Title	On Lagrange multipliers in normal equations
Sub Title	
Author	鷲尾, 泰俊(Washio, Yasutoshi)
Publisher	慶応義塾大学藤原記念工学部
Publication year	1970
Jtitle	Proceedings of the Fujihara Memorial Faculty of Engineering Keio University Vol.23, No.94/95 (1970.) ,p.87- 99
JaLC DOI	
Abstract	When there exist linear constraints among parameters in the linear statistical model, the method of least squares seeks for the values of parameters which minimize the sum of squares of errors subject to the given linear constraints. For this minimization the Lagrange multiplier method is commonly used. This paper treats of the problem of finding whether any Lagrange multipliers in the normal equations are equal to zero or not, before solving the normal equations. A necessary and sufficient condition that the Lagrange multipliers corresponding to any subset of linear constraints can be taken to be equal to zero is given. A sufficient condition for any Lagrange multipliers to be equal to zero is also given. Further a method for finding zero Lagrange multipliers is proposed and is illustrated by some examples.
Notes	
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50001004-00230095-0087

Proc. Fujihara Mem. Fac. Eng. Keio Univ.
Vol. 23, Nos. 3 & 4, 1970, pp. 87–99.

On Lagrange Multipliers in Normal Equations

(Received January 16, 1971)

Yasutoshi WASHIO*

Abstract

When there exist linear constraints among parameters in the linear statistical model, the method of least squares seeks for the values of parameters which minimize the sum of squares of errors subject to the given linear constraints. For this minimization the Lagrange multiplier method is commonly used. This paper treats of the problem of finding whether any Lagrange multipliers in the normal equations are equal to zero or not, before solving the normal equations.

A necessary and sufficient condition that the Lagrange multipliers corresponding to any subset of linear constraints can be taken to be equal to zero is given. A sufficient condition for any Lagrange multipliers to be equal to zero is also given. Further a method for finding zero Lagrange multipliers is proposed and is illustrated by some examples.

1. Introduction

Consider the linear statistical model

$$\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{e}\,, \tag{1.1}$$

where $\boldsymbol{y}$ is an $n\times 1$ vector of observations, X is a known $n\times p$ matrix of rank r, $\boldsymbol{\beta}$ is a $p\times 1$ vector of unknown parameters and $\boldsymbol{e}$ is an $n\times 1$ vector of errors with expectation $E(\boldsymbol{e})=\boldsymbol{0}$ and with variance $E(\boldsymbol{e}\boldsymbol{e}')=\sigma^2 I$, where a prime denotes the transpose of a matrix, σ^2 is an unknown positive constant and I is the $n\times n$ identity matrix. The matrix X may be called a design matrix. This paper deals with the linear estimation of $\boldsymbol{\beta}$ in the case where there exist t linear constraints among parameters

$$B\boldsymbol{\beta}=\boldsymbol{0}\,, \tag{1.2}$$

where B is a known $t\times p$ matrix. This situation arises, for example, in the analysis of variance model or in the problem of testing linear hypothesis in the linear statistical model. The least squares estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is the value of $\boldsymbol{\beta}$ which minimizes $(\boldsymbol{y}-X\boldsymbol{\beta})'(\boldsymbol{y}-X\boldsymbol{\beta})$ subject to the constraints (1. 2). For this minimization the Lagrange multiplier method is commonly used.

---

* 鷲 尾 泰 俊 Associate Professor, Faculty of Engineering, Keio University.

Let $\boldsymbol{\lambda}'=(\lambda_1, \lambda_2, \cdots, \lambda_t)$ be a vector of Lagrange multipliers corresponding to the linear constraints (1.2). Then the normal equations for obtaining the least squares estimates $\hat{\boldsymbol{\beta}}$ are given by

$$\begin{aligned} X'X\hat{\boldsymbol{\beta}}+B'\boldsymbol{\lambda}&=X'\boldsymbol{y} \\ B\hat{\boldsymbol{\beta}} \qquad\quad &=\mathbf{0}. \end{aligned} \tag{1.3}$$

It can be shown that the normal equations (1.3) are consistent and have at least one solution $(\hat{\boldsymbol{\beta}}' \boldsymbol{\lambda}')$ for every $\boldsymbol{y}$ (see, for example, Yamamoto, S. and Fujikoshi, Y. [6]). It can be also shown that a least squares estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ satisfies the normal equations (1.3), and that any solution $\hat{\boldsymbol{\beta}}$ of the normal equations, which is a function of $\boldsymbol{y}$ only, is a least squares estimate of $\boldsymbol{\beta}$.

