Chapter 4
One-Way ANOVA

In this and the following chapters, we apply the general theory of linear models to various special cases. This chapter considers the analysis of one-way ANOVA models. A one-way ANOVA model can be written

$$y_{ij} = \mu + \alpha_i + e_{ij}, \qquad i = 1,\ldots,t, \quad j = 1,\ldots,N_i, \tag{1}$$

where $\mathrm{E}(e_{ij}) = 0$, $\mathrm{Var}(e_{ij}) = \sigma^2$, and $\mathrm{Cov}(e_{ij}, e_{i'j'}) = 0$ when $(i,j) \neq (i',j')$. For finding tests and confidence intervals, the e_{ij}s are assumed to have a multivariate normal distribution. Here α_i is an effect for y_{ij} belonging to the ith group of observations. Group effects are often called *treatment effects* because one-way ANOVA models are used to analyze completely randomized experimental designs.

Section 1 is devoted primarily to deriving the ANOVA table for a one-way ANOVA. The ANOVA table in this case is a device for presenting the sums of squares necessary for testing the reduced model

$$y_{ij} = \mu + e_{ij}, \qquad i = 1,\ldots,t, \quad j = 1,\ldots,N_i, \tag{2}$$

against model (1). This test is equivalent to testing the hypothesis $H_0 : \alpha_1 = \cdots = \alpha_t$.

The main tool needed for deriving the analysis of model (1) is the perpendicular projection operator. The first part of Section 1 is devoted to finding M. Since the ys in model (1) are identified with two subscripts, it will be necessary to develop notation that allows the rows of a vector to be denoted by two subscripts. Once M is found, some comments are made about estimation and the role of side conditions in estimation. Finally, the perpendicular projection operator for testing $H_0 : \alpha_1 = \cdots = \alpha_t$ is found and the ANOVA table is presented. Section 2 is an examination of contrasts. First, contrasts are defined and discussed. Estimation and testing procedures are presented. Orthogonal contrasts are defined and applications of Sections 3.5 and 3.6 are given. Fortunately, many balanced multifactor analysis of variance problems can be analyzed by repeatedly using the analysis for a one-way analysis of variance. For that reason, the results of this chapter are particularly important.

R. Christensen, *Plane Answers to Complex Questions: The Theory of Linear Models*,
Springer Texts in Statistics, DOI 10.1007/978-1-4419-9816-3_4,

4.1 Analysis of Variance

In linear model theory, the main tools we need are perpendicular projection matrices. Our first project in this section is finding the perpendicular projection matrix for a one-way ANOVA model. We will then discuss estimation, side conditions, and the ANOVA table.

Usually, the one-way ANOVA model is written

$$y_{ij} = \mu + \alpha_i + e_{ij}, \qquad i = 1, \ldots, t, \quad j = 1, \ldots, N_i.$$

Let $n = \sum_{i=1}^{t} N_i$. Although the notation N_i is standard, we will sometimes use $N(i)$ instead. Thus, $N(i) \equiv N_i$. We proceed to find the perpendicular projection matrix $M = X(X'X)^{-}X'$.

EXAMPLE 4.1.1. In any particular example, the matrix manipulations necessary for finding M are simple. Suppose $t = 3$, $N_1 = 5$, $N_2 = 3$, $N_3 = 3$. In matrix notation the model can be written

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{15} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{33} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + e.$$

To find the perpendicular projection matrix M, first find

$$X'X = \begin{bmatrix} 11 & 5 & 3 & 3 \\ 5 & 5 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 3 & 0 & 0 & 3 \end{bmatrix}.$$

By checking that $(X'X)(X'X)^{-}(X'X) = X'X$, it is easy to verify that

$$(X'X)^{-} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/5 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 \end{bmatrix}.$$

Then

$$
\begin{aligned}
M &= X(X'X)^-X' \\
&= X\begin{bmatrix}
0&0&0&0&0&0&0&0&0&0&0\\
1/5&1/5&1/5&1/5&1/5&0&0&0&0&0&0\\
0&0&0&0&0&1/3&1/3&1/3&0&0&0\\
0&0&0&0&0&0&0&0&1/3&1/3&1/3
\end{bmatrix}\\
&= \begin{bmatrix}
1/5&1/5&1/5&1/5&1/5&0&0&0&0&0&0\\
1/5&1/5&1/5&1/5&1/5&0&0&0&0&0&0\\
1/5&1/5&1/5&1/5&1/5&0&0&0&0&0&0\\
1/5&1/5&1/5&1/5&1/5&0&0&0&0&0&0\\
1/5&1/5&1/5&1/5&1/5&0&0&0&0&0&0\\
0&0&0&0&0&1/3&1/3&1/3&0&0&0\\
0&0&0&0&0&1/3&1/3&1/3&0&0&0\\
0&0&0&0&0&1/3&1/3&1/3&0&0&0\\
0&0&0&0&0&0&0&0&1/3&1/3&1/3\\
0&0&0&0&0&0&0&0&1/3&1/3&1/3\\
0&0&0&0&0&0&0&0&1/3&1/3&1/3
\end{bmatrix}.
\end{aligned}
$$

