J. SCI. HIROSHIMA UNIV. SER. A-I
31 (1967), 211-219

The Linear Hypotheses and Constraints

Sumiyasu YAMAMOTO and Yasunori FUJIKOSHI

(Received September 20, 1967)

1. Introduction and Summary

The purpose of this paper is to give a unified treatment of the classical least squares theory in the linear hypothesis model. The linear hypothesis model treated here can be summarized as

$$\boldsymbol{y} = \boldsymbol{\theta} + \boldsymbol{e} \tag{1.1}$$

where $\boldsymbol{y}$ is an $n \times 1$ vector of observations, $\boldsymbol{\theta}$ is the expectation of $\boldsymbol{y}$ which is known to belong to a specified linear subspace

$$\Xi = \{\boldsymbol{\theta} \mid \boldsymbol{\theta} = A\boldsymbol{\tau}, \quad B\boldsymbol{\tau} = 0\} \tag{1.2}$$

of the n dimensional Euclidean space E_n, and $\boldsymbol{e}$ is an $n \times 1$ vector of random errors which has the multivariate normal distribution with mean 0 and covariance matrix $\sigma^2 I_n$, σ^2 times the unit matrix I_n.

It is worthwhile to note that in our unified treatment no restriction is imposed on the known $n \times m$ matrix A and the known $l \times m$ matrix B. The matrix A may be called a design matrix. The matrix equation $B\boldsymbol{\tau}=0$ is a set of constraints imposed on the parameter vector $\boldsymbol{\tau}$. $B\boldsymbol{\tau}=0$ is in some cases a set of identifiability constraints of the parameter vector $\boldsymbol{\tau}$, a set of hypotheses to be tested and a set of more complex constraints. The matrices A and B and the parameter vector $\boldsymbol{\tau}$ jointly specify the linear subspace Ξ of E_n.

The least squares estimate of the parameter $\boldsymbol{\tau}$ in the extended sense and the projection operator to the space Ξ obtained by using the generalized inverse matrices are given in the Theorem of section 2. Some properties of the generalized inverse matrices and the projection operators are also given in section 2.

Our general formula given in the Theorem contains as its special cases the following three cases (i), (ii) and (iii):

(i) $\mathrm{rank}(A) = m$,

(ii) $\mathrm{rank}(A) < m$ and $\mathfrak{R}[A'] \supset \mathfrak{R}[B']$,

(iii) $\mathrm{rank}(A) < m$ and $\mathfrak{R}[A'] \cap \mathfrak{R}[B'] = \{0\}$,

where X' denotes the transposed matrix of X and $\mathfrak{R}[X]$ denotes a vector space spanned by the column vectors of a matrix X.

The case (i) is the simplest. The case (ii) has been treated in connection with the theory of testing testable hypotheses by many authors (c.f. Goldman and Zelen [1], Searle [7], Seber [8], [9]). The case (iii) has been treated in connection with the identifiability constraints of the parameter vector $\boldsymbol{\tau}$ (c.f. Reiersøl [5], Seber [9], Scheffé [10]). These results are given in section 3 as special cases of the general results.

Our unified treatment of the problem may be useful in obtaining the least squares estimates under various situations and in determining the likelihood ratio test statistics in the analysis of variance where the identifiability constraints are necessary to be introduced in the underlying models.

2. The main results

The matrix S^g is called a generalized inverse of a matrix S (Rao [3]) if it has the property

$$SS^gS=S. \tag{2.1}$$

The role of the generalized inverse matrices in the least squares theory is well known. The importance of the generalized inverse matrices stems from the fact that a consistent matrix equation

$$SX=Y \tag{2.2}$$

has a general solution

$$X=S^gY+(I-\mathfrak{P}[S'])Z \tag{2.3}$$

where $\mathfrak{P}[S']$ is the projection operator to the $\mathfrak{R}[S']$ and Z is an arbitrary matrix of appropriate dimensions. Although several special types of the generalized inverse matrices have been proposed by several authors [1], [2], [3], and [6], the generalized inverse matrices of the least restrictive type having only the property (2.1) are used in this paper.

