Estimable Functions and Testable Hypotheses in Linear Models

BU-213-M S. R. Searle April, 1966

ABSTRACT

The concept of an estimable function is restated, and the relationship between estimable functions and testable hypotheses discussed. The place of restrictions (constraints) on the elements of the model is also considered. The technique of solving the normal equations by means of a generalized inverse matrix is used throughout.

* Biometrics Unit, Plant Breeding Department, Cornell University.

Estimable Functions and Testable Hypotheses in Linear Models

BU-213-M S. R. Searle April, 1966

Estimable functions

Testable hypotheses are closely related to estimable functions so we begin with a resumé of the latter. At all times we consider linear functions only, we deal solely with the linear model and are concerned with only the fixed effects case thereof.

The equation of the linear model can be written as

$$y = Xb + e \qquad --- (1)$$

where y is a vector of n observations b is a vector of the parameters of the model (k of them), X is a design matrix having rank r ($r \leq k$), and e is a vector of random error terms having zero means, $E(e) = 0$, and variance-covariance matrix $E(ee') = \sigma^2 I$. The symbol E denotes expectation.

The normal equations resulting from the least squares procedure for fitting this model are

$$X'X\hat{b} = X'y \qquad --- (2)$$

where $\hat{b}$ is the solution corresponding to the parameter vector b. With X'X having rank less than its order, $r < k$, there is an infinite number of solutions $\hat{b}$ to these equations. Attention is therefore directed not to the solutions themselves but to linear functions of their elements.

Consider a linear function q'b of the parameters in b, where q' is a known vector. It is defined as being an estimable function if there exists some linear combination of the observations $y_1, y_2, \ldots, y_n$ whose expected value is q'b; i.e. if there exists a vector t' such that the expected value of t'y is q'b, then q'b is said to be estimable. It is called an estimable function.

* Biometrics Unit, Plant Breeding Department, Cornell University.

Now in general parlance "estimable" means "can be estimated" and yet in this context the definition of "estimable function" just given does not seem altogether relevant. However, four theorems stemming from the definition bring to light its importance. Numbers 2, 3 and 4 correspond approximately to theorems 11.1, 11.3 and 11.2 of Graybill (1961).

Theorem 1. The function $q'b$ is estimable if and only if there exists a vector t such that $q' = t'X$.

Proof. If $q'b$ is estimable then, by definition, there exists a vector t' such that $E(t'y) = q'b$. Therefore

$$q'b = E(t'y) = t'E(y) = t'E(Xb + e) = t'Xb$$

and since this is true for all b, $q' = t'X$. Conversely, if $q' = t'X$ then

$$q'b = t'Xb = t'E(y) = E(t'y)$$

and hence $q'b$ is estimable.

Theorem 2. The function $q'b$ is estimable if and only if the equation $X'Xu = q$ has a solution for u.

Proof (i). If $q'b$ is estimable then t exists such that $X't = q$. Therefore the equations $X'w = q$ have the solution $w = t$ and hence (using $r(A)$ to denote the rank of A)

$$r(X'X) = r(X') = r(X' \quad q) .$$

But
$$(X'X \quad q) \equiv \left[(X' \quad q) \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} \right] ,$$

so that
$$r(X'X \quad q) \le r(X' \quad q)$$
$$\le r(X'X) .$$

But
$$r(X'X \quad q) \ge r(X'X) \text{ , by the definition of rank.}$$

Therefore
$$r(X'X \quad q) = r(X'X)$$

and so $X'Xu = q$ has a solution for u.

Proof (ii). If $X'Xu = q$ has a solution for u

$$q'b = u'X'Xb = u'X'(Xb) = u'X'E(y) = E[(Xu)'y] \ ;$$

i.e., there is a linear function of the q's whose expected value is $q'b$. Hence $q'b$ is estimable and the theorem is proved.

It is readily seen that if the necessary and sufficient condition of Theorem 2 is met then so is that of Theorem 1. For, if u exists such that $q' = u'X'X$, then clearly t exists ($= u'X'$) such that $q' = t'X$. For vectors q' of this nature the next theorem indicates an optimum estimator of the estimable function $q'b$.

Theorem 3. If $q'b$ is estimable, then $q'\hat{b}$ is an unbiased estimator of it that is invariant to which solution of $X'Xb = X'y$ is used for $\hat{b}$.

Proof. For $q'b$ being estimable $X'Xu = q$ has solutions for u. Let u_1 be one such solution. Then, if $\hat{b}_1$ and $\hat{b}_2$ are two different solutions of the normal equations, $X'X\hat{b}_1 = X'y = X'X\hat{b}_2 = X'y$ and

$$\begin{aligned} u_1'X'y = u_1'X'X\hat{b}_1 &= (u_1'X'X)\hat{b}_1 = q'\hat{b}_1 \\ = u_1'X'X\hat{b}_2 &= (u_1'X'X)\hat{b}_2 = q'\hat{b}_2 \end{aligned}$$

and hence $q'\hat{b}_1 = q'\hat{b}_2$. Furthermore, if u_2 is another solution to $X'Xu = q$ then $u_2'X'y = q'\hat{b}_1 = q'\hat{b}_2$. Thus $q'\hat{b}$ is invariant to which solution of $X'X\hat{b} = X'y$ is used for $\hat{b}$. Furthermore,

$$E(q'\hat{b}) = E(u'X'X\hat{b}) = E(u'X'y) = u'X'E(y) = u'X'Xb = q'b \ ,$$

showing that $q'\hat{b}$ is unbiased. Thus is the theorem proved.

The importance of this theorem is that if $q'b$ is estimable we only need find any one solution $\hat{b}$ of the normal equations $X'X\hat{b} = X'y$ and then $q'\hat{b}$ will be an unbiased estimator of $q'b$, and $q'\hat{b}$ will have the same value no matter which solution of the normal equations is used for $\hat{b}$. An additional property of this estimator, its minimum variance property, is now noted.

Theorem 4. Of all unbiased linear estimators of the estimable function $q'b$, that having the smallest variance is $q'\hat{b}$.

Proof. When $q'b$ is estimable, $q'\hat{b}$ can be written as $q'\hat{b} = u'X'y$ where $X'Xu = q$, as in the proof of Theorem 2. Therefore, because $\text{var}(a'y) = a'a\sigma^2$ for any vector a of order n,

$$\text{var}(q'\hat{b}) = \text{var}(u'X'y) = u'X'Xu\sigma^2 .$$

Suppose an unbiased estimator of $q'b$ other than $q'\hat{b}$ is $q'\hat{b} + a'y = u'X'y + a'y$; because it is unbiased

$$E(u'X'y + a'y) = q'b ,$$

$$q'b + a'Xb = q'b$$

and so

$$a'X = 0 .$$

Therefore

$$\text{var}(u'X'y + a'y) = (u'X' + a')(Xu + a)\sigma^2$$

$$= (u'X'Xu + a'a)\sigma^2 .$$

Now $a'a\sigma^2$ is always positive, and $u'X'Xu\sigma^2 = \text{var}(q'\hat{b})$. Consequently

$$\text{var}(u'X'y + a'y) > \text{var}(q'\hat{b}) ;$$

i.e. $q'\hat{b}$ has smaller variance than any other unbiased estimator of $q'b$.

From these four theorems it is evident that although the basic definition of estimability may have appeared abstruse initially, the implications derived from it are of great importance: if $q'b$ is estimable, $q'\hat{b}$ is an unbiased estimator of it that is invariant to the choice of which solution to the normal equations is used for $\hat{b}$; and of all unbiased estimators of $q'b$, $q'\hat{b}$ has the smallest variance. Thus it is that $q'\hat{b}$ is referred to as the best linear unbiased (BLU) estimator of $q'b$, "best" in this sense meaning having smallest variance (from among all linear unbiased estimators).

We can notice here that linear functions of estimable functions are themselves estimable. For, if $q_1'b$ and $q_2'b$ are both estimable, with $q_1' = u'X'X$ and $q_2' = v'X'X$, then for λ_1 and λ_2 being scalars

$$\lambda_1 q_1'b + \lambda_2 q_2'b = (\lambda_1 q_1' + \lambda_2 q_2')b$$

and

$$\lambda_1 q_1' + \lambda_2 q_2' = (\lambda_1 u' + \lambda_2 v')X'X .$$

Therefore $\lambda_1 q_1'b + \lambda_2 q_2'$ is estimable.

As shown in Theorem 4, the BLU estimator of the estimable function $q'b$ is $q'\hat{b} = u'X'y$ where $u'X'X = q'$. But this formulation is of little help in ascertaining values of q' for which $q'b$ is estimable, or in finding the estimator $q'\hat{b}$. We attend to these matters by solving the normal equations $X'X\hat{b} = X'y$ using a generalized inverse matrix as described earlier (Searle, 1965 and 1966). Brief outline is given below.

Generalized inverse matrices

Solutions for the normal equations

$$X'X\hat{b} = X'y$$

are developed from a generalized inverse of $X'X$, namely a matrix G such that

$$X'XGX'X = X'X \, . \qquad --- (3)$$

One method of deriving such a matrix G, based on the equivalent canonical form of $X'X$, is given in Searle (1966). Other methods are referred to in Rao (1962). Note that by transposing (3) we find that

$$X'XG'X'X = X'X \, , \qquad --- (4)$$

so that if G is a generalized inverse of $X'X$ then so is G'.