Thus, in this example, M is Blk diag$[N_i^{-1}J_{N(i)}^{N(i)}]$, where J_r^c is a matrix of 1s with r rows and c columns. In fact, we will see below that this is the general form for M in a one-way ANOVA when the observation vector Y has subscripts changing fastest on the right.

A somewhat easier way of finding M is as follows. Let Z be the model matrix for the alternative one-way analysis of variance model

$$y_{ij} = \mu_i + e_{ij},$$

$i = 1,\dots,t$, $j = 1,\dots,N_i$. (See Example 3.1.1.) Z is then just a matrix consisting of the last t columns of X, i.e., $X = [J,Z]$. Clearly $C(X) = C(Z)$, $Z'Z = \text{Diag}(N_1,N_2,\dots,N_t)$, and $(Z'Z)^{-1} = \text{Diag}(N_1^{-1},N_2^{-1},\dots,N_t^{-1})$. It is easy to see that $Z(Z'Z)^-Z' = \text{Blk diag}[N_i^{-1}J_{N(i)}^{N(i)}]$.

We now present a rigorous derivation of these results. The ideas involved in Example 4.1.1 are perfectly general. A similar computation can be performed for any values of t and the N_is. The difficulty in a rigorous general presentation lies entirely in being able to write down the model in matrix form. The elements of Y are the y_{ij}s. The y_{ij}s have two subscripts, so a pair of subscripts must be used to specify each row of the vector Y. The elements of the model matrices X and Z are determined entirely by knowing the order in which the y_{ij}s have been listed in the Y vector. For example, the row of Z corresponding to y_{ij} would have a 1 in the ith column and 0s everywhere else. Clearly, it will also be convenient to use a pair of subscripts to specify the rows of the model matrices. Specifically, let

$$Y' = (y_{11}, y_{12}, \dots, y_{1N(1)}, y_{21}, \dots, y_{tN(t)}),$$

where y_{21} is the $N_1 + 1$ row of Y. In general, write a vector S as $S = [s_{ij}]$, where for $j = 1, \dots, N_i$, s_{ij} denotes the $(N_1 + \cdots + N_{i-1} + j)$th row of S. The discussion of tensors in Appendix B may help the reader feel more comfortable with our use of subscripts.

To specify the model matrices X and Z we must identify the columns of X and Z. Write $X = [J, X_1, X_2, \dots, X_t]$ and $Z = [X_1, X_2, \dots, X_t]$. Note that the kth column of Z can be written

$$X_k = [t_{ij}], \qquad \text{where } t_{ij} = \delta_{ik} \tag{1}$$

with δ_{ik} equal to 0 if $i \neq k$ and 1 if $i = k$. This means that if the observation in the ij row belongs to the kth group, the ij row of X_k is 1. If not, the ij row of X_k is zero.

Our goal is to find $M = Z(Z'Z)^{-1}Z'$. To do this we need to find $(Z'Z)$ and $(Z'Z)^{-1}$. Noting that $(X_k)'(X_q)$ is a real number, we can write the elements of $Z'Z$ as

$$(Z'Z) = \left[(X_k)'(X_q)\right]_{t\times t}.$$

Now, from (1)

$$(X_k)'(X_k) = \sum_{ij} \delta_{ik}\delta_{ik} = \sum_{i=1}^{t}\sum_{j=1}^{N_i} \delta_{ik} = \sum_{i=1}^{t} N_i\delta_{ik} = N_k$$

and for $k \neq q$

$$(X_k)'(X_q) = \sum_{i=1}^{t}\sum_{j=1}^{N_i} \delta_{ik}\delta_{iq} = \sum_{i=1}^{t} N_i\delta_{ik}\delta_{iq} = 0.$$

It follows that

$$(Z'Z) = \mathrm{Diag}(N_i)$$

and clearly

$$(Z'Z)^{-1} = \mathrm{Diag}(N_i^{-1}).$$

We can now find $Z(Z'Z)^{-1}$.