LEMMA 1. *Let X, Y and Z be $n\times m$, $l\times m$ and $n\times t$ matrices, respectively, and suppose $\mathfrak{R}[X']\supset\mathfrak{R}[Y']$.* Then we have:

(i) *$X(X'X)^gX'$ is a symmetric idempotent matrix and is the projection operator to $\mathfrak{R}[X]$.*

(ii) $Y(X'X)^gX'X=Y$.

(iii) *$Y(X'X)^gY'$ is a symmetric matrix.*

(iv) $\mathfrak{R}[Y(X'X)^gY']=\mathfrak{R}[Y(X'X)^gX']=\mathfrak{R}[Y(X'X)^g]=\mathfrak{R}[Y]$.

(v) $\mathfrak{P}[X:Z]=\mathfrak{P}[X]+\mathfrak{P}[(I-\mathfrak{P}[X])Z]$.

PROOF. (i) The proof of (i) can be seen in Rao [4], p. 187. We, however, give a direct proof of (i).

Suppose that $\mathrm{rank}(X)=r$. Then there exist orthogonal matrices H_1 and

H_2 such that

$$H_1XH_2=\left(\begin{array}{c:c} D_r & O_{r,m-r} \\ \hdashline O_{n-r,r} & O_{n-r,m-r} \end{array}\right) \tag{2.4}$$

where D_r is a diagonal matrix of order r with $|D_r|\neq 0$ and $O_{p,q}$ is the $p\times q$ null matrix. Simple matrix multiplication gives

$$X'X=H_2\left(\begin{array}{c:c} D_r^2 & O_{r,m-r} \\ \hdashline O_{m-r,r} & O_{m-r,m-r} \end{array}\right)H_2' \tag{2.5}$$

Since $(X'X)^g$ satisfies $X'X(X'X)^gX'X=X'X$, $(X'X)^g$ must be of the form

$$H_2\left(\begin{array}{c:c} D_r^{-2} & U \\ \hdashline V & W \end{array}\right)H_2' \tag{2.6}$$

where U, V and W are arbitrary matrices of order $r\times(m-r)$, $(m-r)\times r$ and $(m-r)\times(m-r)$, respectively. From (2.4) and (2.6) we have

$$X(X'X)^gX'=H_1'\left(\begin{array}{c:c} I_r & O_{r,n-r} \\ \hdashline O_{n-r,r} & O_{n-r,n-r} \end{array}\right)H_1 \tag{2.7}$$

Hence we can see that $X(X'X)^gX'$ is a symmetric idempotent matrix and the projection operator to $\mathfrak{R}[X]$.

(ii) Since $\mathfrak{R}[X']\supset\mathfrak{R}[Y']$, there exists an $l\times n$ matrix K such that $Y=KX$. Thus, using (i) we have $Y(X'X)^gX'X=KX(X'X)^gX'X=KX=Y$.

(iii) Using (i) we have $\{Y(X'X)^gY'\}'=\{KX(X'X)^gX'K'\}'=KX(X'X)^gX'K'$ $=Y(X'X)^gY'$.

(iv) Since $\mathfrak{R}[Y(X'X)^gY']\subset\mathfrak{R}[Y(X'X)^gX']\subset\mathfrak{R}[Y(X'X)^g]\subset\mathfrak{R}[Y]$, it is sufficient to show that $\operatorname{rank}(Y)=\operatorname{rank}(Y(X'X)^gY')$. In fact, we have $\operatorname{rank}(Y)\geq$ $\operatorname{rank}(Y(X'X)^gY')=\operatorname{rank}\{(Y(X'X)^gX')(Y(X'X)^gX')'\}=\operatorname{rank}(Y(X'X)^gX')=$ $\operatorname{rank}(Y(X'X)^gX'X)=\operatorname{rank}(Y)$.

(v) Since $\mathfrak{P}[X]\mathfrak{P}[X:Z]=\mathfrak{P}[X]$, $\mathfrak{R}[\mathfrak{P}[X:Z]-\mathfrak{P}[X]]=\mathfrak{R}[(I-\mathfrak{P}[X])\cdot$ $\mathfrak{P}[X:Z])]=\mathfrak{R}[(I-\mathfrak{P}[X])[X:Z]]=\mathfrak{R}[(I-\mathfrak{P}[X])Z]$. Hence we have $\mathfrak{P}[X:Z]$ $=\mathfrak{P}[X]+\mathfrak{P}[(I-\mathfrak{P}[X])Z]$.