Other properties of G are worthy of note.

On defining $$H = GX'X \, , \qquad --- (5)$$

we find that H is idempotent,

and $XH = X$, equivalent to $XGX'X = X$. --- (6)

The idempotency of H is seen from utilizing (3) in the expansion of H^2; and the equality of XH and X is demonstrated by applying the result that $M'M = 0$ only if $M = 0$. For,

$$(XH - X)'(XH - X) = H'X'XH - X'XH - H'XX + X'X = 0$$

by (3), (4) and (5) and so $XH = X$ as in (6). It can also be shown that

$$r(G) = r(H) = r(X) = r$$

and $$r(H - I) = k - r \, . \qquad --- (7)$$

Obtaining estimable functions

As shown in Searle (1966), using arguments derived from Rao (1962), solutions to (2) are readily obtained by means of a matrix G as just defined. They are

$$\hat{b} = GX'y + (H - I)z$$

where z is any arbitrary vector of order k. The simplest form of $\hat{b}$ is with $z = 0$:

$$\hat{b}_o = GX'y \qquad --- (8)$$

enabling $\hat{b}$ to be written as

$$\hat{b} = \hat{b}_o + (H - I)z \ . \qquad --- (9)$$

Now it is a direct outcome of this method of solving equations that for any solution $\hat{b}$ derived from (8) a linear function of the elements of $\hat{b}$, $m'\hat{b}$ say, will be invariant to which form of $\hat{b}$ is used provided $m'H = m'$. But invariance with respect to $\hat{b}$ is just what is encountered in the BLU of an estimable function. This leads to a fifth theorem:

Theorem 5. With G a generalized inverse of $X'X$ and $H = GX'X$, then $q'b$ is estimable if and only if $q'H = q'$.

Proof. Given that $q'b$ is estimable, $q' = u'X'X$ for some u and

$$q'H = u'X'XH = u'X'XGX'X = u'X'X = q' \ .$$

Conversely, given $q'H = q'$,

$$q'b = q'Hb = q'GX'Xb = q'GX'E(y) = E[(q'GX')y] \ ,$$

so defining $q'b$ as being estimable.

Thus when

$$q'H = q' \qquad --- (10)$$

$q'b$ is estimable with BLU estimator $q'\hat{b}$. By taking (8) as the form of $\hat{b}$ this estimator is

$$q'\hat{b} = q'GX'y = q'\hat{b}_o \ . \qquad --- (11)$$

Ascertaining whether a function $q'b$ is estimable or not is therefore achieved by seeing if q' satisfies (10); if it does the BLU of $q'b$ is given by (11). This begs the question as to what functions are estimable; to which the answer is those functions $q'b$ for which q' satisfies (10). And from the idempotency of H we see that vectors q' of the form

$$q' = w'H \text{ for arbitrary } w \qquad --- (12)$$

do indeed satisfy (10). Consequently, for any arbitrary k-order vector w'

$$q'b = w'Hb \qquad --- (13)$$

is an estimable function, and its BLU estimator given by (11) is

$$\begin{aligned} q'\hat{b} &= w'H\hat{b}_0 \\ &= w'GX'XGX'y \\ &= w'GX'y \quad \text{from (6)} \\ &= w'\hat{b}_0 \ . \qquad --- (14) \end{aligned}$$

Also, $\quad \text{var}(q'\hat{b}) = w'Gw\sigma^2$.

Thus, for any values used as elements of w' in (13), the expression $(q'b =)w'Hb$ is estimable; and the same values used in w' in (14) give its BLU estimator $(q'\hat{b} =)w'\hat{b}_0$. And finally, because the rank of H is r there are only r linearly independent vectors $q' = w'H$, and hence from (12) there are only r linearly independent estimable functions.

One particularly useful concept can be derived from the estimable function $w'Hb$ and its BLU estimator $w'\hat{b}_0$. By letting w' be each row in turn of the $k \times k$ identity matrix we see that Hb is estimable and its BLU estimator is $\hat{b}_0$; i.e. the estimable function corresponding to the solution $\hat{b}_0$ is Hb. This, of course, can also be seen by noting that $E(\hat{b}_0) = E(GX'y) = GX'Xb = Hb$. The consequence is that, for any generalized inverse G, the solution $GX'y$ used as $\hat{b}_0$ has Hb as the corresponding estimable function. This means that although $\hat{b}_0$ is, of course, not a BLU of b, the function of b for which it is the BLU is easily found, namely Hb.

Operationally it seems very important to notice that nowhere in the estimation process here described has any extraneous "constraint" or "restriction" been placed on the parameters of the model or on the solutions of the normal equations, "in order to get a solution". Although the concept of constraints (or restrictions) - often the "usual constraint", $\Sigma t_i = 0$ for example - is so often used in statistical texts, it has not entered the discussion here. And it is quite unnecessary. Once the concept of estimability and the method of generalized inverses are assimilated the "usual" constraints become quite "unusual". The solutions of the normal equations can be established by using a generalized inverse, and from this can be derived the BLU estimators, invariant to which solution is used, of estimable functions. All other functions have estimators that do not have the optimal properties of BLU estimators.

In dismissing the use of constraints "in order to get a solution" of the normal equations so peremptorily, we are not dismissing situations in which there may truly be constraints or restrictions on the elements of the model. These are considered anon.

Estimability and Experimenters

The value of the estimability concept to experimenters is worth re-emphasizing. The normal equations have many different solutions. But certain linear functions of the parameters, known as estimable functions, have BLU estimators that are the same for all solutions of the normal equations. Only if $q'b$ is estimable will its BLU estimator $q'\hat{b}$ have the one value no matter which solution of the normal equations is used for $\hat{b}$. And if $q'b$ represents something of interest to an experimenter he will surely want to estimate it from his data; and if, from these data, a series of different estimates can be obtained by using different solutions of the normal equations, the experimenter will, to say the least, be confused and no doubt dismayed at any statistical methodology that gives him many different estimates. What he wants is one, single estimate of $q'b$ from the data - the best that can be provided. By interpreting "best" as the BLU estimator, the statistician gives to the experimenter just what he seeks, an estimator that has but one, invariant value no matter what solution of the normal equations is used, an estimator that is also unbiased and one that has smaller variance than any other unbiased estimator. But this can be done only if $q'b$, the function of interest to the experimenter, is estimable. Thus it is that the statistician demands of the experimenter that he be interested only in functions that are estimable. This is why experimenters (and all data collectors) should be made aware of this situation before collecting their data - to make sure that functions of the parameters that are of interest to them will indeed be estimable. For, as shown in Theorem 2, and again in Theorem 5, $q'b$ is estimable only if q' has a certain form - a form that depends on X, which itself is governed entirely by the form of the data that will be collected. When the data are to be obtained from a well designed experiment there is little problem - for the consequences of the resulting X are usually well-known and available in numerous texts on experimental design and analysis. But when data are to come from any kind of survey, or if there will be numerous "missing observations", then the experimenter must be certain that the things he is interested in will be estimable from the data he plans to gather. If they are not he will have to change the form of data he intends accumulating.

Example. Several numerical examples of the methods given above are to be found in Searle (1966). Another is given below. It is used subsequently in discussing the relationship of testable hypotheses to this development of estimable functions.

Suppose that the two observations on each of three treatments are as follows.

	Treatments		
	t_1	t_2	t_3
	8	5	12
	6	3	14
Totals	14	8	26

The equation of the model, $y = Xb + e$, is

$$\begin{bmatrix} 8 \\ 6 \\ 5 \\ 3 \\ 12 \\ 14 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ t_1 \\ t_2 \\ t_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \\ e_{31} \\ e_{32} \end{bmatrix} ,$$

and the normal equations $X'X\hat{b} = X'y$ are

$$\begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \hat{b} = \begin{bmatrix} 48 \\ 14 \\ 8 \\ 26 \end{bmatrix} . \qquad \text{--- (15)}$$

One value of G such that $X'XGX'X = X'X$ is

$$G = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \qquad \text{--- (15a)}$$

with

$$H = GX'X = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} . \qquad \text{--- (16)}$$

Thus by (13), with $w' = (w_0 \;\; w_1 \;\; w_2 \;\; w_3)$, an estimable function is

$$q'b = w'Hb = (w_1 + w_2 + w_3)\mu + w_1t_1 + w_2t_2 + w_3t_3 . \qquad \text{--- (17)}$$

This is estimable for any values given to the w's. And for the solution to the normal equations given by (8)

$$\hat{b}_o = GX'y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 48 \\ 14 \\ 8 \\ 26 \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} \qquad \text{--- (18)}$$

the BLU estimator of $q'b$ is, by (14),

$$q'\hat{b} = w'\hat{b}_o = 7w_1 + 4w_2 + 13w_3 . \qquad \text{--- (19)}$$

Thus for any w-values used in (17) q'b so derived is estimable and has, as its BLU estimator, the value given by (19) using the same w's. Examples are shown below.