$$\begin{aligned} Z(Z'Z)^{-1} &= [X_1, X_2, \dots, X_t]\mathrm{Diag}(N_i^{-1}) \\ &= \left[N_1^{-1}X_1, N_2^{-1}X_2, \dots, N_t^{-1}X_t\right]. \end{aligned}$$

Finally, we are in a position to find $M = Z(Z'Z)^{-1}Z'$. We denote the columns of an $n \times n$ matrix using the convention introduced above for denoting rows, i.e., by using two subscripts. Then the matrix M can be written

$$M = [m_{ij,i'j'}].$$

We now find the entries of this matrix. Note that $m_{ij,i'j'}$ is the ij row of $Z(Z'Z)^{-1}$ times the $i'j'$ column of Z' (i.e., the $i'j'$ row of Z). The ij row of $Z(Z'Z)^{-1}$ is $(N_1^{-1}\delta_{i1}, \dots, N_t^{-1}\delta_{it})$. The $i'j'$ row of Z is $(\delta_{i'1}, \dots, \delta_{i't})$. The product is

$$
\begin{aligned}
m_{ij,i'j'} &= \sum_{k=1}^{t} N_k^{-1}\delta_{ik}\delta_{i'k} \\
&= N_i^{-1}\delta_{ii'}.
\end{aligned}
$$

These values of $m_{ij,i'j'}$ determine a block diagonal matrix

$$M = \text{Blk diag}(N_i^{-1}J_{N(i)}^{N(i)}),$$

just as in Example 4.1.1.

The notation and methods developed above are somewhat unusual, but they are necessary for giving a rigorous treatment of ANOVA models. The device of indicating the rows of vectors with multiple subscripts will be used extensively in later discussions of multifactor ANOVA. It should be noted that the arguments given above really apply to any order of specifying the entries in a vector $S = [s_{ij}]$; they do not really depend on having $S = (s_{11}, s_{12}, \ldots, s_{tN(t)})'$. If we specified some other ordering, we would still get the perpendicular projection matrix M; however, M might no longer be block diagonal.

Exercise 4.1 To develop some facility with this notation, let

$$T_r = [t_{ij}], \qquad \text{where } t_{ij} = \delta_{ir} - \frac{N_r}{n}$$

for $r = 1, \ldots, t$. Find $T_r'T_r$, $T_r'T_s$ for $s \neq r$, and $J'T_r$.

A very important application of this notation is in characterizing the vector MY. As discussed in Section 3.1, the vector MY is the base from which all estimates of parametric functions are found. A second important application involves the projection operator

$$M_\alpha = M - \frac{1}{n}J_n^n .$$

M_α is useful in testing hypotheses and is especially important in the analysis of multifactor ANOVAs. It is therefore necessary to have a characterization of M_α.

Exercise 4.2 Show that

$$MY = [t_{ij}], \qquad \text{where } t_{ij} = \bar{y}_{i\cdot}$$

and

$$M_\alpha Y = [u_{ij}], \qquad \text{where } u_{ij} = \bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot} .$$

Hint: Write $M_\alpha Y = MY - (\frac{1}{n}J_n^n)Y$.

Knowing these characterizations MY and $M_\alpha Y$ tell us how to find Mv and $M_\alpha v$ for any vector v. In fact, they completely characterize the perpendicular projection operators M and M_α.

EXAMPLE 4.1.1 CONTINUED. In this example,

$$MY = (\bar{y}_{1\cdot}, \bar{y}_{1\cdot}, \bar{y}_{1\cdot}, \bar{y}_{1\cdot}, \bar{y}_{1\cdot}, \bar{y}_{2\cdot}, \bar{y}_{2\cdot}, \bar{y}_{2\cdot}, \bar{y}_{3\cdot}, \bar{y}_{3\cdot}, \bar{y}_{3\cdot})'$$

and

$$M_\alpha Y = (\bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot}, \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot}, \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot}, \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot}, \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot},$$
$$\bar{y}_{2\cdot} - \bar{y}_{\cdot\cdot}, \bar{y}_{2\cdot} - \bar{y}_{\cdot\cdot}, \bar{y}_{2\cdot} - \bar{y}_{\cdot\cdot}, \bar{y}_{3\cdot} - \bar{y}_{\cdot\cdot}, \bar{y}_{3\cdot} - \bar{y}_{\cdot\cdot}, \bar{y}_{3\cdot} - \bar{y}_{\cdot\cdot})'.$$