LEMMA 2. *The following three conditions* (i), (ii) *and* (iii) *are equivalent.*

(i) *The equations* $A\boldsymbol{\tau}=\boldsymbol{\theta}$, $B\boldsymbol{\tau}=0$ *have a solution* $\boldsymbol{\tau}$ *for every* $\boldsymbol{\theta}\in\mathfrak{R}[A]$.

(ii) $\mathfrak{R}[A']\cap\mathfrak{R}[B']=\{0\}$.

(iii) $(I-\mathfrak{P}[B_2])B_1=0$ *where* $B_1'=\mathfrak{P}[A']B'$ *and* $B_2'=(I-\mathfrak{P}[A'])B'$.

PROOF. It is well known that the conditions (i) and (ii) are equivalent (c.f. Seber [9], p. 101). We show that the conditions (ii) and (iii) are equivalent. We have

$$
\begin{aligned}
\mathfrak{R}[B'] &= \{\boldsymbol{\tau} \mid B\boldsymbol{\tau}=0\}^{\perp} \\
&= \left\{\boldsymbol{\tau} \,\middle|\, \begin{bmatrix} (I-\mathfrak{P}[B_2])B_1 \\ \mathfrak{P}[B_2]B_1+B_2 \end{bmatrix} \boldsymbol{\tau}=0 \right\}^{\perp} \\
&= \mathfrak{R}[B_1'(I-\mathfrak{P}[B_2]) : B_1'\mathfrak{P}[B_2]+B_2'] \\
&= \mathfrak{R}[B_1'(I-\mathfrak{P}[B_2])]+\mathfrak{R}[B_1'\mathfrak{P}[B_2]+B_2']
\end{aligned}
\tag{2.8}
$$

where $V^{\perp}$ denotes the orthogonal complement of a linear subspace V. It can be easily seen that

$$\mathfrak{R}[A'] \supset \mathfrak{R}[B_1'(I-\mathfrak{P}[B_2])] \tag{2.9}$$

and

$$\mathfrak{R}[A'] \cap \mathfrak{R}[B_1'\mathfrak{P}[B_2]+B_2'] = \{0\}. \tag{2.10}$$

Hence the conditions (ii) and (iii) are equivalent.

The least squares estimate $\hat{\boldsymbol{\tau}}$ of $\boldsymbol{\tau}$ is the value of $\boldsymbol{\tau}$ which minimizes

$$(\boldsymbol{y}-A\boldsymbol{\tau})'(\boldsymbol{y}-A\boldsymbol{\tau})$$

subject to $B\boldsymbol{\tau}=0$. By using the Lagrange multiplier method we can see that $\hat{\boldsymbol{\tau}}$ is identical with the solution of the equations

$$A'A\boldsymbol{\tau}+B'\boldsymbol{\lambda}=A'\boldsymbol{y} \tag{2.11}$$

and

$$B\boldsymbol{\tau}=0 \tag{2.12}$$

where $\boldsymbol{\lambda}$ is a vector of Lagrange multipliers. It is known but not explicitly stated that the equations (2.11) and (2.12) are consistent and have at least one solution $[\hat{\boldsymbol{\tau}}' : \hat{\boldsymbol{\lambda}}']$ for every $\boldsymbol{y}$ in E_n. The results, however, seem not to be so trivial. So, we give an algebraic proof of the consistency of equations (2.11) and (2.12) by using Lemma 2.