Examples of estimable functions

Example	Values of w's			Estimable function Equation (17)	BLU estimator Equation (19)
	w_1	w_2	w_3	$(w_1 + w_2 + w_3)\mu + w_1t_1 + w_2t_2 + w_3t_3$	$7w_1 + 4w_2 + 13w_3$
1	1	-1	0	$t_1 - t_2$	3
2	0	1	-1	$t_2 - t_3$	-9
3	1	1	0	$2\mu + t_1 + t_2$	11
4	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\mu + (t_1 + t_2 + t_3)/3$	8
5	1	0	0	$\mu + t_1$	7
6	$\frac{1}{2}$	$\frac{1}{2}$	-1	$\frac{1}{2}(t_1 + t_2) - t_3$	$-7\frac{1}{2}$

Writing each of the estimable functions shown in the above table as q'b it is easily shown that q'H = q'. But for the following functions that equality is not satisfied.

Examples of non-estimable functions

$$\mu$$
$$t_1$$
$$t_1 + t_2$$
$$2t_1 + 7t_2 - 2t_3$$

For each of these, q'H $\neq$ q'; and so no numerical values can be allocated to w_1, w_2 and w_3 in such a way that (17) reduces to any of these expressions.

Testable hypotheses

We confine ourselves to hypotheses formulated by equating a linear function $q'b$ to a pre-assigned constant, m say. Thus in our example, with $b' = (\mu, t_1, t_2, t_3)$, one hypothesis might be $t_1 - t_2 = 7$ and another could be $\mu = 17$.

A hypothesis can also involve several such statements, s of them say, represented as $Q'b = m$, where Q' is $s \times k$ and m is $s \times 1$. For example, the hypothesis $t_1 = t_2 = t_3$ can be expressed as

$$\begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} \mu \\ t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} .$$

In hypotheses of this nature Q', having s rows, is taken to consist of linearly independent rows; i.e. $r(Q) = s$.

Elston and Bush (1964) define a hypothesis $q'b = m$ as being testable if and only if $q'b$ is estimable. This, of course, is correct. But it gives no answer to the question "why can we not test a hypothesis about $q'b$ if $q'b$ is not estimable"? To some this question may appear trite, but to experimenters steeped in data it is not a question asked lightly, because it is only a special case of the more general question "why can't I test any hypothesis I want to"? There is one overriding reason; only if $q'b$ is estimable will $q'\hat{b}$ be invariant to the choice of $\hat{b}$; and by the manner in which $q'\hat{b}$ is involved in the customary F-test for testing the hypothesis $q'b = m$ it is necessary to have $q'\hat{b}$ invariant for F to be invariant also. Clearly a test using F would be of no value unless it was so invariant. For this reason then, as we will show, the hypothesis $q'b = m$ is testable only when $q'b$ is estimable.

In the same way, the hypothesis $Q'b = m$ is testable only if each row of $Q'b$ is estimable; in which case we speak of $Q'b$ as being estimable. There are s rows in Q', and they are linearly independent; and with each row of $Q'b$ being estimable this means $s \leq r \leq k$ because there are only r linearly independent estimable functions.

Derivation of F-test

We will not give here the full development of the F-test as given for example, in chapter 6 of Graybill (1961) for a regression model. It suffices to say that the F-value for testing the hypothesis is the ratio of two mean squares; the numerator is the mean square attributable to the hypothesis and the denominator is the residual mean square when fitting the model without the hypothesis. We consider the latter first.

If fitting the model $y = Xb + e$, known as the full model in contrast to the reduced model soon to be defined, the analysis of variance is as follows.

Analysis of Variance for Full Model

Source of variation	d.f.	Sum of Squares
Model (including mean)	r	$SSM_0 = \hat{b}_0 X'y$
Residual	$n - r$	$SSR_0 = y'y - \hat{b}_0 X'y$
Total	n	$SST_0 = y'y$

The derivation of SSR_0 in this form comes from

$$y - X\hat{b} = y - X[\hat{b}_0 + (H - I)z], \quad \text{from (9)}$$

$$= y - X\hat{b}_0, \quad \text{from (6)}$$

and so

$$SSR_0 = (y - X\hat{b})'(y - X\hat{b})$$

$$= (y - X\hat{b}_0)'(y - X\hat{b}_0)$$

$$= y'y - 2\hat{b}_0'X'y + \hat{b}_0'X'X\hat{b}_0$$

$$= y'y - 2\hat{b}_0'X'y + \hat{b}_0(X'y) \text{ because of (2)}$$

$$= y'y - \hat{b}_0'X'y \; . \qquad --- (20)$$

The denominator of the F-test is the residual mean square

$$\hat{\sigma}^2 = SSR_0/(n - r) \; . \qquad --- (21)$$

To test the hypothesis $Q'b = m$ we fit the model

$$y = X'b + e \quad \text{subject to } Q'b = m . \qquad (22)$$

This is called the reduced model. With Q' being of rank s the analysis of variance for fitting this model is as follows.

Analysis of Variance for Reduced Model

Source of Variation	d.f.	Sum of Squares
Model	$r - s$	SSM_H
Residual	$n - r + s$	$SSR_H = y'y - SSM_H$
Total	n	

The F-test for the hypothesis is then

$$F = \frac{SSR_H - SSR_O}{s\hat{\sigma}^2} .$$

We proceed to derive an expression for SSR_H.

To fit model (22) we minimize $(y - Xb)'(y - Xb)$ subject to the condition $Q'b = m$. Introducing an s-order vector 2λ of Lagrange multipliers leads to the equations

$$X'X\tilde{b} + Q\lambda = X'y \qquad (23)$$

and

$$Q'\tilde{b} = m \qquad (24)$$

where $\tilde{b}$ distinguishes the solution from $\hat{b}$. Now, were λ known, equation (23) would be solved in the same manner as (2), namely as

$$\tilde{b} = G(X'y - Q\lambda) + (H - I)z$$

$$= \hat{b}_0 + (H - I)z - GQ\lambda \qquad --- (25)$$

$$= \hat{b} - GQ\lambda . \qquad --- (26)$$

Using this in (24) gives

$$Q'\hat{b} - Q'GQ\lambda = m$$

and so

$$Q'GQ\lambda = Q'\hat{b} - m . \qquad --- (27)$$

These results hold true for any matrix Q. But two cases are now distinguishable; when Q'b is estimable and when it is not.

We first consider the case of Q'b being estimable. Then, for some U', Q' = U'X'X, and so

$$Q'GQ = U'X'XGX'XU = U'X'XU = (XU)(XU) .$$

Therefore $\quad r(Q'GQ) = r(XU)$.

But $\quad Q' = (XU)'X$.

Therefore $\quad r(Q) = s \leq r(XU)$.

$\therefore \quad r(Q'GQ) \geq s$

But $\quad r(Q'GQ) \leq s = r(Q)$

Hence $r(Q'GQ) = s$, as does its order, and so $Q'GQ$ is non-singular. Therefore in (27)

$$\lambda = (Q'GQ)^{-1}(Q'\hat{b} - m) . \qquad --- (28)$$

Furthermore, because $Q'b$ is estimable $Q'\hat{b} = Q'\hat{b}_0$ and so

$$\lambda = (Q'GQ)^{-1}(Q'\hat{b}_0 - m) , \qquad --- (29)$$

and hence in (26)

$$\tilde{b} = \hat{b} - GQ(Q'GQ)^{-1}(Q'\hat{b}_0 - m) . \qquad --- (30)$$

In $\hat{b}$ the vector z is arbitrary; taking $z = 0$ reduces $\hat{b}$ to $\hat{b}_0$, as in (8), and denoting the resulting value of $\tilde{b}$ by $\tilde{b}_0$ we have, from (26) and (30)

$$\tilde{b}_0 = \hat{b}_0 - GQ\lambda \qquad --- (31)$$

$$= \hat{b}_0 - GQ(Q'GQ)^{-1}(Q'\hat{b}_0 - m) . \qquad --- (32)$$

and also, from (25) and (26)

$$y - X\tilde{b} = y - X[\hat{b}_0 + (H - I)z - GQ\lambda]$$

$$= y - X(\hat{b}_0 - GQ\lambda) ,$$

so that, in fitting the reduced model,

$$SSR_H = (y - X\tilde{b})'(y - X\tilde{b})$$

$$= (y - X\hat{b}_0 + XGQ\lambda)'(y - X\hat{b}_0 + XGQ\lambda)$$

$$= (y - X\hat{b}_0)'(y - X\hat{b}_0) + 2\lambda'Q'G'X'(y - X\hat{b}_0) + \lambda'Q'G'X'XGQ\lambda .$$

Now $(y - X\hat{b}_0)'(y - X\hat{b}_0) = SSR_0$;

$$X'(y - X\hat{b}_0) = X'y - X'X\hat{b}_0 = 0 ;$$

and because $Q'b$ is estimable $Q' = U'X'X$ (analogous to $q' = u'X'X$ for estimable $q'b$), so that

$$\lambda'Q'G'X'XGQ\lambda = \lambda'U'X'XG'X'XGX'XU\lambda$$

$$= \lambda'U'X'XGX'XU\lambda$$

$$= \lambda'Q'GQ\lambda .$$

Therefore

$$SSR_H = SSR_0 + \lambda' Q' G Q \lambda$$

$$= SSR_0 + (Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m) \qquad --- (33)$$

on substituting for λ from (29). Hence

$$F = \frac{SSR_H - SSR_0}{s\hat{\sigma}^2}$$

$$= \frac{(Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m)}{s\hat{\sigma}^2} \qquad --- (34)$$

It is evident from this that $Q'\hat{b}_0$ must be invariant to the choice of $\hat{b}_0$ in order for F to be similarly invariant; i.e. $Q'b$ must be estimable in order for $Q'b = m$ to be testable.