At this situation, it is of importance for us to know before solving the normal equations whether each λ_i is equal to zero (or can be taken to be equal to zero) for every $\boldsymbol{y}$ or not. Because, when some of λ_i, $i=1.2, \cdots, t$, are equal to zero, the normal equations become simpler and so it facilitates us to solve them.

This problem has been investigated by many authors. Mann, H. B. [1] gave a sufficient condition for the Lagrange multiplier corresponding to a particular constraint to be equal to zero for every $\boldsymbol{y}$ under some assumptions on the design matrix and the constraints. A generalization of the Mann's theorem was given by Masuyama, M. [2]. A set of linear constraints $B\boldsymbol{\beta}=\mathbf{0}$ is called a set of identifiability constraints if the matrix B satisfies the following conditions (a) and (b): (a) The composite matrix $\begin{bmatrix} X \\ B \end{bmatrix}$ has rank p. (b) No linear combination of the rows of B is a linear combination of the rows of X except $\mathbf{0}'$. It is known that if the constraints $B\boldsymbol{\beta}=\mathbf{0}$ is a set of identifiability constraints, then the Lagrange multipliers corresponding to these constraints are equal to zero for every $\boldsymbol{y}$ (see, for example, Plackett, R. L. [3] and Seber, G. A. F. [5]). Reiersøl, O. [4] gave a necessary and sufficient condition for the Lagrange multipliers corresponding to any subset of linear constraints to be equal to zero for every $\boldsymbol{y}$, assuming that all the constraints are linearly independent. Yamamoto, S. and Fujikoshi, Y. [6] gave a necessary and sufficient condition that $B'\boldsymbol{\lambda}=\mathbf{0}$ for every $\boldsymbol{y}$ in the normal equations (1.3), without assuming linear independency of the linear constraints $B\boldsymbol{\beta}=\mathbf{0}$.

In this paper, we shall give a necessary and sufficient condition that the Lagrange multipliers corresponding to any subset of the given linear constraints can be taken to be equal to zero for every $\boldsymbol{y}$, without assuming linear independency of the constraints, in Section 2. A sufficient condition for any Lagrange multipliers to be equal to zero is also given in Section 2. In Section 3, a method for finding zero Lagrange multipliers by making use of our results is proposed and is illustrated by some examples. Section 4 contains some remarks on Lagrange multipliers.

2. Conditions That Lagrange Multipliers Are Equal to Zero or Can Be Taken to Be Equal to Zero

Throughout this paper, a linear combination of row vectors of the coefficient matrix B in the linear constraints $B\boldsymbol{\beta}=\mathbf{0}$ plays an important role. Denote this linear combination by $\boldsymbol{l}'$. Then

$$\boldsymbol{l}=B'\boldsymbol{c}=c_1\boldsymbol{b}_1+c_2\boldsymbol{b}_2+\cdots+c_t\boldsymbol{b}_t\,,$$

where $B=\begin{bmatrix}\boldsymbol{b}_1'\\ \boldsymbol{b}_2'\\ \vdots\\ \boldsymbol{b}_t'\end{bmatrix}$ and $\boldsymbol{c}'=(c_1, c_2, \cdots, c_t)$ is an arbitrary vector.

The range space of a matrix A, or the subspace spanned by the columns of A will be denoted by $\mathscr{R}[A]$. The orthogonal complement of $\mathscr{R}[A]$ will be denoted by $\{\mathscr{R}[A]\}^{\perp}$.