Multiplying (2.11) on the left by $\mathfrak{P}[A']$ and $(I-\mathfrak{P}[A'])$, respectively, the equation (2.11) is equivalent to

$$A'A\boldsymbol{\tau}+B_1'\boldsymbol{\lambda}=A'\boldsymbol{y} \tag{2.13}$$

and

$$B_2'\boldsymbol{\lambda}=0. \tag{2.14}$$

Since $\mathfrak{R}[A']=\mathfrak{R}[A'A : B_1']$, Lemma 2 shows that the consistency of (2.13), (2.14) and (2.12) can be proved by showing that

$$\mathfrak{R}\begin{bmatrix} A'A \\ B_1 \end{bmatrix} \cap \mathfrak{R}\begin{bmatrix} O & B' \\ B_2 & O \end{bmatrix} = \{0\}. \tag{2.15}$$

Suppose that

$$\begin{bmatrix} A'A \\ B_1 \end{bmatrix} \boldsymbol{u} = \begin{bmatrix} O & B' \\ B_2 & O \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{bmatrix}$$

holds for some vectors $\boldsymbol{u}$, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Then, since $B' = B_1'(I - \mathfrak{P}[B_2]) + (B_1'\mathfrak{P}[B_2] + B_2')$ and $\mathfrak{R}[B_1'\mathfrak{P}[B_2] + B_2'] \cap \mathfrak{R}[A'] = \{0\}$, we have

$$A'A\boldsymbol{u} = B_1'(I - \mathfrak{P}[B_2])\boldsymbol{v}_2 \tag{2,16}$$

and

$$B_1\boldsymbol{u} = B_2\boldsymbol{v}_1 \tag{2.17}$$

Solving (2.16) for $\boldsymbol{u}$, we have

$$\boldsymbol{u} = (A'A)^g B_1'(I - \mathfrak{P}[B_2])\boldsymbol{v}_2 + (I - \mathfrak{P}[A'])\boldsymbol{v}_3 \tag{2.18}$$

where $\boldsymbol{v}_3$ is an arbitrary vector. Substituting (2.18) into the equation $(I - \mathfrak{P}[B_2])B_1\boldsymbol{u} = 0$ obtained from (2.17), we have

$$(I - \mathfrak{P}[B_2])B_1(A'A)^g B_1'(I - \mathfrak{P}[B_2])\boldsymbol{v}_2 = 0.$$

Using Lemma 1. (ii), we get

$$B_1'(I - \mathfrak{P}[B_2])\boldsymbol{v}_2 = 0.$$

This implies $A'A\boldsymbol{u} = 0$. Hence $B_1\boldsymbol{u} = 0$. This completes the proof.

In general the solution $\hat{\boldsymbol{\tau}}$ is not unique. We, however, call $\hat{\boldsymbol{\tau}}$ the least squares estimate of $\boldsymbol{\tau}$ in the extended sense. Since $(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})'A'(\boldsymbol{y} - A\hat{\boldsymbol{\tau}}) = 0$ for all $\boldsymbol{\tau}$ satisfying $B\boldsymbol{\tau} = 0$, we have $(\boldsymbol{y} - \mathrm{A}\boldsymbol{\tau})'(\boldsymbol{y} - A\boldsymbol{\tau}) = (\boldsymbol{y} - A\hat{\boldsymbol{\tau}})'(\boldsymbol{y} - A\hat{\boldsymbol{\tau}}) + (\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})'A'A(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}) \geq (\boldsymbol{y} - A\hat{\boldsymbol{\tau}})'(\boldsymbol{y} - A\hat{\boldsymbol{\tau}})$ for a solution $\hat{\boldsymbol{\tau}}$ of the equations (2.11) and (2.12). Thus $\hat{\boldsymbol{\tau}}$ actually minimizes $(\boldsymbol{y} - A\boldsymbol{\tau})'(\boldsymbol{y} - A\boldsymbol{\tau})$ subject to $B\boldsymbol{\tau} = 0$. Moreover, if we have two solutions $(\hat{\boldsymbol{\tau}}_1, \hat{\boldsymbol{\lambda}}_1)$ and $(\hat{\boldsymbol{\tau}}_2, \hat{\boldsymbol{\lambda}}_2)$ of (2.11) and (2.12), we get $(\hat{\boldsymbol{\tau}}_1 - \hat{\boldsymbol{\tau}}_2)'A'A(\hat{\boldsymbol{\tau}}_1 - \hat{\boldsymbol{\tau}}_2) = 0$. This implies that $A\hat{\boldsymbol{\tau}}$, the projection of $\boldsymbol{y}$ to the space $\varXi$, is unique.