We can now obtain a variety of estimates. Recall that estimable functions are linear combinations of the rows of $X\beta$, e.g., $\rho' X\beta$. Since

$$X\beta = \big(\mu+\alpha_1, \mu+\alpha_1, \mu+\alpha_1, \mu+\alpha_1, \mu+\alpha_1,$$
$$\mu+\alpha_2, \mu+\alpha_2, \mu+\alpha_2, \mu+\alpha_3, \mu+\alpha_3, \mu+\alpha_3\big)',$$

if ρ' is taken to be $\rho' = (1,0,0,0,0,0,0,0,0,0,0)'$, then it is easily seen that $\mu+\alpha_1 = \rho' X\beta$ is estimable. The estimate of $\mu+\alpha_1$ is $\rho' MY = \bar{y}_{1\cdot}$. Similarly, the estimates of $\mu+\alpha_2$ and $\mu+\alpha_3$ are $\bar{y}_{2\cdot}$ and $\bar{y}_{3\cdot}$, respectively. The contrast $\alpha_1 - \alpha_2$ can be obtained as $\rho' X\beta$ using $\rho' = (1,0,0,0,0,-1,0,0,0,0,0)'$. The estimate of $\alpha_1 - \alpha_2$ is $\rho' MY = \bar{y}_{1\cdot} - \bar{y}_{2\cdot}$. Note that for this contrast $\rho' MY = \rho' M_\alpha Y$.

Estimation is as easy in a general one-way ANOVA as it is in Example 4.1.1. We have found M and MY, and it is an easy matter to see that, for instance, $\mu+\alpha_i$ is estimable and the estimate of $\mu+\alpha_i$ is

$$\{\hat{\mu} + \hat{\alpha}_i\} = \bar{y}_{i\cdot}\,.$$

The notation $\{\hat{\mu} + \hat{\alpha}_i\}$ will be used throughout this chapter to denote the estimate of $\mu+\alpha_i$.

For computational purposes, it is often convenient to present one particular set of least squares estimates. In one-way ANOVA, the traditional *side condition* on the parameters is $\sum_{i=1}^t N_i \alpha_i = 0$. With this condition, one obtains

$$\mu = \frac{1}{n}\sum_{i=1}^t N_i(\mu+\alpha_i)$$

and an estimate

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^t N_i\{\hat{\mu}+\hat{\alpha}_i\} = \frac{1}{n}\sum_{i=1}^t \frac{N_i}{N_i}\sum_{j=1}^{N_i} y_{ij} = \bar{y}_{\cdot\cdot},$$

which is the mean of all the observations. Similarly,

$$\hat{\alpha}_i = \{\hat{\mu}+\hat{\alpha}_i\} - \hat{\mu} = \bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}\,.$$

Exercise 4.9 involves finding parameter estimates that satisfy a different commonly used side condition. Fortunately, all side conditions lead to the same estimates of identifiable functions, so one choice of a side condition is as good as any other. The best choice of a side condition is the most convenient choice. However, different side conditions do lead to different "estimates" of nonidentifiable parameters. Do not be lulled into believing that an arbitrary side condition allows you to say anything meaningful about nonidentifiable parameters. That is just silly!

We now derive the analysis of variance table for the one-way ANOVA. The analysis of variance table is a device for displaying an orthogonal breakdown of the *total sum of squares* of the data (*SSTot*), i.e., $Y'Y$. Sometimes, the total sum of squares corrected for fitting the *grand mean* is broken down. The *sum of squares for fitting the grand mean* (*SSGM*), also known as the correction factor (*C*), is just the sum of squares accounted for by the model

$$Y = J\mu + e.$$

The *total sum of squares corrected for the grand mean* ($SSTot - C$) is the error sum of squares for this model, i.e., $Y'(I - [1/n]J_n^n)Y$. Included in an ANOVA table is information to identify the sums of squares (Source), the degrees of freedom for the sums of squares (*df*), the sums of squares (*SS*), and the mean squares (*MS*). The mean squares are just the sums of squares divided by their degrees of freedom. Sometimes the expected values of the mean squares are included. From the expected mean squares, the hypotheses tested by the various sums of squares can be identified. Recall that, when divided by σ^2, the sums of squares have χ^2 distributions and that there is a very close relationship between the expected mean square, the expected sum of squares, the noncentrality parameter of the χ^2 distribution, and the noncentrality parameter of an F distribution with the mean square in the numerator. In particular, if the expected mean square is $\sigma^2 + \pi/df$, then the noncentrality parameter is $\pi/2\sigma^2$. Assuming the full model is true, the null hypothesis being tested is that the noncentrality parameter of the F distribution is zero.