THEOREM 1. *Suppose that there exists a set of linear constraints among parameters*

$$B\boldsymbol{\beta}=\begin{bmatrix}B_1\\ \hdashline B_2\end{bmatrix}\boldsymbol{\beta}=\mathbf{0} \tag{2. 1}$$

in the linear statistical model (1. 1) *(B_2 may be 0), and let $\boldsymbol{\lambda}'=[\boldsymbol{\lambda}_1' \vdots \boldsymbol{\lambda}_2']$ be a vector of Lagrange multipliers corresponding to the constraints* (2. 1). *Then $\boldsymbol{\lambda}_1$ can be taken to be equal to $\mathbf{0}$ for every $\boldsymbol{y}$ as a consistent solution to the normal equations*) if and only if*

$$\boldsymbol{l}\in\mathscr{R}[X'] \quad \textit{implies} \quad \boldsymbol{l}\in\mathscr{R}[B_2']\,. \tag{2. 2}$$

Proof. The normal equations can be writtem as

$$\begin{aligned} X'X\hat{\boldsymbol{\beta}}+B_1'\boldsymbol{\lambda}_1+B_2'\boldsymbol{\lambda}_2&=X'\boldsymbol{y}\\ B_1\hat{\boldsymbol{\beta}}\qquad\qquad\qquad&=\mathbf{0}\\ B_2\hat{\boldsymbol{\beta}}\qquad\qquad\qquad&=\mathbf{0}\,. \end{aligned} \tag{2. 3}$$

Then $B_1'\boldsymbol{\lambda}_1=\mathbf{0}$ for every $\boldsymbol{y}$ in (2. 3) if and only if the equations

$$\begin{bmatrix}X'X & B_2'\\ B_1 & 0\\ B_2 & 0\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\beta}}\\ \boldsymbol{\lambda}_2\end{bmatrix}=\begin{bmatrix}X'\boldsymbol{y}\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix} \tag{2. 4}$$

are consistent for every $\boldsymbol{y}$. The last statement is equivalent to each of the following

*) In what follows, we shall use the expression such that λ_1 can be taken to be equal to $\mathbf{0}$ for every $\boldsymbol{y}$ as a consistent solution to the normal equations instead of $B_1'\lambda_1=\mathbf{0}$ for every $\boldsymbol{y}$ in the normal equations.

statements (2. 5) and (2. 6).

$$\left\{\mathcal{R}\begin{bmatrix} X'X & B_2' \\ B_1 & 0 \\ B_2 & 0 \end{bmatrix}\right\}^{\perp} = \left\{\mathcal{R}\begin{bmatrix} X'X & B_2' & X'\boldsymbol{y} \\ B_1 & 0 & \mathbf{0} \\ B_2 & 0 & \mathbf{0} \end{bmatrix}\right\}^{\perp} \quad \text{for every } \boldsymbol{y}. \tag{2. 5}$$

$$\begin{aligned} &X'X\boldsymbol{u}+B_1'\boldsymbol{v}_1+B_2'\boldsymbol{v}_2=\mathbf{0} \quad \text{and} \quad B_2\boldsymbol{u}=\mathbf{0} \\ &\text{imply} \quad X\boldsymbol{u}=\mathbf{0} \quad \text{and} \quad B_1'\boldsymbol{v}_1+B_2'\boldsymbol{v}_2=\mathbf{0}\,. \end{aligned} \tag{2. 6}$$

Now we shall prove that the statements (2. 6) is equivalent to the following statement (2. 7).

$$\begin{aligned} &X'\boldsymbol{p}+B_1'\boldsymbol{q}_1+B_2'\boldsymbol{q}_2=\mathbf{0} \quad \text{implies} \\ &B_1'\boldsymbol{q}_1+B_2'\boldsymbol{q}_2=-B_2'\boldsymbol{v}_2 \quad \text{for some } \boldsymbol{v}_2\,. \end{aligned} \tag{2. 7}$$