In the following theorem, we give an explicit solution of the least squares estimate $\hat{\boldsymbol{\tau}}$ of $\boldsymbol{\tau}$ in the extended sense and the projection operator $\mathrm{P}_{\varXi}$ to the space $\varXi$.

Theorem. *The least squares estimate $\hat{\boldsymbol{\tau}}$ of $\boldsymbol{\tau}$ in the extended sense and the projection operator $\mathrm{P}_{\varXi}$ to the space $\varXi$ are given by*

$$\begin{aligned} \hat{\boldsymbol{\tau}} = {} & \{I - (I - \mathfrak{P}[A'])B_2^g T\}[(A'A)^g A' - (A'A)^g S'\{S(A'A)^g S'\}^g S(A'A)^g A']\boldsymbol{y} \\ & + (I - \mathfrak{P}[A'])(I - \mathfrak{P}[B_2'])\boldsymbol{v} \end{aligned} \tag{2.19}$$

and

$$P_{\Xi} = A(A'A)^g A' - A(A'A)^g S'\{S(A'A)^g S'\}^g S(A'A)^g A' \tag{2.20}$$

where $S' = B_1'(I - \mathfrak{P}[B_2])$, $T' = B_1'\mathfrak{P}[B_2] + B_2'$, $B_1' = \mathfrak{P}[A']B'$, $B_2' = (I - \mathfrak{P}[A'])B'$ *and* $\boldsymbol{v}$ *is an arbitrary vector.*

PROOF. Solving (2.13) and (2.14), we obtain

$$\hat{\boldsymbol{\lambda}} = (I - \mathfrak{P}[B_2])\boldsymbol{u}_1 \tag{2.21}$$

and

$$\hat{\boldsymbol{\tau}} = (A'A)^g A'\boldsymbol{y} - (A'A)^g S'\boldsymbol{u}_1 + (I - \mathfrak{P}[A'])\boldsymbol{u}_2 \tag{2.22}$$

where $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are arbitrary vectors. To arrive at the desired solution, we have to adjust $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ so that (2.12) holds. Substituting (2.22) into (2.12), we have

$$B(A'A)^g S'\boldsymbol{u}_1 - B_2\boldsymbol{u}_2 = B(A'A)^g A'\boldsymbol{y}. \tag{2.23}$$

Multiplying on the left by $(I - \mathfrak{P}[B_2])$ and $\mathfrak{P}[B_2]$, respectively, the equation (2.23) is equivalent to

$$S(A'A)^g S'\boldsymbol{u}_1 = S(A'A)^g A'\boldsymbol{y} \tag{2.24}$$

and

$$T(A'A)^g S'\boldsymbol{u}_1 - B_2\boldsymbol{u}_2 = T(A'A)^g A'\boldsymbol{y}. \tag{2.25}$$

Solving (2.24) by using Lemma 1. (iv), we can write

$$\boldsymbol{u}_1 = \{S(A'A)^g S'\}^g S(A'A)^g A'\boldsymbol{y} + (I - \mathfrak{P}[S])\boldsymbol{u}_3 \tag{2.26}$$

and solving (2.25), we have

$$\begin{aligned}\boldsymbol{u}_2 = & -B_2^g T(A'A)^g A'\boldsymbol{y} + B_2^g T(A'A)^g S'\{S(A'A)^g S'\}^g S(A'A)^g A'\boldsymbol{y} \\ & + (I - \mathfrak{P}[B_2'])\boldsymbol{v}\end{aligned} \tag{2.27}$$

where $\boldsymbol{u}_3$ and $\boldsymbol{v}$ are arbitrary vectors. From (2.22), (2.26) and (2.27) we have a general solution of $\hat{\boldsymbol{\tau}}$ given in (2.19). Since $A\hat{\boldsymbol{\tau}} = P_{\Xi}\boldsymbol{y}$ for every $\boldsymbol{y}$ in E_n, we have an explicit formula of P_{Ξ} given in (2.20).

From our explicit formulas given in the Theorem we have the following corollaries.