The usual orthogonal breakdown for a one-way ANOVA is to isolate the effect of the grand mean (μ), and then the effect of fitting the treatments ($\alpha_i s$) after fitting the mean. The *sum of squares for treatments* (*SSTrts*) is just what is left after removing the sum of squares for μ from the sum of squares for the model. In other words, the sum of squares for treatments is the sum of squares for testing the reduced model (4.0.2) against model (4.0.1). As we have seen earlier, the projection operator for fitting the grand mean is based on the first column of X, i.e., J. The projection operator is $(1/n)J_n^n = (1/n)JJ'$. The projection operator for the treatment sum of squares is then

$$M_\alpha = M - \frac{1}{n}J_n^n.$$

The sum of squares for fitting treatments after μ, $Y'M_\alpha Y$, is the difference between the sum of squares for fitting the full model, $Y'MY$, and the sum of squares for fitting the model with just the mean, $Y'([1/n]J_n^n)Y$.

Table 1 gives an ANOVA table and indicates some common notation for the entries.

Table 4.1 One-Way Analysis of Variance Table

Matrix Notation		
Source	df	SS
Grand Mean	1	$Y'\left(\frac{1}{n}J_n^n\right)Y$
Treatments	$t-1$	$Y'\left(M-\frac{1}{n}J_n^n\right)Y$
Error	$n-t$	$Y'(I-M)Y$
Total	n	$Y'Y$
Source	SS	$\mathrm{E}(MS)$
Grand Mean	$SSGM$	$\sigma^2+\beta'X'\left(\frac{1}{n}J_n^n\right)X\beta$
Treatments	$SSTrts$	$\sigma^2+\beta'X'\left(M-\frac{1}{n}J_n^n\right)X\beta/(t-1)$
Error	SSE	σ^2
Total	$SSTot$	

Algebraic Notation		
Source	df	SS
Grand Mean	$dfGM$	$n^{-1}y_{\cdot\cdot}^2 = n\bar{y}_{\cdot\cdot}^2$
Treatments	$dfTrts$	$\sum_{i=1}^t N_i\left(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot}\right)^2$
Error	dfE	$\sum_{i=1}^t\sum_{j=1}^{N_i}\left(y_{ij}-\bar{y}_{i\cdot}\right)^2$
Total	$dfTot$	$\sum_{i=1}^t\sum_{j=1}^{N_i}y_{ij}^2$
Source	MS	$\mathrm{E}(MS)$*
Grand Mean	$SSGM$	$\sigma^2+n(\mu+\bar{\alpha}_\cdot)^2$
Treatments	$SSTrts/(t-1)$	$\sigma^2+\sum_{i=1}^t N_i\left(\alpha_i-\bar{\alpha}_\cdot\right)^2/(t-1)$
Error	$SSE/(n-t)$	σ^2
Total		

$*\bar{\alpha}_\cdot = \sum_{i=1}^t N_i\alpha_i/n$

Exercise 4.3 Verify that the estimate of $\mu + \alpha_i$ is $\bar{y}_{i\cdot}$ and that the algebraic formulas for the sums of squares in the ANOVA table are correct.

Hint: To find, for example, $Y(M - [1/n]J_n^n)Y = Y'M_\alpha Y$, use Exercise 4.2 to get $M_\alpha Y$ and recall that $Y'M_\alpha Y = [M_\alpha Y]'[M_\alpha Y]$.

Exercise 4.4 Verify that the formulas for expected mean squares in the ANOVA table are correct.

Hint: Use Theorem 1.3.2 and Exercise 4.3.

The techniques suggested for Exercises 4.3 and 4.4 are very useful. The reader should make a point of remembering them.

4.2 Estimating and Testing Contrasts

In this section, contrasts are defined and characterizations of contrasts are given. Estimates of contrasts are found. The numerator sum of squares necessary for doing an F test of a contrast and the form of the F test are given. The form of a confidence interval for a contrast is presented and the idea of orthogonal contrasts is discussed. Finally, the results of Section 3.5 are reviewed by deriving them anew for the one-way ANOVA model.

A contrast in the one-way ANOVA (4.0.1) is a function $\sum_{i=1}^t \lambda_i \alpha_i$, with $\sum_{i=1}^t \lambda_i = 0$. In other words, the vector λ' in $\lambda'\beta$ is $(0, \lambda_1, \lambda_2, \ldots, \lambda_t)$ and $\lambda' J_{t+1} = 0$. To establish that $\lambda'\beta$ is estimable, we need to find ρ such that $\rho' X = \lambda'$. Write

$$\rho' = (\lambda_1/N_1, \ldots, \lambda_1/N_1, \lambda_2/N_2, \ldots, \lambda_2/N_2, \lambda_3/N_3, \ldots, \lambda_t/N_t),$$

where ρ' is a $1 \times n$ vector and the string of λ_i/N_is is N_i long. In the alternate notation that uses two subscripts to denote a row of a vector, we have