We first prove that (2. 6) implies (2. 7). Suppose that

$$X'\boldsymbol{p}+B_1'\boldsymbol{q}_1+B_2'\boldsymbol{q}_2=\mathbf{0} \tag{2. 8}$$

holds. By a property of normal equations, the equations

$$\begin{bmatrix} X'X & B_2' \\ B_2 & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v}_2 \end{bmatrix} = \begin{bmatrix} X'\boldsymbol{p} \\ \mathbf{0} \end{bmatrix} \tag{2. 9}$$

have a solution $(\boldsymbol{u}', \boldsymbol{v}_2')$ for every $\boldsymbol{p}$. From (2. 8) and (2. 9), we have

$$\begin{aligned} &X'X\boldsymbol{u}+B_1'\boldsymbol{q}_1+B_2'(\boldsymbol{v}_2+\boldsymbol{q}_2)=\mathbf{0} \\ &B_2\boldsymbol{u} \qquad\qquad\qquad\qquad\quad =\mathbf{0}\,. \end{aligned}$$

Then, from the statement (2. 6), we get

$$B_1'\boldsymbol{q}_1+B_2'\boldsymbol{q}_2=-B_2'\boldsymbol{v}_2\,.$$

Thus the statement (2. 7) holds. Conversely, suppose that the statement (2. 7) holds and that

$$\begin{aligned} &X'X\boldsymbol{u}+B_1'\boldsymbol{v}_1+B_2'\boldsymbol{v}_2=\mathbf{0} \\ &B_2\boldsymbol{u} \qquad\qquad\qquad\quad =\mathbf{0}\,. \end{aligned} \tag{2. 10}$$

From the statement (2. 7), we can write as

$$B_1'\boldsymbol{v}_1+B_2'\boldsymbol{v}_2=-B_2'\boldsymbol{s} \qquad \text{for some } \boldsymbol{s}\,. \tag{2. 11}$$

Substituting (2. 11) into (2. 10), we get

$$X'X\boldsymbol{u}-B_2'\boldsymbol{s}=\mathbf{0}\,. \tag{2. 12}$$

Multiplying (2. 12) on the left by $\boldsymbol{u}'$ and making use of (2. 10), we have $X\boldsymbol{u}=\mathbf{0}$. Substituting $X\boldsymbol{u}=\mathbf{0}$ into (2. 10), we obtain

$$B_1'\boldsymbol{v}_1+B_2'\boldsymbol{v}_2=\boldsymbol{0}\,.$$

Thus the statement (2. 6) holds.

The equivalence between the statements (2. 7) and (2. 2) is clear. It can be verified that the above argument also holds for $B_2=0$. Thus the proof is complete.

Theorem 1 contains the results obtained so far as special cases. A necessary and sufficient condition that $B'\boldsymbol{\lambda}=\boldsymbol{0}$ for every $\boldsymbol{y}$ in the normal equations (1. 3) is obtained by putting $B_2=0$ in Theorem 1. The result can be written as follows:

COROLLARY 1. *Suppose that there exists a set of linear constraints $B\boldsymbol{\beta}=\boldsymbol{0}$ among parameters in the linear statistical model* (1. 1), *and let $\boldsymbol{\lambda}$ be a vector of Lagrange multipliers corresponding to the constraints. Then $\boldsymbol{\lambda}$ can be taken to be equal to zero for every $\boldsymbol{y}$ as a consistent solution to the normal equations* (1. 3) *if and only if no linear combination of the rows of B is a linear combination of the rows of X except $\boldsymbol{0}'$.*

Corollary 1 is identical with the result given in Yamamoto, S. and Fujikoshi, Y. [6]. We can conclude from Corollary 1 that the case where *all* the Lagrange multipliers can be ignored in solving normal equations is only the case where the given set of constraints is a set or a subset of identifiability constraints.

When the linear constraints are linearly independent, the solution $\boldsymbol{\lambda}$ to the normal equations (1. 3) is unique, and the condition (2. 2) in Theorem 1 turns out that

$$\boldsymbol{l}=B_1'\boldsymbol{c}_1+B_2'\boldsymbol{c}_2\in\mathcal{R}[X'] \qquad \text{implies} \quad \boldsymbol{c}_1=\boldsymbol{0}\,.$$

Therefore, Theorem 1 leads to the next.