Example (continued)

As indicated earlier, $y' = (8\ 5\ 6\ 3\ 12\ 14)$ and so

$$y'y = 474 \; ; \qquad --- (35)$$

and from (20), (18) and (15)

$$SSR_0 = 474 - (0 \quad 7 \quad 4 \quad 13)\begin{bmatrix} 48 \\ 14 \\ 8 \\ 26 \end{bmatrix} = 474 - 468$$

$$= 6 \; . \qquad --- (36)$$

Hence, from (21)

$$\hat{\sigma}^2 = \frac{6}{6 - 3} = 2 \qquad --- (37)$$

Now consider the hypothesis $t_1 - t_2 = 7$. It can be written as

$$(0 \quad 1 \quad -1 \quad 0)b = 7$$

so that $$Q'\hat{b}_0 - m = (0 \quad 1 \quad -1 \quad 0)\begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} - 7 = -4$$

and Q'GQ, using G from (15a), is

$$Q'GQ = \tfrac{1}{2} + \tfrac{1}{2} = 1 .$$

The value of s is 1 and so

$$F = \frac{(-4)(1)^{-1}(-4)}{1(2)} = \frac{16}{2} = 8 .$$

Using the numerator of F from (33) and the value of SSR_0 given in (36) we get

$$SSR_H = SSR_0 + 16 = 6 + 16 = 22 . \qquad --- (37a)$$

Further, from (30), (9) and (18)

$$\tilde{b} = \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} + \begin{bmatrix} -z_1 \\ z_1 \\ z_1 \\ z_1 \end{bmatrix} - \begin{bmatrix} 0 \\ \tfrac{1}{2} \\ -\tfrac{1}{2} \\ 0 \end{bmatrix} (1)^{-1}(-4)$$

$$= \begin{bmatrix} -z_1 \\ 9 + z_1 \\ 2 + z_1 \\ 13 + z_1 \end{bmatrix} \qquad --- (38)$$

A negative "sum of squares"

The analysis of variance table given earlier for the reduced model can be misleading, for as set out there it is possible to have the residual sum of squares SSR_H greater than the total sum of squares y'y. This would make SSM_H negative. Discussion of this possibility is therefore merited, and is provided by means of an example.

Suppose we have two observations on each of two treatments:

Treatment	Observations		Totals
1	8	6	14
2	5	3	8
			22

The normal equations are

$$\begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} \mu \\ t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} 22 \\ 14 \\ 8 \end{bmatrix} ,$$

and with $G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$ a solution is $\hat{b}_0 = \begin{bmatrix} \hat{\mu} \\ \hat{t}_1 \\ \hat{t}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \\ 4 \end{bmatrix}$.

Sums of squares are

$$y'y = 64 + 36 + 25 + 9 = 134$$

$$\hat{b}_0'X'y = 7(14) + 4(8) \qquad = 130$$

and the analysis of variance is

Source	d.f.	S.S.
Model	2	$SSM_0 = 130$
Residual	2	$SSR_0 = \ \ 4$
Total	4	134

Consider testing the testable hypothesis $t_1 - t_2 = m$. Written as $Q'b = m$ we have $Q' = (0 \ \ 1 \ \ -1)$, $Q'\hat{b}_0 = 7 - 4 = 3$, and $Q'GQ = 1$. Therefore

$$SSR_H = SSR_0 + (Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m)$$

$$= 4 + (3 - m)^2$$

and the analysis of variance for the reduced model as given above becomes

Source	d.f.	S.S.
Reduced model	1	$130 - (3 - m)^2 = SSM_H$
Residual	3	$4 + (3 - m)^2 = SSR_H$
Total	4	$134 \qquad = y'y$

Clearly SSR_H can exceed the total sum of squares if m is large enough; and equivalently, SSM_H can be negative. And this is quite feasible since m is any pre-assigned number we care to use in the hypothesis.

The interpretation of this situation is that SSM_H is not a true sum of squares. Indeed, its expected value can be negative:

$$E(SSM_H) = E[\hat{b}_0'X'y - (Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m)]$$

and for our particular example this is

$$E(SSM_H) = E[(0 \quad \bar{y}_1 \quad \bar{y}_2)\begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \end{bmatrix} - (\bar{y}_1 - \bar{y}_2 - m)^2]$$

$$= E\{2(\bar{y}_1^2 + \bar{y}_2^2) - [\bar{y}_1^2 + \bar{y}_2^2 - 2\bar{y}_1\bar{y}_2 + m^2 - 2m(\bar{y}_1 - \bar{y}_2)]\}$$

$$= E\{\bar{y}_1^2 + \bar{y}_2^2 + 2\bar{y}_1\bar{y}_2 - m^2 + 2m(\bar{y}_1 - \bar{y}_2)\} .$$

Now $E(\bar{y}_1) = \mu + t_1$,

$$E(\bar{y}_1^2) = (\mu + t_1)^2 + \tfrac{1}{2}\sigma^2 ,$$

and $E(\bar{y}_1\bar{y}_2) = (\mu + t_1)(\mu + t_2)$.

Thus $E(SSM_H) = \sigma^2 + (2\mu + t_1 + t_2)^2 - m^2 + 2m(t_1 - t_2)$

and for large m this can be negative.

An intuitive interpretation of SSM_H having some satisfaction is that it is simply the deviation from SSM_0 of $(SSR_H - SSR_0)$, the sum of squares due to the hypothesis - in this case $(3 - m)^2$; and if the hypothesis is such that $(3 - m)^2$ is large then this deviation can be negative.

Before deriving a sum of squares that is truly due to fitting the reduced model, notice that the possibility of SSM_H being negative in no way affects the test of the hypothesis: for the F-value for this test is

$$F = \frac{SSR_H - SSR_0}{(1)SSR_0/2} = \tfrac{1}{2}(3 - m)^2$$

no matter what value m has.

The analysis of variance for the reduced model must now be reconsidered. In this model we assume the null hypothesis $t_1 - t_2 = m$ is true. The equations of the model are then

$$\begin{aligned} 8 &= \mu + m + t_2 + e_1 \\ 6 &= \mu + m + t_2 + e_2 \\ 5 &= \mu + t_2 + e_3 \\ 3 &= \mu + t_2 + e_4 \end{aligned} \quad \text{equivalent to} \quad \begin{aligned} 8 - m &= \mu + t_2 + e_1 \\ 6 - m &= \mu + t_2 + e_2 \\ 5 &= \mu + t_2 + e_3 \\ 3 &= \mu + t_2 + e_4 . \end{aligned}$$

It is to the second form of these equations, not the first, that the least squares procedure applies, equations in which y' is (8-m, 6-m, 5, 3) not (8, 6, 5, 3). The resulting normal equations are

$$\begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} \begin{bmatrix} \mu \\ t_2 \end{bmatrix} = \begin{bmatrix} 22 - 2m \\ 22 - 2m \end{bmatrix}$$

for which $\mu = 0$ and $t_2 = 5\frac{1}{2} - \frac{1}{2}m$ is a solution. Hence

$$y'y = (8 - m)^2 + (6 - m)^2 + 25 + 9 = 134 - 28m + 2m^2$$

$$\hat{b}_0'X'y = (5\tfrac{1}{2} - \tfrac{1}{2}m)(22 - 2m) \qquad = 121 - 22m + m^2$$

and so the analysis of variance is

<u>Analysis of Variance for Reduced Model</u>

Source	d.f.	Sums of Squares
Model	1	$121 - 22m + m^2 = SSM_H$
Residual	3	$13 - 6m + m^2 = SSR_H$
Total	4	$134 - 28m + 2m^2$

We see that

$$SSR_H = 13 - 6m + m^2 = 4 + (3 - m)^2 \quad \text{as before,}$$

but now

$$SSM_H = 121 - 22m + m^2 = (11 - m)^2 ,$$

different from the earlier value because $y'y$ is now based on the correct value of y in the reduced model, namely $y' = (8-m, 6-m, 5, 3)$. The sameness of SSR_H with this y is to be emphasized, indicating the appropriatness of the F-test regardless of the value of m, as has already been mentioned.