$$\rho = [t_{ij}], \qquad \text{where } t_{ij} = \lambda_i/N_i. \tag{1}$$

Recall from Section 2.1 that, while other choices of ρ may exist with $\rho' X = \lambda'$, the vector $M\rho$ is unique. As shown in Exercise 4.10, for ρ as in (1), $\rho \in C(X)$; so $\rho = M\rho$. Thus, for any contrast $\lambda'\beta$, the vector $M\rho$ has the structure

$$M\rho = [t_{ij}], \qquad \text{where } t_{ij} = \lambda_i/N_i. \tag{2}$$

We now show that the contrasts are precisely the estimable functions that do not involve μ. Note that since J is the column of X associated with μ, $\rho' X\beta$ does not involve μ if and only if $\rho' J = 0$.

Proposition 4.2.1. $\rho' X\beta$ is a contrast if and only if $\rho' J = 0$.

PROOF. Clearly, a contrast does not involve μ, so $\rho' J = 0$. Conversely, if $\rho' J = 0$, then $\rho' X\beta = \rho'[J, Z]\beta$ does not involve μ; so we need only show that $0 = \rho' X J_{t+1}$. This follows because $XJ_{t+1} = 2J_n$, and we know that $\rho' J_n = 0$. □

We now show that the contrasts are the estimable functions that impose constraints on $C(M_\alpha)$. Recall that the constraint imposed on $C(X)$ by $\rho' X\beta = 0$ is that $\mathrm{E}(Y) \in C(X)$ and $\mathrm{E}(Y) \perp M\rho$, i.e., $\mathrm{E}(Y)$ is constrained to be orthogonal to $M\rho$. By definition, $\rho' X\beta$ puts a constraint on $C(M_\alpha)$ if $M\rho \in C(M_\alpha)$.

Proposition 4.2.2. $\rho' X\beta$ is a contrast if and only if $M\rho \in C(M_\alpha)$.

PROOF. Using Proposition 4.2.1 and $J \in C(X)$, we see that $\rho' X\beta$ is a contrast if and only if $0 = \rho' J = \rho' MJ$, i.e., $J \perp M\rho$. However, $C(M_\alpha)$ is everything in $C(X)$ that is orthogonal to J; thus $J \perp M\rho$ if and only if $M\rho \in C(M_\alpha)$. □

Finally, we can characterize $C(M_\alpha)$.

Proposition 4.2.3. $C(M_\alpha) = \left\{\rho \middle| \rho = [t_{ij}],\ t_{ij} = \lambda_i/N_i,\ \sum_{i=1}^t \lambda_i = 0\right\}$.

PROOF. Any vector ρ with the structure of (1) and $\sum_i \lambda_i = 0$ has $\rho' J = 0$ and by Proposition 4.2.1 determines a contrast $\rho' X\beta$. By Proposition 4.2.2, $M\rho \in C(M_\alpha)$. However, vectors that satisfy (1) also satisfy $M\rho = \rho$, so $\rho \in C(M_\alpha)$. Conversely, if $\rho \in C(M_\alpha)$, then $\rho' J = 0$; so $\rho' X\beta$ determines a contrast. It follows that $M\rho$ must be of the form (2), where $\lambda_1 + \cdots + \lambda_t = 0$. However, since $\rho \in C(M_\alpha)$, $M\rho = \rho$; so ρ must be of the form (1) with $\lambda_1 + \cdots + \lambda_t = 0$.

Exercise 4.5 Show that $\alpha_1 = \alpha_2 = \cdots = \alpha_t$ if and only if all contrasts are zero.

We now consider estimation and testing for contrasts. The least squares estimate of a contrast $\sum_{i=1}^t \lambda_i \alpha_i$ is easily obtained. Let $\hat{\mu}, \hat{\alpha}_1, \ldots, \hat{\alpha}_t$ be any choice of least squares estimates for the nonidentifiable parameters $\mu, \alpha_1, \ldots, \alpha_t$. Since $\sum_{i=1}^t \lambda_i = 0$, we can write

$$\sum_{i=1}^t \lambda_i \hat{\alpha}_i = \sum_{i=1}^t \lambda_i \{\hat{\mu} + \hat{\alpha}_i\} = \sum_{i=1}^t \lambda_i \bar{y}_{i\cdot},$$