COROLLARY 2. *Suppose that $B\boldsymbol{\beta}=\boldsymbol{0}$ and $\boldsymbol{\lambda}'=[\boldsymbol{\lambda}_1' \vdots \boldsymbol{\lambda}_2']$ are as in Theorem* 1, *and further that the linear constraints $B\boldsymbol{\beta}=\boldsymbol{0}$ are linearly independent. Then $\boldsymbol{\lambda}_1$ is equal to $\boldsymbol{0}$ for every $\boldsymbol{y}$ if and only if*

$$\boldsymbol{l}=B_1'\boldsymbol{c}_1+B_2'\boldsymbol{c}_2\in\mathcal{R}[X'] \qquad \textit{implies} \quad \boldsymbol{c}_1=\boldsymbol{0}\,. \tag{2. 13}$$

Reiersøl, O. [4] gave the following condition (2. 14) instead of (2. 13).

$$\mathcal{R}[X' \vdots B_2'] \cap \mathcal{R}[B_1'] = \{\boldsymbol{0}\}\,. \tag{2. 14}$$

We shall prove the equivalence between (2. 13) and (2. 14). If (2. 14) is false, then there exist vectors $\boldsymbol{z}_1$, $\boldsymbol{z}_2$ and $\boldsymbol{z}_3$ such that

$$B_1'\boldsymbol{z}_1=B_2'\boldsymbol{z}_2+X'\boldsymbol{z}_3 \qquad \text{and} \qquad B_1'\boldsymbol{z}_1\neq\boldsymbol{0}\,. \tag{2. 15}$$

Since $\boldsymbol{z}_1\neq\boldsymbol{0}$, (2. 15) shows there exists a vector $\boldsymbol{l}$ such that

$$\boldsymbol{l}\in\mathcal{R}[X'] \qquad \text{and} \qquad \boldsymbol{c}_1\neq\boldsymbol{0}\,,$$

which is the negation of (2. 13). Thus (2. 13) implies (2. 14). Conversely, if (2. 13) is false, then there exist vectors $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ and $\boldsymbol{c}_3$ such that $\boldsymbol{c}_1 \neq \boldsymbol{0}$ and

$$[X' \vdots B_2'] \begin{bmatrix} \boldsymbol{c}_3 \\ \\ -\boldsymbol{c}_2 \end{bmatrix} = B_1' \boldsymbol{c}_1 . \tag{2. 16}$$

Since the columns of B_1' are linearly independent, $B_1'\boldsymbol{c}_1 \neq \boldsymbol{0}$. Hence (2. 16) shows the negation of (2. 14). Thus, (2. 14) implies (2. 13), and the equivalence is proved.

The next theorem gives a sufficient condition for $\boldsymbol{\lambda}_1$ *to be equal to* $\boldsymbol{0}$ for every $\boldsymbol{y}$.

THEOREM 2. *Suppose that there exists a set of linear constraints among parameters*

$$B\boldsymbol{\beta} = \begin{bmatrix} B_1 \\ \hdashline B_2 \end{bmatrix} \boldsymbol{\beta} = \boldsymbol{0} \tag{2. 17}$$

in the linear statistical model (1. 1) (B_2 *may be* 0), *and let* $\boldsymbol{\lambda}' = [\boldsymbol{\lambda}_1' \vdots \boldsymbol{\lambda}_2']$ *be a vector of Lagrange multipliers corresponding to the constraints* (2. 17). *If*

$$\boldsymbol{l} = B_1'\boldsymbol{c}_1 + B_2'\boldsymbol{c}_2 \in \mathscr{R}[X'] \qquad \textit{implies} \quad \boldsymbol{c}_1 = \boldsymbol{0} , \tag{2. 18}$$

then $\boldsymbol{\lambda}_1 = \boldsymbol{0}$ *for every* $\boldsymbol{y}$.

Proof. The normal equations are given by (2. 3). From the first equations in (2. 3), we have

$$B_1'\boldsymbol{\lambda}_1 + B_2'\boldsymbol{\lambda}_2 = X'(\boldsymbol{y} - X\hat{\boldsymbol{\beta}}) . \tag{2. 19}$$

Since the normal equations are consistent, (2. 19) implies that $B_1'\boldsymbol{\lambda}_1 + B_2'\boldsymbol{\lambda}_2$ must belong to $\mathscr{R}[X']$. Hence, by the condition (2. 18), $\boldsymbol{\lambda}_1 = \boldsymbol{0}$.

