + SS_{fac2} + SS_{int} + SS_{res}$$

The corresponding degrees of freedom associated with the sums of squares in the breakup in (6-57) are

$$gbn - 1 = (g - 1) + (b - 1) + (g - 1)(b - 1) + gb(n - 1) \qquad (6\text{-}58)$$

The ANOVA table takes the following form:

ANOVA Table for Comparing Effects of Two Factors and Their Interaction

Source of variation	Sum of squares (SS)	Degrees of freedom (d.f.)
Factor 1	$SS_{fac1} = \sum_{\ell=1}^{g} bn(\bar{x}_{\ell\cdot} - \bar{x})^2$	$g - 1$
Factor 2	$SS_{fac2} = \sum_{k=1}^{b} gn(\bar{x}_{\cdot k} - \bar{x})^2$	$b - 1$
Interaction	$SS_{int} = \sum_{\ell=1}^{g}\sum_{k=1}^{b} n(\bar{x}_{\ell k} - \bar{x}_{\ell\cdot} - \bar{x}_{\cdot k} + \bar{x})^2$	$(g - 1)(b - 1)$
Residual (Error)	$SS_{res} = \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(x_{\ell kr} - \bar{x}_{\ell k})^2$	$gb(n - 1)$
Total (corrected)	$SS_{cor} = \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(x_{\ell kr} - \bar{x})^2$	$gbn - 1$

The F-ratios of the mean squares, $SS_{fac1}/(g - 1)$, $SS_{fac2}/(b - 1)$, and $SS_{int}/(g - 1)(b - 1)$ to the mean square, $SS_{res}/(gb(n - 1))$ can be used to test for the effects of factor 1, factor 2, and factor 1–factor 2 interaction, respectively. (See [11] for a discussion of univariate two-way analysis of variance.)

Multivariate Two-Way Fixed-Effects Model with Interaction

Proceeding by analogy, we specify the two-way fixed-effects model for a *vector* response consisting of p components [see (6-54)]

$$\begin{aligned} \mathbf{X}_{\ell kr} &= \boldsymbol{\mu} + \boldsymbol{\tau}_\ell + \boldsymbol{\beta}_k + \boldsymbol{\gamma}_{\ell k} + \mathbf{e}_{\ell kr} \\ \ell &= 1, 2, \ldots, g \\ k &= 1, 2, \ldots, b \\ r &= 1, 2, \ldots, n \end{aligned} \qquad (6\text{-}59)$$

where $\sum_{\ell=1}^{g} \boldsymbol{\tau}_\ell = \sum_{k=1}^{b} \boldsymbol{\beta}_k = \sum_{\ell=1}^{g} \boldsymbol{\gamma}_{\ell k} = \sum_{k=1}^{b} \boldsymbol{\gamma}_{\ell k} = \mathbf{0}$. The vectors are all of order $p \times 1$, and the $\mathbf{e}_{\ell kr}$ are independent $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ random vectors. Thus, the responses consist of p measurements replicated n times at each of the possible combinations of levels of factors 1 and 2.

Following (6-56), we can decompose the observation vectors $\mathbf{x}_{\ell kr}$ as

$$\mathbf{x}_{\ell kr} = \bar{\mathbf{x}} + (\bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}}) + (\bar{\mathbf{x}}_{\cdot k} - \bar{\mathbf{x}}) + (\bar{\mathbf{x}}_{\ell k} - \bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}}_{\cdot k} + \bar{\mathbf{x}}) + (\mathbf{x}_{\ell kr} - \bar{\mathbf{x}}_{\ell k}) \qquad (6\text{-}60)$$

where $\bar{\mathbf{x}}$ is the overall average of the observation vectors, $\bar{\mathbf{x}}_{\ell\cdot}$ is the average of the observation vectors at the ℓth level of factor 1, $\bar{\mathbf{x}}_{\cdot k}$ is the average of the observation vectors at the kth level of factor 2, and $\bar{\mathbf{x}}_{\ell k}$ is the average of the observation vectors at the ℓth level of factor 1 *and* the kth level of factor 2.

Straightforward generalizations of (6-57) and (6-58) give the breakups of the sum of squares and cross products and degrees of freedom:

$$\begin{aligned} \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(\mathbf{x}_{\ell kr} - \bar{\mathbf{x}})(\mathbf{x}_{\ell kr} - \bar{\mathbf{x}})' &= \sum_{\ell=1}^{g} bn(\bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}})(\bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}})' \\ &+ \sum_{k=1}^{b} gn(\bar{\mathbf{x}}_{\cdot k} - \bar{\mathbf{x}})(\bar{\mathbf{x}}_{\cdot k} - \bar{\mathbf{x}})' \\ &+ \sum_{\ell=1}^{g}\sum_{k=1}^{b} n(\bar{\mathbf{x}}_{\ell k} - \bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}}_{\cdot k} + \bar{\mathbf{x}})(\bar{\mathbf{x}}_{\ell k} - \bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}}_{\cdot k} + \bar{\mathbf{x}})' \\ &+ \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(\mathbf{x}_{\ell kr} - \bar{\mathbf{x}}_{\ell k})(\mathbf{x}_{\ell kr} - \bar{\mathbf{x}}_{\ell k})' \end{aligned} \qquad (6\text{-}61)$$

$$gbn - 1 = (g - 1) + (b - 1) + (g - 1)(b - 1) + gb(n - 1) \qquad (6\text{-}62)$$

Again, the generalization from the univariate to the multivariate analysis consists simply of replacing a scalar such as $(\bar{x}_{\ell\cdot} - \bar{x})^2$ with the corresponding matrix $(\bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}})(\bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}})'$.