because $\mu + \alpha_i$ is estimable and its unique least squares estimate is $\bar{y}_{i\cdot}$. This result can also be seen by examining $\rho' Y = \rho' MY$ for the ρ given earlier in (1). To test the hypothesis that $\lambda' \beta = 0$, we have seen that the numerator of the F test statistic is $(\rho' MY)^2/\rho' M\rho$. However, $\rho' MY = \rho' X\hat{\beta} = \lambda' \hat{\beta} = \sum_{i=1}^t \lambda_i \bar{y}_{i\cdot}$. We also need to find $\rho' M\rho$. The easiest way is to observe that, since $M\rho$ has the structure of (2),

$$\begin{aligned} \rho' M\rho &= [M\rho]'[M\rho] \\ &= \sum_{i=1}^t \sum_{j=1}^{N_i} \lambda_i^2/N_i^2 \end{aligned}$$

$$= \sum_{i=1}^{t} \lambda_i^2 / N_i.$$

The numerator sum of squares for testing the contrast is

$$SS(\lambda'\beta) \equiv \left(\sum_{i=1}^{t} \lambda_i \bar{y}_{i\cdot}\right)^2 \Bigg/ \left(\sum_{i=1}^{t} \lambda_i^2 / N_i\right).$$

The α level test for $H_0 : \sum_{i=1}^{t} \lambda_i \alpha_i = 0$ is to reject H_0 if

$$\frac{\left(\sum_{i=1}^{t} \lambda_i \bar{y}_{i\cdot}\right)^2 \Big/ \left(\sum_{i=1}^{t} \lambda_i^2 / N_i\right)}{MSE} > F(1-\alpha, 1, dfE).$$

Equivalently, $\sum_{i=1}^{t} \lambda_i \bar{y}_{i\cdot}$ has a normal distribution, $\mathrm{E}(\sum_{i=1}^{t} \lambda_i \bar{y}_{i\cdot}) = \sum_{i=1}^{t} \lambda_i(\mu + \alpha_i) = \sum_{i=1}^{t} \lambda_i \alpha_i$, and $\mathrm{Var}(\sum_{i=1}^{t} \lambda_i \bar{y}_{i\cdot}) = \sum_{i=1}^{t} \lambda_i^2 \mathrm{Var}(\bar{y}_{i\cdot}) = \sigma^2 \sum_{i=1}^{t} \lambda_i^2 / N_i$, so we have a t test available. The α level test is to reject H_0 if

$$\frac{\left|\sum_{i=1}^{t} \lambda_i \bar{y}_{i\cdot}\right|}{\sqrt{MSE \sum_{i=1}^{t} \lambda_i^2 / N_i}} > t\left(1 - \frac{\alpha}{2}, dfE\right).$$

Note that since $\sum_{i=1}^{t} \lambda_i \bar{y}_{i\cdot} = \rho' MY$ is a function of MY and MSE is a function of $Y'(I-M)Y$, we have the necessary independence for the t test. In fact, all tests and confidence intervals follow as in Exercise 2.1.

In order to break up the sums of squares for treatments into $t-1$ orthogonal single degree of freedom sums of squares, we need to find $t-1$ contrasts $\lambda_1'\beta, \ldots, \lambda_{t-1}'\beta$ with the property that $\rho_r' M \rho_s = 0$ for $r \neq s$, where $\rho_r' X = \lambda_r'$ (see Section 3.6). Let $\lambda_r' = (0, \lambda_{r1}, \ldots, \lambda_{rt})$ and recall that $M\rho_r$ has the structure of (2). The condition required is

$$\begin{aligned} 0 &= \rho_r' M \rho_s \\ &= [M\rho_r]'[M\rho_s] \\ &= \sum_{i=1}^{t} \sum_{j=1}^{N_i} (\lambda_{ri}/N_i)(\lambda_{si}/N_i) \\ &= \sum_{i=1}^{t} \lambda_{ri}\lambda_{si}/N_i. \end{aligned}$$

With any set of contrasts $\sum_{i=1}^{t} \lambda_{ri}\alpha_i$, $r = 1, \ldots, t-1$, for which $0 = \sum_{i=1}^{t} \lambda_{ri}\lambda_{si}/N_i$ for all $r \neq s$, we have a set of $t-1$ orthogonal constraints on the test space so that the sums of squares for the contrasts add up to the $SSTrts$. Contrasts that determine orthogonal constraints are referred to as *orthogonal contrasts*.