Hypotheses based on non-estimable functions

Having shown that the hypothesis $Q'b = m$ can be tested when $Q'b$ is estimable we now consider what happens when $Q'b$ is not estimable. We start with the case of just one row in Q', namely the hypothesis $q'b = m$. Just as in (23) and (24), we arrive at normal equations for the reduced model:

$$X'X\tilde{b} + q\lambda = X'y$$
$$\text{and} \qquad q'\tilde{b} = m \qquad --- (39)$$

where λ is here a scalar. These are k+1 equations in (39) in k+1 unknowns; $X'X$ has rank r, and because $q'b$ is non-estimable $q' \neq t'X'X$ and so the rank of the equations is r+1. Hence for $r < k$, equations (39) have an infinite number of solutions. Let us confine ourselves to those for which $\lambda = 0$. Then (39) reduce to

$$X'X\tilde{b} = X'y$$
$$q'\tilde{b} = m \qquad --- (40)$$

and because $q' \neq t'X'X$ these equations are consistent and accordingly have a solution. Thus any solution to (40) together with $\lambda = 0$ is a solution of (39). But from the point of view of testing the hypothesis λ is of no interest: therefore we need only solve (40). The first of these is the same as the normal equations (2) so that, as in (25),

$$\tilde{b} = GX'y + (H - I)z$$

where z is arbitrary. The second equation of (40) can be satisfied by appropriate choice of z: we want

$$q'\tilde{b} = q'GX'y + q'(H - I)z = m .$$

Therefore z is chosen as z^* say, where

$$q'(H - I)z^* = m - q'GX'y = m - q'\hat{b}_0 \qquad --- (41)$$

and then

$$\tilde{b} = \hat{b}_0 + (H - I)z^* . \qquad --- (42)$$

This procedure demands that (42) have a solution for z^*. Since $(H - I)$ has order k and rank k-r (equation (7)), $(H - I)z$ has k-r elements that are arbitrary. But (42) imposes only a single linear constraint on the elements of $(H - I)z^*$ and therefore (42) can be solved.

In (41) and (42) we have established the estimator $\tilde{b}$ for the reduced model involving the hypothesis $q'b = m$ when $q'b$ is non-estimable. But by the nature of (42) the $\tilde{b}$ so obtained is simply one of the solutions of $X'X\hat{b} = X'y$, the normal equations of the full model. This means that SSR_H for this reduced model will be the same as SSR_0, and consequently the F-value will be identically zero for all data, no matter what they are. We therefore conclude that the hypothesis cannot be tested. Thus when $q'b$ is not estimable the hypothesis $q'b = m$ is not testable; and conversely, only when $q'b$ is estimable can the hypothesis $q'b = m$ be tested.

This result is now extended to the hypothesis $Q'b = m$ when $Q'b$ is not estimable. Clearly, if every row of $Q'b = m$ is simply a series of non-testable hypotheses such as just been discussed, and the whole hypothesis cannot be tested. Likewise if some rows of $Q'b$ are estimable and some are not that part of the hypothesis $Q'b = m$ involving the non-estimable rows cannot be tested and so the whole hypothesis cannot. <u>In toto</u> then, the only hypotheses that can be tested are those involving estimable functions.

It is enlightening to consider further the case of a hypothesis composed partly of estimable functions and partly of non-estimable functions. Suppose the hypothesis $Q'b = m$ can be written as

$$\begin{bmatrix} S'b \\ q'b \end{bmatrix} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$$

where S' and m_1 have t rows, m_2 is scalar and where $S'b$ is estimable but $q'b$ is not. Then for the vector $(\lambda_1' \quad \lambda_2)$ of Lagrange multipliers, λ_1' being of order t and λ_2 a scalar, the normal equations of the reduced model, similar to (23) and (24) are

$$\begin{aligned} X'X\tilde{b} + S\lambda_1 + q\lambda_2 &= X'y \\ S'\tilde{b} \qquad\qquad &= m_1 \\ q'\tilde{b} \qquad\qquad &= m_2 \end{aligned}$$

These are k+t+1 equations in as many variables, having rank r+1. ($S' = U'X'X$ for some U, because $S'b$ is estimable and $q' \neq u'X'X'$ because $q'b$ is not estimable.) Therefore they have many solutions, and by confining ourselves to those for which $\lambda_2 = 0$ the equations become

$$\begin{aligned} X'X\tilde{b} + S\lambda_1 &= X'y \\ S'\tilde{b} &= m_1 \\ q'\tilde{b} &= m_2 \end{aligned} \qquad --- (43)$$

The first two of these equations are similar to (23) and (24) and so can be solved in the manner of (30):

$$\tilde{b} = \hat{b} - GS(S'GS)^{-1}(S'\hat{b}_0 - m_1) .$$

The third equation of (43) is satisfied by choosing the arbitrary z implicit in $\hat{b}$ in such a way that $q'\tilde{b} = m_2$:

$$q'[\hat{b}_0 + (H - I)z^*] - q'GS(S'GS)^{-1}(S'\hat{b}_0 - m_1) = m_2 .$$

Henze z^* is determined from the equation

$$q'(H - I)z^* = m_2 - q'[\hat{b}_0 - GS(S'GS)^{-1}(S'\hat{b}_0 - m_1)] \qquad --- (44)$$

where the vector inside the square bracket is analogous to the $\tilde{b}_0$ of (32). With the value of z^* obtained from (44) the solution of (43) is then

$$\tilde{b} = \hat{b}_0 + (H - I)z^* - GS(S'GS)^{-1}(S'\hat{b}_0 - m_1) . \qquad --- (45)$$

But this is simply one of the solutions of the first two equations of (43). Consequently SSR_H for the hypothesis $\begin{bmatrix} S' \\ q' \end{bmatrix} b = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$ where q'b is estimable is identically the same as SSR_H for the hypothesis S'b = m. Hence the complete hypothesis $\begin{bmatrix} S' \\ q' \end{bmatrix} b = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$ cannot be tested. Thus if one row of a hypothesis relates to a non-estimable function then the hypothesis is not testable. And clearly the same is true for hypotheses having more than one row relating to non-estimable functions.

The condition for the hypothesis Q'b = m to be testable is that every row of Q'b be estimable, i.e. that Q'H equal Q'. Before setting out to test Q'b = m we must therefore be assured that Q' satisfies Q'H = Q'. Suppose however, that this is not done, and that we proceed at once to calculate

$$SSR_H = SSR_0 + (Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m) .$$

This exists, in particular $(Q'GQ)^{-1}$ exists, so long as $Q'b$ is estimable. But estimability of $Q'b$ is not a necessary condition for the existence of $(Q'GQ)^{-1}$; for in some cases $(Q'GQ)^{-1}$ can also exist when $Q'b$ is not estimable. In such cases we must establish the meaning of the computed value of SSR_H. If $Q'b = m$ is not testable what hypothesis, we ask ourselves, is being tested when using SSR_H?

The answer is simple. If $Q'b = m$ is not testable and we carry out the calculations for testing it then the hypothesis that is being tested is $Q'Hb = m$; or at least the calculations are indistinguishable from those for testing $Q'Hb = m$. The reason for this is readily elucidated. The residual sum of squares under the null hypothesis $Q'b = m$ is

$$SSR_H = SSR_0 + (Q'\hat{b}_0 - n)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m)$$

and that under the hypothesis $Q'Hb = m$ is

$$SSR_H^* = SSR_0 + (Q'H\hat{b}_0 - m)'(Q'HGH'Q)^{-1}(Q'H\hat{b}_0 - m) .$$

Now $$Q'H\hat{b}_0 = Q'HGX'y = Q'GX'XGX'y ,$$

and $$Q'HGH'Q = Q'GX'XGX'XG'Q = Q'GX'XG'Q .$$

<u>Lemma</u>. XGX' is invariant to which generalized inverse of $X'X$ is used for G.

<u>Proof</u>. Let F be a generalized inverse of $X'X$ different from G. Then

$$\begin{aligned}(XGX' - XFX')(XGX' - XFX')' &= XGX'XG'X - XGX'XF'X - XFX'XG'X + XFX'XF'X \\ &= XGX' - XGX' - XFX' + XFX' \\ &= 0\end{aligned}$$

Therefore $$XGX' = XFX'.$$

Now because G is a generalized inverse of $X'X$ so is G'; therefore, by the above lemma, $XGX' = XG'X'$, and hence

$$Q'H\hat{b}_0 = Q'GX'XG'X'\mathbf{y} .$$

Furthermore, because both G and G' are generalized inverses of $X'X$ it is clear that $GX'XG'$ is also. Denote $GX'XG'$ by G^*. Then

$$Q'H\hat{b}_0 = Q'G^*X'y = Q'\hat{b}^* \text{ and } Q'HGH'GQ = Q'G^*Q$$

where $\hat{b}^*$ is the solution of $X'X\hat{b} = X'y$ corresponding to G^*. Thus

$$SSR_H^* = SSR_0 + (Q'\hat{b}^* - m)'(Q'G^*Q)^{-1}(Q'\hat{b}^* - m) \ .$$

In this form SSR_H^*, using G^* instead of G, is immediately recognizable as the residual sum of squares for the reduced model under the hypothesis $Q'b = m$. But this sum of squares is invariant to the choice of G. Therefore the procedure for testing $Q'b = m$ is identical to that for testing $Q'Hb = m$. Hence carrying out the calculations for testing $Q'b = m$ when $Q'b$ is non-estimable leads to testing $Q'Hb = m$ — provided, of course, that $(Q'GQ)^{-1}$ does exist. This result is not unexpected, for if $Q'b$ is estimable $Q'b = Q'Hb$; and as in equation (13) $Q'Hb$ is estimable for any Q.