COROLLARY 1. *The following conditions are equivalent.*

(i) $P_{\Xi} = A(A'A)^g A'$.

(ii) $S = (I - \mathfrak{P}[B_2])B_1 = 0$.

(iii) $\mathfrak{N}[A'] \cap \mathfrak{N}[B'] = \{0\}$.

(iv) $B'\hat{\boldsymbol{\lambda}} = 0$ *for every* $\boldsymbol{y}$.

PROOF. From (2.20) we have that $P_{\Xi}=A(A'A)^g A'$ if and only if $A(A'A)^g S'\{S(A'A)^g S'\}^g S(A'A)^g A'=\mathfrak{P}[A(A'A)^g S']=0$. Therefore

$$(\mathrm{i})\Leftrightarrow A(A'A)^g S'=0\Leftrightarrow 0=A'A(A'A)^g S'=S' \qquad \text{(by Lemma 1. (ii))}$$

Hence (i) and (ii) are equivalent. Lemma 2 shows that (ii) and (iii) are equivalent. From (2.21) and (2.26) we have

$$B'\hat{\boldsymbol{\lambda}}=S'\{S(A'A)^g S'\}^g S(A'A)^g A'\boldsymbol{y}. \tag{2.28}$$

Now, if (ii) is true, $B'\hat{\boldsymbol{\lambda}}=0$. Conversely, if (iv) is true, $S'\{S(A'A)^g S'\}^g S(A'A)^g A'=0$. Hence we get $S=0$. This completes the proof of Corollary 1.

If $\operatorname{rank}(B)=l$, the condition (iv) turns out to that $\hat{\boldsymbol{\lambda}}=0$. This special case was treated by Reiersøl [5].

COROLLARY 2. *The least squares estimate $\hat{\boldsymbol{\tau}}$ is uniquely determined if and only if* $\operatorname{rank}[A':B']=m$.

PROOF. Suppose that $\hat{\boldsymbol{\tau}}$ is uniquely determined. From (2.19) and Lemma 1. (v) we have $I-\mathfrak{P}[A':B']=0$. Hence $\operatorname{rank}[A':B']=m$. Conversely, suppose that $\operatorname{rank}[A':B']=m$. We have that $(A'A+B'B)\hat{\boldsymbol{\tau}}=A'A\hat{\boldsymbol{\tau}}=A'P_{\Xi}\boldsymbol{y}$ and $(A'A+B'B)$ is nonsingular. Hence $\hat{\boldsymbol{\tau}}$ is unique.

3. Some special cases

In this section we apply our Theorem to some special cases. Given that $\mathcal{E}\boldsymbol{y}=\boldsymbol{\theta}\in\Omega$, a linear subspace in E_n, then a linear hypothesis is a hypothesis which states that $\boldsymbol{\theta}\in\omega$, a linear subspace of Ω. The role of P_Ω and P_ω, the projection operators to Ω and ω, in testing a hypothesis is well known (c.f. Seber [8], [9]).

(a) Consider the case where underlying assumption is

$$\boldsymbol{\theta}\in\Omega=\{\boldsymbol{\theta}\,|\,\boldsymbol{\theta}=A\boldsymbol{\tau}\} \tag{3.1}$$

and suppose we wish to test a hypothesis $H\boldsymbol{\tau}=0$:

$$\boldsymbol{\theta}\in\omega=\{\boldsymbol{\theta}\,|\,\boldsymbol{\theta}=A\boldsymbol{\tau},\quad H\boldsymbol{\tau}=0\} \tag{3.2}$$

(i) $\operatorname{rank}(A)=m$: In this case we have $(A'A)^g=(A'A)^{-1}$ and for $\boldsymbol{\theta}\in\omega$ we have $B_1=H$, $B_2=0$ and $S=H$. Thus we have that $\hat{\boldsymbol{\tau}}_\Omega=(A'A)^{-1}A'\boldsymbol{y}$, $P_\Omega=A(A'A)^{-1}A'$ and $P_\omega=A(A'A)^{-1}A'-A(A'A)^{-1}H'\{H(A'A)^{-1}H'\}^g H(A'A)^{-1}A'$, where $\hat{\boldsymbol{\tau}}_\Omega$ denotes the least squares estimate of $\boldsymbol{\tau}$ under Ω. Therefore, we have