In later analyses, we will need to use the fact that the analysis developed here depends only on the projection matrix onto the space for testing $\alpha_1 = \alpha_2 = \cdots = \alpha_t$. That projection matrix is $M_\alpha = M - (1/n)J_n^n$. Note that $M = (1/n)J_n^n + M_\alpha$. For any contrast $\lambda'\beta$ with $\rho'X = \lambda'$, we know that $\rho'J_n = 0$. It follows that $\rho'M = \rho'(1/n)J_n^n + \rho'M_\alpha = \rho'M_\alpha$. There are two main uses for this fact. First,

$$\sum_{i=1}^{t} \lambda_i \hat{\alpha}_i = \sum_{i=1}^{t} \lambda_i \bar{y}_{i\cdot} = \rho'MY = \rho'M_\alpha Y,$$

$$\sum_{i=1}^{t} \lambda_i^2/N_i = \lambda'(X'X)^- \lambda = \rho'M\rho = \rho'M_\alpha \rho,$$

so estimation, and therefore tests, depend only on the projection M_α. Second, the condition for contrasts to give an orthogonal breakdown of $SSTrts$ is

$$0 = \sum_{i=1}^{t} \lambda_{ri}\lambda_{si}/N_i = \rho_r'M\rho_s = \rho_r'M_\alpha \rho_s,$$

which depends only on M_α. This is just a specific example of the theory of Section 3.5.

Exercise 4.6 Using the theories of Sections 3.3 and 2.6, respectively, find the F test and the t test for the hypothesis $H_0 : \sum_{i=1}^{t} \lambda_i \alpha_i = d$ in terms of the MSE, the $\bar{y}_{i\cdot}$s, and the λ_is.

Exercise 4.7 Suppose $N_1 = N_2 = \cdots = N_t \equiv N$. Rewrite the ANOVA table incorporating any simplifications due to this assumption.

Exercise 4.8 If $N_1 = N_2 = \cdots = N_t \equiv N$, show that two contrasts $\lambda_1'\beta$ and $\lambda_2'\beta$ are orthogonal if and only if $\lambda_1'\lambda_2 = 0$.

Exercise 4.9 Find the least squares estimates of μ, α_1, and α_t using the side condition $\alpha_1 = 0$.

Exercise 4.10 Using ρ as defined by (1) and X as defined in Section 1, especially (4.1.1), show that

(a) $\rho'X = \lambda'$, where $\lambda' = (0, \lambda_1, \ldots, \lambda_t)$.

(b) $\rho \in C(X)$.

4.3 Additional Exercises

Exercise 4.3.1 An experiment was conducted to see which of four brands of blue jeans were most resistant to wearing out as a result of students kneeling before their linear models instructor begging for additional test points. In a class of 32 students, 8 students were randomly assigned to each brand of jeans. Before being informed of their test score, each student was required to fall to his/her knees and crawl 3 meters to the instructor's desk. This was done after each of 5 mid-quarter and 3 final exams. (The jeans were distributed along with each of the 8 test forms and were collected again 36 hours after grades were posted.) A fabric wear score was determined for each pair of jeans. The scores are listed below.

Brand 1:	3.41	1.83	2.69	2.04	2.83	2.46	1.84	2.34
Brand 2:	3.58	3.83	2.64	3.00	3.19	3.57	3.04	3.09
Brand 3:	3.32	2.62	3.92	3.88	2.50	3.30	2.28	3.57
Brand 4:	3.22	2.61	2.07	2.58	2.80	2.98	2.30	1.66

(a) Give an ANOVA table for these data, and perform and interpret the F test for the differences between brands.

(b) Brands 2 and 3 were relatively inexpensive, while Brands 1 and 4 were very costly. Based on these facts, determine an appropriate set of orthogonal contrasts to consider in this problem. Find the sums of squares for the contrasts.

(c) What conclusions can be drawn from these data? Perform any additional computations that may be necessary

Exercise 4.3.2 After the final exam of spring quarter, 30 of the subjects of the previous experiment decided to test the sturdiness of 3 brands of sport coats and 2 brands of shirts. In this study, sturdiness was measured as the length of time before tearing when the instructor was hung by his collar out of his second-story office window. Each brand was randomly assigned to 6 students, but the instructor was occasionally dropped before his collar tore, resulting in some missing data. The data are listed below.

Coat 1:	2.34	2.46	2.83	2.04	2.69	
Coat 2:	2.64	3.00	3.19	3.83		
Coat 3:	2.61	2.07	2.80	2.58	2.98	2.30
Shirt 1:	1.32	1.62	1.92	0.88	1.50	1.30
Shirt 2:	0.41	0.83	0.53	0.32	1.62	

(a) Give an ANOVA table for these data, and perform and interpret the F test for the differences between brands.

(b) Test whether, on average, these brands of coats are sturdier than these brands of shirts.

(c) Give three contrasts that are mutually orthogonal and orthogonal to the contrast used in (b). Compute the sums of squares for all four contrasts.

(d) Give a 95% confidence interval for the difference in sturdiness between shirt Brands 1 and 2. Is one brand significantly sturdier than the other?