<u>Examples</u> (continued)

(a) A non-testable hypothesis.

Consider testing the hypothesis $t_1 + t_2 = 0$. With H as in (16) and $\hat{b}_0$ as in (18), $\hat{b}$ of (9) is

$$\hat{b} = \hat{b}_0 + (H - I)z = \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} + \begin{bmatrix} -z_1 \\ z_1 \\ z_1 \\ z_1 \end{bmatrix} \qquad --- (46)$$

using z_1 as the first element of z. Now

$$t_1 + t_2 = 0 \text{ is } (0 \quad 1 \quad 1 \quad 0)b = 0$$

and so (41) for determining z^*,

$$q'(H - I)z^* = m - q'\hat{b}_0 \ ,$$

is
$$(0 \quad 1 \quad 1 \quad 0)\begin{bmatrix} -z_1^* \\ z_1^* \\ z_1^* \\ z_1^* \end{bmatrix} = 0 - (0 \quad 1 \quad 1 \quad 0)\begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} \ .$$

giving $z_1^* = -5\frac{1}{2}$.

On substituting this for z_1 in (46) we get (42):

$$\tilde{b} = \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} + \begin{bmatrix} +5\frac{1}{2} \\ -5\frac{1}{2} \\ -5\frac{1}{2} \\ -5\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 5\frac{1}{2} \\ 1\frac{1}{2} \\ -1\frac{1}{2} \\ 7\frac{1}{2} \end{bmatrix} \qquad \text{--- (46a)}$$

Since this is exactly the same as (46) with z_1 replaced by $-5\frac{1}{2}$, SSR_H is the same as SSR_0 given in (36), namely 6, and the hypothesis cannot be tested. This is as we would expect, because $t_1 + t_2$ is not estimable.

(b) A non (partially) testable hypothesis.

Take the hypothesis $\begin{cases} t_1 - t_2 = 7 \\ t_2 \quad\;\; = 0 \end{cases}$.

It can be written as

$$\begin{bmatrix} S' \\ q' \end{bmatrix} b = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} b = \begin{bmatrix} 7 \\ 0 \end{bmatrix} .$$

Knowing that $S'b$ so defined is estimable and that $q'b$ is not, we use (44) and (45) to estimate b; (44) requires

$$S'\hat{b}_0 - m_1 = (0 \;\; 1 \;\; -1 \;\; 0) \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} - 7 = -4 .$$

$$GS = \begin{bmatrix} 0 \\ \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix} , \text{ and } S'GS = 1 .$$

With these values (44),

$$q'(H - I)z^* = m_2 - q'\hat{b}_0 + q'GS(S'GS)^{-1}(S'\hat{b}_0 - m_1) ,$$

is

$$(0\ 0\ 1\ 0)\begin{bmatrix}-z_1^*\\ z_1^*\\ z_1^*\\ z_1^*\end{bmatrix} = 0 - (0\ 0\ 1\ 0)\begin{bmatrix}0\\7\\4\\13\end{bmatrix} + (0\ 0\ 1\ 0)\begin{bmatrix}0\\ \frac{1}{2}\\ -\frac{1}{2}\\ 0\end{bmatrix}(1)^{-1}(-4)$$

giving $z_1^* = -4 + 2 = -2$. Substitution in (45) yields

$$\tilde{b} = \begin{bmatrix}0\\7\\4\\13\end{bmatrix} + \begin{bmatrix}2\\-2\\-2\\-2\end{bmatrix} - \begin{bmatrix}0\\ \frac{1}{2}\\ -\frac{1}{2}\\ 0\end{bmatrix}(1)^{-1}(-4)$$

$$= \begin{bmatrix}0\\9\\2\\13\end{bmatrix} + \begin{bmatrix}2\\-2\\-2\\-2\end{bmatrix}$$

$$= \begin{bmatrix}2\\7\\0\\11\end{bmatrix}.$$

This is exactly the same as $\tilde{b}$ given in (38) only with z_1 replaced by -2, and there we were testing the hypothesis $t_1 - t_2 = 7$. Thus the value of SSR_0 for the hypothesis $\begin{cases} t_1 - t_2 = 7 \\ t_2 = 0 \end{cases}$ is identical to that for the hypothesis $t_1 - t_2 = 7$, so illustrating the principle that only the estimable portion of a hypothesis involving both estimable and non-estimable functions can be tested.

Suppose, however, that we were unaware that the hypothesis $\begin{cases} t_1 - t_2 = 7 \\ t_2 = 0 \end{cases}$ was not testable, and that we proceeded to carry out the calculations for testing it. We have

$$Q' = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ and } m = \begin{bmatrix} 7 \\ 0 \end{bmatrix},$$

$$Q'\hat{b}_0 - m = \begin{bmatrix} 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 7 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \end{bmatrix},$$

$$GQ = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix}, \quad Q'GQ = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \text{ and } (Q'GQ)^{-1} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}.$$

Hence for this hypothesis

$$SSR_H = 6 + (Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m) = 6 + (-4 \ \ 4)\begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} -4 \\ 4 \end{bmatrix}$$

$$= 6 + 32 = 38 .$$

This, it will be found, is identical to SSR_H for the hypothesis $Q'Hb = m$. For

$$Q'H = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix},$$

and for this it will be found that

$$(Q'H)\hat{b}_0 - m = \begin{bmatrix} -4 \\ 4 \end{bmatrix} = Q'\hat{b}_0 - m;$$

$$\text{and} \quad [(Q'H)G(Q'H)']^{-1} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix} = (Q'GQ)^{-1} .$$

Thus SSR_H for the hypothesis $Q'Hb = m$ is the same as that for the hypothesis $Q'b = 0$.

Restricted models

Reference was made earlier to the fact that sometimes a linear model as defined in (1) may include restrictions on the elements of the parameter vector. Such restrictions are quite different from the "usual constraints" introduced in many texts for the sole purpose of getting a solution to the normal equations. The need for such constraints has already been dismissed. The restrictions envisaged here are considered to be an integral part of the model and as such must be taken into account in the estimation and testing processes.

The discussion so far has been in terms of models whose parameters have been very loosely defined. Indeed no formal definition has been made. In writing the equation of the model as $y = Xb + e$ we simply described b as being the vector of the "parameters of the model" and left it at that. Implicit in this, in terms of our simple example, μ is a general mean and t_1, t_2 and t_3 are the effects on yield arising from three different treatments. No further definition is implied. Sometimes, however, more explicit definitions inherent in the model result in there being relationships (or restrictions) among the parameters of the model. These are considered part and parcel of the model. For example, the situation may be such that the parameters of the model satisfy the relation $t_1 + t_2 + t_3 = 0$.

The models already discussed, those that contain no restrictions of the kind just referred to, shall be referred to as unrestricted models. And models that do include restrictions of this nature shall be called restricted models. The question then arises as to how the estimation and hypothesis testing processes developed for unrestricted models apply to restricted models.

In general we consider the set of restrictions $P'b = \alpha$ as part of the model. The restricted full model is then $y = Xb + e$, restricted by the condition $P'b = \alpha$. Fitting this model is operationally identical to fitting the unrestricted model under the null hypothesis $P'b = \alpha$; we call this model the unrestricted reduced model. The sum of squares $(y - Xb)'(y - Xb)$ has to be minimized subject to $P'b = \alpha$. Equations

$$X'X\hat{b}^* + P\lambda = X'y$$

and $$P'\hat{b}^* = \alpha ,$$

of the same form as (23) and (24) have to be solved, $\hat{b}^*$ distinguishing their solution from $\hat{b}$ and $\tilde{b}$. Thus

$$\hat{b}^* = \hat{b} - GP\lambda$$

and

$$P'GP\lambda = P'\hat{b} - \alpha ,$$

similar to (26) and (27). As there, two cases must be distinguished: when $P'b$ is estimable and when it is not. First, when $P'b$ is estimable. As in (30), the solution for $\hat{b}^*$ is

$$\hat{b}^* = \hat{b} - GP(P'GP)^{-1}(P'\hat{b}_0 - \alpha)$$

where of course, $\hat{b} = \hat{b}_0 + (H - I)z$, so that

$$\hat{b}^* = \hat{b}_0 - GP(P'GP)^{-1}(P'\hat{b}_0 - \alpha) + (H - I)z . \qquad --- (47)$$

To see what functions are estimable in this restricted model we observe from (47) that

$$\hat{b}^* = G[X'y - P(P'GP)^{-1}(P'\hat{b}_0 - \alpha)] + (H - I)z ,$$

and from this we conclude that $Q'b$ is estimable if $Q'(H - I) = 0$, just as in the unrestricted full model. Thus all functions that are estimable in the unrestricted full model are also estimable in the restricted full model, when the restriction involves estimable functions. The functions will, of course be amended by the restrictions. Thus if $Q'b$ was estimable earlier it is still estimable, but its interpretation is subject to the restriction $P'b = \alpha$, i.e. $Q'b$ becomes $Q'b + \lambda P'b$ for any constant λ. For example, if $Q'b$ is $2\mu + t_1 + t_2$ and if $P'b = \alpha$ is $t_1 - t_2 = 0$ then $Q'b$ becomes $2(\mu + t_1)$. It is clear in this situation that because both $Q'b$ and $P'b$ are estimable then so is $Q'b + \lambda P'b$.