$$P_\Omega-P_\omega=A(A'A)^{-1}H'\{H(A'A)^{-1}H'\}^g H(A'A)^{-1}A' \tag{3.3}$$

and

$$y'(\mathrm{P}_\Omega-\mathrm{P}_\omega)y=(H\hat{\tau}_\Omega)'\{\mathrm{cov}(H\hat{\tau}_\Omega,\ \hat{\tau}'_\Omega H')\}^g H\hat{\tau}_\Omega\sigma^2. \tag{3.4}$$

When the rows of H are linearly independent, the generalized inverses of the matrices in (3.3) and (3.4) can be replaced by the inverses of them.

(ii) $\mathrm{rank}(A)<m$ and $\mathfrak{R}[A']\supset\mathfrak{R}[H']$: In this case, since $B_1=H\mathfrak{P}[A']=H$, $B_2=0$ and $S=H$ for $\boldsymbol{\theta}\in\omega$, we have that $\hat{\tau}_\Omega=(A'A)^gA'y+(I-\mathfrak{P}[A'])v$, $\mathrm{P}_\Omega=A(A'A)^gA'$ and $\mathrm{P}_\omega=A(A'A)^gA'-A(A'A)^gH'\{H(A'A)^gH'\}^gH(A'A)^gA'$. Therefore, we have

$$\mathrm{P}_\Omega-\mathrm{P}_\omega=A(A'A)^gH'\{H(A'A)^gH'\}^gH(A'A)^gA' \tag{3.5}$$

and

$$y'(\mathrm{P}_\Omega-\mathrm{P}_\omega)y=(H\tau_\Omega)'\{\mathrm{cov}(H\hat{\tau}_\Omega,\ \tau'_\Omega H')\}^g(H\hat{\tau}_\Omega)\sigma^2. \tag{3.6}$$

When the rows of H are linearly independent, Lemma 1. (iv) shows that $\{H(A'A)^gH'\}^g$ and $\{\mathrm{cov}(H\hat{\tau}_\Omega,\ \hat{\tau}'_\Omega H')\}^g$ can be replaced by $\{H(A'A)^gH'\}^{-1}$ and $\{\mathrm{cov}(H\hat{\tau}_\Omega,\ \hat{\tau}'_\Omega H')\}^{-1}$, respectively.

(b) Consider the case where identifiability constraints are introduced in underlying assumption:

$$\boldsymbol{\theta}\in\Omega=\{\boldsymbol{\theta}\,|\,\boldsymbol{\theta}=A\tau,\quad C\tau=0\} \tag{3.7}$$

where $C\tau=0$ is a set of identifiability constraints (c.f. Seber [9], Scheffé [10]), i.e., the matrix C satisfies the conditions $\mathfrak{R}[A']\cap\mathfrak{R}[C']=\{0\}$ and $\mathrm{rank}[A'\,:\,C']=m$. Suppose we wish to test the hypothesis $H\tau=0$:

$$\boldsymbol{\theta}\in\omega=\{\boldsymbol{\theta}\,|\,\boldsymbol{\theta}=A\tau,\ \begin{bmatrix}C\\H\end{bmatrix}\tau=0\}. \tag{3.8}$$

Using Lemma 2, Theorem and Corollary 1, we have $\hat{\tau}_\Omega=\{I-(I-\mathfrak{P}[A'])C_2^gC\}\cdot(A'A)^gA'y$, $\mathrm{P}_\Omega=A(A'A)^gA'$ and $\mathrm{P}_\omega=A(A'A)^gA'-A(A'A)^gS'\{S(A'A)^gS'\}^gS(A'A)^gA'$, where $C_2'=(I-\mathfrak{P}[A'])C'$, $S=(I-\mathfrak{P}[B_2])B_1$, $B_1'=\mathfrak{P}[A']B'$, $B_2'=(I-\mathfrak{P}[A'])B'$ and $B'=[C'\,:\,H']$.