In fitting the restricted full model the residual sum of squares, to be denoted by $SSR_{0'}$, will be $(y - X\hat{b}^*)'(y - X\hat{b}^*)$ where $\hat{b}^*$ is given by (47). Reduction of this expression can be made in exactly the same way as SSR_H was obtained in the unrestricted reduced model. Hence, from (33),

$$SSR_{0'} = SSR_0 + (P'\hat{b}_0 - \alpha)'(P'GP)^{-1}(P'\hat{b}_0 - \alpha) , \qquad --- (48)$$

where, as always,

$$SSR_0 = y'y - \hat{b}_0X'y . \qquad --- (49)$$

The degrees of freedom for $SSR_{0'}$ are $n - (r - p)$ and those for SSR_0 are $n - r$.

Now consider hypothesis testing in the restricted model. The hypothesis $Q'b = m$ can be considered only if it is consistent with $P'b = \alpha$; for example, if $P'b = \alpha$ were $t_1 - t_2 = 0$ one obviously could not consider the hypothesis $t_1 - t_2 = 4$. With this proviso then, the hypothesis $Q'b = m$ is tested in the restricted model having restrictions $P'b = \alpha$, by considering the unrestricted full model under the hypothesis

$$\begin{bmatrix} P' \\ Q' \end{bmatrix} b = \begin{bmatrix} \alpha \\ m \end{bmatrix} .$$

We assume throughout that the p rows of P' are linearly independent and that the s rows of Q' are likewise; and that the rows of P' and Q' are linearly independent of each other. Thus P' and Q' have rank p and s respectively and on defining

$$T' = \begin{bmatrix} P' \\ Q' \end{bmatrix} \text{ and } \tau = \begin{bmatrix} \alpha \\ m \end{bmatrix}$$

T is of rank $p + s$. We now consider what can be called the restricted reduced model, $y = Xb + e$ with $T'b = \tau$. In this case the estimator of b is, analogous to (30),

$$\tilde{b}^* = \hat{b}_0 - GT(T'GT)^{-1}(T'\hat{b}_0 - \tau) + (H - I)z \qquad --- (51)$$

and the residual sum of squares, similar to (32), is

$$SSR_{H'} = SSR_0 + (T'\hat{b}_0 - \tau)'(T'GT)^{-1}(T'\hat{b}_0 - \tau) . \qquad --- (52)$$

The degrees of freedom for $SSR_{H'}$ are clearly $n - (r - s - p)$, and the F-value for testing the hypothesis $Q'b = m$ in the restricted model is

$$F = \frac{SSR_{H'} - SSR_{0'}}{s} \Bigg/ \frac{SSR_{0'}}{n - (r - n)} \qquad --- (53)$$

where $SSR_{H'}$ and $SSR_{0'}$ are calculated as in (52) and (48) respectively.

<u>Example</u>. Consider the restricted model $y = Xb + e$ having

$$P'b = \alpha \text{ as } t_1 - t_2 = 7 .$$

Before considering it recall that in fitting the unrestricted full model $SSR_0 = 6$, and in fitting the unrestricted reduced model under the hypothesis $t_1 - t_2 = 7$ the value of SSR_H was $6 + 16 = 22$, as in equation (37a). In the restricted full model now being discussed $SSR_{0'}$ is therefore, by (48),

$$SSR_{0'} = 22 ;$$

and the solution for $\hat{b}^*$ given by (47) will be that shown in (38)

$$\tilde{b}^* = \begin{bmatrix} -z_1 \\ 9 + z_1 \\ 2 + z_1 \\ 13 + z_1 \end{bmatrix}$$

Now, in this restricted model let us test the hypothesis $\mu + t_1 = 2$; i.e., having

$$Q'b = m \text{ as } (1 \quad 1 \quad 0 \quad 0)b = 2 .$$

For obtaining $SSR_{H'}$ from (52)

$$T' = \begin{bmatrix} P' \\ Q' \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \text{ and } \tau = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$$

and so

$$T'\hat{b}_0 - \tau = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} - \begin{bmatrix} 7 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 5 \end{bmatrix}$$

$$GT = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix}, \quad T'GT = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \text{ and } (T'CT)^{-1} = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}.$$

Hence (52) is

$$SSR_{H'} = 6 + (-4 \quad 5)\begin{bmatrix} 2 & -2 \\ 12 & 4 \end{bmatrix}\begin{bmatrix} -4 \\ 5 \end{bmatrix}$$

$$= 6 + [16(2) + 25(4) - 20(-4)]$$

$$= 218 \ .$$

Hence in (53), with $n = 6$, $r = 3$, $p = 1$ and $s = 1$

$$F = \frac{218 - 22}{1} \Big/ \frac{22}{4} = \frac{392}{11} \ . \qquad \text{--- (54)}$$

And in (49) the estimator of b in the restricted reduced model is

$$\tilde{b}^* = \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} -4 \\ 5 \end{bmatrix} + \begin{bmatrix} -z_1 \\ z_1 \\ z_1 \\ z_1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} -18 \\ 28 \end{bmatrix} + \begin{bmatrix} -z_1 \\ z_1 \\ z_1 \\ z_1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} - \begin{bmatrix} 0 \\ 5 \\ 9 \\ 0 \end{bmatrix} + \begin{bmatrix} -z_1 \\ z_1 \\ z_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} -z_1 \\ 2 + z_1 \\ -5 + z_1 \\ 13 + z_1 \end{bmatrix}$$

We see in this solution that both $\tilde{t}_1^* - \tilde{t}_2^* = 7$ and $\tilde{\mu}^* + \tilde{t}_1^* = 2$.

Note that in the presence of the restriction $t_1 - t_2 = 7$ the hypothesis $\mu + t_1 = 2$ is equivalent to $\mu + t_2 = -5$. For <u>this</u> hypothesis

$$T = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \text{ and } \tau = \begin{bmatrix} 7 \\ -5 \end{bmatrix}$$

so that $$T'\hat{b}_0 - \tau = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} - \begin{bmatrix} 7 \\ -5 \end{bmatrix} = \begin{bmatrix} -4 \\ 9 \end{bmatrix}$$

and $$T'GT = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}, \text{ with } (T'GT)^{-1} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix} .$$

With these values

$$\begin{aligned} SSR_{H'} &= 6 + (-4 \quad 9) \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} -4 \\ 9 \end{bmatrix} \\ &= 6 + [16(2) + 81(4) - 36(4)] \\ &= 218 \quad \text{as before.} \end{aligned}$$

Finally we consider restricted models having restrictions that involve non-estimable functions. Fitting a restricted full model of this nature is equivalent to fitting the unrestricted reduced model under the null hypothesis $P'b = \alpha$ for $P'b$ non-estimable. Hence the solution for $\hat{b}^*$ comes from (42),

$$\hat{b}^* = \hat{b}_0 + (H - I)z^*$$

with (41) giving z^*:

$$P'(H - I)z^* = \alpha - P'\hat{b}_0 . \qquad \text{--- (55)}$$

From these equations it is clear that $\hat{b}^*$ is simply a subset of the solutions $\hat{b}$. Hence functions that are estimable in the unrestricted model are also estimable in the restricted model; the residual sums of squares in the two models are the same; and so are their degrees of freedom (because expectations in the full model are unaltered). Thus in fitting the restricted model, $SSR_{0''} = SSR_0$ with $n - r$ degrees of freedom, and $SSR_{H''} = SSR_H$ with $n - (r - s)$ degrees of freedom, and the F-test for the testable hypothesis $Q'b = m$ is the same as that in the unrestricted model. So it is that restrictions $P'b = \alpha$ for $P'b$ non-estimable restrict the solutions of the normal equations to a subset of $\tilde{b}$ but for estimable $Q'b$ they do not alter the estimate of $Q'b$ nor the F-value for testing the hypothesis $Q'b = m$.

Should the solutions to the normal equations and the residual sums of squares be explicitly required in the restricted model having $P'b$ non-estimable, the procedure is to solve the equations of the unrestricted model and then apply the restrictions. Thus for the full model first obtain $\hat{b} = \hat{b}_0 + (H - I)z$; and then the solution for the restricted full model is $\hat{b}^*$, which is simply $\hat{b}$ using z^* for z where z^* comes from the restrictions $P'\hat{b} = \alpha$, i.e. $P'(H - I)z^* = \alpha - P'\hat{b}_0$ as in (55). And for the reduced model the solution $\tilde{b} = \hat{b}_0 + (H - I)z - GQ(Q'GQ)^{-1}(Q'\hat{b}_0 - m)$ is first obtained; and then $\tilde{b}^*$ for the restricted reduced model is just $\tilde{b}$ using z^* for z where $P'\tilde{b} = \alpha$ determines z^*. In both cases the residual sums of squares are unaffected by the restrictions; i.e. $SSR_{0''} = SSR_0$ and $SSR_{H''} = SSR_H$.