(i) $\mathfrak{R}[A']\supset\mathfrak{R}[H']$: In this case we have $S=\left(I-\mathfrak{P}\begin{bmatrix}C_2\\0\end{bmatrix}\right)\begin{bmatrix}C_1\\H\end{bmatrix}=(I-\mathfrak{P}[C_2])C_1+H=H$, where $C_1'=\mathfrak{P}[A']C'$. Thus we have

$$\mathrm{P}_\Omega-\mathrm{P}_\omega=A(A'A)^gH'\{H(A'A)^gH'\}^gH(A'A)^gA' \tag{3.9}$$

and

$$y'(\mathrm{P}_\Omega-\mathrm{P}_\omega)y=(H\hat{\tau}_\Omega)'\{\mathrm{cov}(H\hat{\tau}_\Omega,\ \hat{\tau}'_\Omega H')^gH\hat{\tau}_\Omega\sigma^2. \tag{3.10}$$

We note that if $\mathfrak{R}[A']\supset\mathfrak{R}[H']$, or if $H\tau$ is a set of estimable parametric functions, then P_Ω- P_ω does not depend on the identifiability constraints.

(ii) $\mathfrak{R}[A']\not\supset\mathfrak{R}[H']$: In this case

$$S=\left(I-\mathfrak{P}\begin{bmatrix}C_2\\H_2\end{bmatrix}\right)\begin{bmatrix}C_1\\H_1\end{bmatrix} \tag{3.11}$$

where $\begin{bmatrix} C_1 \\ H_1 \end{bmatrix} = \begin{bmatrix} C \\ H \end{bmatrix} \mathfrak{P}(A')$ and $\begin{bmatrix} C_2 \\ H_2 \end{bmatrix} = \begin{bmatrix} C \\ H \end{bmatrix} (I - \mathfrak{P}(A'))$. Thus we have

$$\mathrm{P}_\Omega - \mathrm{P}_\omega = A(A'A)^g S' \{S(A'A)^g S'\}^g S(A'A)^g A' \tag{3.12}$$

and

$$\boldsymbol{y}'(\mathrm{P}_\Omega - \mathrm{P}_\omega)\boldsymbol{y} = (S\hat{\boldsymbol{\tau}}_\Omega)' \{\mathrm{cov}(S\hat{\boldsymbol{\tau}}_\Omega, \hat{\boldsymbol{\tau}}'_\Omega S')\}^g (S\hat{\boldsymbol{\tau}}_\Omega)\sigma^2 \tag{3.13}$$

These results show that the sum of squares appropriate for testing the hypothesis $H\boldsymbol{\tau}=0$ depends upon how the parameter vector $\boldsymbol{\tau}$ is defined by the identifiability constraints. Such a situation happens in multi-way analysis of variance where interactions are assumed in the models.

References

[1] Goldman, A. J. and Zelen, M. (1964), Weak generalized inverses and minimum variance linear unbiased estimators. *J. Res. Nat. Bur. Stand.* **68** B, 151–172.

[2] Penrose, R. (1955), A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.*, **51**, 406–418.

[3] Rao, C. R. (1962), A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. *J. Roy. Statist. Soc., Ser.* B, **24**, 152–158.

[4] Rao, C. R. (1965), *Linear statistical inference and its applications.* Wiley, New York.

[5] Reiersøl, O. (1963), Identifiability, estimability, phenorestricting specifications, and zero Lagrange multipliers in the analysis of variance. *Skand. Aktuariietidsk., Häft.* **3–4**, 131–142.

[6] Rohde, C. A. (1966), Some results on generalized inverses. *Siam Review*, **8**, 201–205.

[7] Searle, S. R. (1965), Additional results concerning estimable functions and generalized inverse matrices. *J. Roy. Statist. Soc.*, **27**, 486–490.

[8] Seber, G. A. F. (1964), The linear hypothesis and idempotent matrices. *J. Roy. Statist. Soc., Ser.* B, **26**, 773–779.

[9] Seber, G. A. F. (1966), *The linear hypothesis.* Griffin, London.

[10] Scheffé, H. (1959), *The analysis of variance.* Wiley, New, York.

Department of Mathematics
Faculty of Science
Hiroshima University