Restrictions do, of course, alter hypotheses, and in this way hypotheses that are not testable in an unrestricted model can sometimes be tested in restricted models. For example, the hypothesis $\mu = 0$ cannot be tested in the unrestricted model; but if the restriction $t_1 + t_2 + t_3 = 0$ is part of the model it can, because $3\mu + t_1 + t_2 + t_3$ is estimable. Indeed, in the presence of the restriction $P'b = \alpha$ the hypothesis $Q'b = m$ can be interpreted as $(Q' + LP')b = m + L\alpha$ for any matrix L of order $s \times p$. Equivalently, for any function $Q'b$ that is estimable in the unrestricted model the function $(Q' + LP')b$ is estimable in the restricted model. Or, from another viewpoint, the function $R'b$ may not be estimable in the unrestricted model, and yet if R' has the form $R' = Q' + LP'$, $R'b$ will be estimable in the restricted model.

Example (continued)

The solution $\hat{b}$ in the unrestricted full model is

$$\hat{b} = \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} + \begin{bmatrix} -z_1 \\ z_1 \\ z_1 \\ z_1 \end{bmatrix} .$$

In the model restricted so that $t_1 + t_2 + t_3 = 0$ the solution is $\tilde{b}$, being $\hat{b}$ with z^* used for z where

$$(0 \quad 1 \quad 1 \quad 1)\hat{b} = 0 ;$$

i.e., $24 + 3z_1^* = 0$, so giving $z_1^* = -8$

and hence

$$\hat{b}^* = \begin{bmatrix} 8 \\ -1 \\ -4 \\ 5 \end{bmatrix} .$$

Clearly, $\hat{b}^*$ is simply a specific value of $\hat{b}$ and hence $SSR_{0''}$ will be the same as SSR_0:

$$\begin{aligned} SSR_{0''} &= y'y - \hat{b}^{*\prime}X'y \\ &= 474 - [8(48) - 1(14) - 4(8) + 5(26)] \\ &= 474 - 468 \\ &= 6 \\ &= SSR_0 . \end{aligned}$$

In testing the hypothesis $3\mu + t_1 + t_2 + t_3 = 18$ in the unrestricted model the solution for $\tilde{b}$ is

$$\tilde{b} = \begin{bmatrix} 0 \\ 7 \\ 4 \\ 13 \end{bmatrix} + \begin{bmatrix} -z_1 \\ z_1 \\ z_1 \\ z_1 \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} (1\tfrac{1}{2})^{-1}(24 - 18) = \begin{bmatrix} 0 \\ 5 \\ 2 \\ 11 \end{bmatrix} + \begin{bmatrix} -z_1 \\ z_1 \\ z_1 \\ z_1 \end{bmatrix} .$$

Hence in the restricted model having $t_1 + t_2 + t_3 = 0$, $\tilde{b}^*$ is $\tilde{b}$ with z^* for z where

$$(0 \quad 1 \quad 1 \quad 1)\tilde{b} = 0 ;$$

i.e. $18 + 3z_1^* = 0$, giving $z_1^* = -6$

and so

$$\tilde{b}^* = \begin{bmatrix} 6 \\ -1 \\ -4 \\ 5 \end{bmatrix} .$$

For the residual sum of squares SSR_H we have

$$\begin{aligned} SSR_H &= SSR_0 + (Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m) \\ &= 6 + (24 - 18)(1\tfrac{1}{2})^{-1}(24 - 18) = 6 + 24 \\ &= 30 . \end{aligned}$$

That this is the same as $SSR_{H'}$ can be verified from writing

$$SSR_{H'} = (y - X\tilde{b}^*)'(y - X\tilde{b}^*)$$

with
$$(y - X\tilde{b}^*) = \begin{bmatrix} 8 \\ 6 \\ 5 \\ 4 \\ 12 \\ 14 \end{bmatrix} - \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ -1 \\ -4 \\ 5 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \\ 5 \\ 4 \\ 12 \\ 14 \end{bmatrix} - \begin{bmatrix} 5 \\ 5 \\ 2 \\ 2 \\ 11 \\ 11 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 3 \\ 1 \\ 1 \\ 3 \end{bmatrix}$$

so that

$$SSR_{H'} = 9 + 1 + 9 + 1 + 1 + 9 = 30 = SSR_H .$$

The hypothesis $3\mu + t_1 + t_2 + t_3 = 18$ in the presence of the restriction $t_1 + t_2 + t_3 = 0$ then reduces to the hypothesis $\mu = 6$. That this is testable in the restricted model is evident from considering just the first of the normal equations:

$$6\hat{\mu} + 2\hat{t}_1 + 2\hat{t}_2 + 2\hat{t}_3 = 48$$

In the presence of the restriction this reduces to $\hat{\mu}^* = 3$, so indicating the estimability of μ and the testability of the hypothesis $\mu = 6$ in this restricted model.

Summary

Without redefining notation we here summarize most of the salient results.

1. $q'b$ is estimable if and only if $q' = u'X'X$ for some u'. (Theorem 2)

2. If $q'b$ is estimable its BLU estimator is $q'\hat{b}$ where $\hat{b}$ is any solution to $X'X\hat{b} = X'y$. (Theorems 3 and 4)

3. Solutions to $X'X\hat{b} = X'y$ are $\hat{b} = \hat{b}_0 + (H - I)z$ where G is a generalized inverse of $X'X$, $H = GX'X$, $\hat{b}_0 = GX'y$ and z is arbitrary. (Equation 9)

4. $q'b$ is estimable if and only if $q'H = q'$. (Equation 12)

5. For any arbitrary vector w, $q'b = w'Hb$ is estimable with BLU estimator $q'\hat{b} = w'\hat{b}_0$. (Equations 13 and 14)

6. The number of linearly independent estimable functions is the rank of X. (Page 7)

7. Hb is the BLU estimator of $\hat{b}_0$. (Page 7)

8. Testable hypotheses are those involving estimable functions. (Page 22)

9. The residual sum of squares under the full model is $SSR_0 = y'y - \hat{b}_0X'y$; $\hat{\sigma}^2 = SSR_0/(N - r)$. (Equations 20 and 21)

10. The residual sum of squares under the null hypothesis $Q'b = m$ is $SSR_H = SSR_0 + (Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m)$. (Equation 33)

11. The F-value for testing the testable hypothesis $Q'b = m$ is

$$F = \frac{(Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m)}{s\hat{\sigma}^2} .$$ (Equation 34)

12. The F-value given in item 11 above tests the hypothesis $Q'Hb = m$ when $Q'b$ is not estimable. (Page 25)

13. Functions that are estimable in an unrestricted model are also estimable in a restricted model. (Pages 30 and 34)

14. Hypotheses that are testable in an unrestricted model are also testable in a restricted model. (Pages 32 and 34)

15. In restricted models, functions of parameters and hypotheses relating to them are conditioned by the restrictions, such that if $Q'b$ is estimable in the unrestricted model then $(Q' + LP')b$, for L of appropriate order, is estimable in the restricted model. (Pages 33 and 35)

16. The accompanying table summarizes the expressions for residual sums of squares.

Residual Sum of Squares in hypothesis testing

Model $y = Xb + e$	Full Model $y = Xb + e$	Reduced Model under the estimable hypothesis $Q'b = m$.
Unrestricted	$SSR_0 = y'y - \hat{b}_0'X'y$ d.f. $= n - r$	$SSR_H = SSR_0 + (Q'\hat{b}_0 - m)'(Q'GQ)^{-1}(Q'\hat{b}_0 - m)$ d.f. $= n - (r - s)$
Restricted Restrictions $P'b = \alpha$, with $P'b$ estimable.	$SSR_{0'} = SSR_{H \rightarrow P'b=\alpha}$ $= SSR_0 + (P'\hat{b}_0 - \alpha)'(P'GP)^{-1}(P'\hat{b}_0 - \alpha)$ d.f. $= n - (r - p)$	$\begin{bmatrix} P' \\ Q' \end{bmatrix} b = \begin{bmatrix} \alpha \\ m \end{bmatrix} \equiv T'b = \tau$ $SSR_{H'} = SSR_0 + (T'\hat{b}_0 - \tau)'(T'GT)^{-1}(T'\hat{b}_0 - \tau)$ d.f. $= n - (r - s - p)$
Restricted Restrictions $P'b = \alpha$, with $P'b$ non-estimable.	$SSR_{0''} = SSR_0$ d.f. $= n - r$	$SSR_{H''} = SSR_H$ d.f. $= n - (r - s)$

Acknowledgements

Thanks go to D. S. Robson, N. S. Urquhart and the 1966 class of Plant Breeding 517 for assistance on numerous points of detail.

References

Elston and Bush (1964). The hypotheses that can be tested when there are interactions in an analysis of variance model. Biometrics 20, 681-698.

Graybill, Franklin A. (1961). An Introduction to Linear Statistical Models. Vol. I, McGraw-Hill, New York.

Rao, C. R. (1962). A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. J. Roy. Stat. Soc. (B) 24, 152-158.

Searle, S. R. (1965). Additional results concerning estimable functions and generalized inverse matrices. J. Roy. Stat. Soc. (B) 27, 480-490.

Searle, S. R. (1966). Matrix Algebra for the Biological Sciences (Including Applications in Statistics). John Wiley and Sons, New York, 296 pp.