The next theorem gives a relationship among Lagrange multipliers and will be helpful in obtaining the values of non-zero Lagrange multipliers.

THEOREM 3. *Suppose that there exists a set of linear constraints among parameters*

$$B\boldsymbol{\beta} = \begin{bmatrix} B_1 \\ \hdashline B_2 \end{bmatrix} \boldsymbol{\beta} = \boldsymbol{0} \tag{2. 20}$$

in the linear statistical model (1. 1), *and let* $\boldsymbol{\lambda}' = [\boldsymbol{\lambda}_1' \vdots \boldsymbol{\lambda}_2']$ *be a vector of Lagrange multipliers corresponding to the constraints* (2. 20). *If* $\boldsymbol{l} = B_1'\boldsymbol{c}_1 + B_2'\boldsymbol{c}_2 \in \mathscr{R}[X']$ *implies that* $\boldsymbol{c}_1 = \boldsymbol{0}$ *and that there exists a vector* $\boldsymbol{c}_2$ *such that* $B_2'\boldsymbol{c}_2 \neq \boldsymbol{0}$ *and that there exists a unique relationship among components of vector* $\boldsymbol{c}_2$, *say* $\boldsymbol{a}_2'\boldsymbol{c}_2 = 0$, *then* $\boldsymbol{\lambda}_1 = \boldsymbol{0}$ *and a relationship* $\boldsymbol{a}_2'\boldsymbol{\lambda}_2 = 0$ *holds.*

Proof. The normal equations are given by (2. 3). By Theorem 2, we have $\boldsymbol{\lambda}_1 = \boldsymbol{0}$. Substituting $\boldsymbol{\lambda}_1 = \boldsymbol{0}$ into (2. 3), we get

$$B_2'\boldsymbol{\lambda}_2 = X'(\boldsymbol{y} - X\hat{\boldsymbol{\beta}}) .$$

Then, by the assumptions of the theorem, the relationship $\boldsymbol{a}_2'\boldsymbol{\lambda}_2 = 0$ must hold.

3. A Method for Finding Zero Lagrange Multipliers

We shall propose a method for finding zero Lagrange multipliers in practice by making use of theorems and corollaries given in Section 2.

When the rank of the design matrix X is p, $\boldsymbol{l}$ belongs to $\mathscr{R}[X']$ for any $\boldsymbol{c}$. It is, therefore, easy to apply the theorems and corollaries in finding zero Lagrange multipliers. For example, if the constraints $B\boldsymbol{\beta}=\boldsymbol{0}$ are linearly independent, then by Corollary 2 each of the components of $\boldsymbol{\lambda}$ is not equal to zero for every $\boldsymbol{y}$. Then we shall confine ourselves to the situation when the rank of the design matrix is less than p.

In order to examine whether each of the Lagrange multipliers can be taken to be equal to zero for every $\boldsymbol{y}$ or not, we at first want to know the set of values of $\boldsymbol{c}'=(c_1, c_2, \cdots, c_t)$ when $\boldsymbol{l}$ belongs to $\mathscr{R}[X']$. A matrix, say D, is introduced for this purpose. If the rank of $X, r,$ is less than p, then there exists a $p\times(p-r)$ matrix D of rank $(p-r)$ such that $XD=0$. The matrix D may be called a matrix which shows linear dependency among columns of X. In the analysis of variance model it will be easy to find out the matrix D.

Since the columns of D can be regarded as a basis for $\{\mathscr{R}[X']\}^{\perp}$, a p-dimensional vector $\boldsymbol{l}$ belongs to $\mathscr{R}[X']$ if and only if $\boldsymbol{l}'D=\boldsymbol{0}'$. Hence the set of solutions $\boldsymbol{c}'=(c_1, c_2, \cdots, c_t)$ to the equations $\boldsymbol{l}'D=\boldsymbol{0}'$ gives the set of values of $\boldsymbol{c}$, when $\boldsymbol{l}$ belongs to $\mathscr{R}[X']$.

Therefore, we can find zero Lagrange multipliers by checking the following two or three steps:

Step 1. Find a matrix D which shows linear dependency among columns of X from the design matrix X.

Step 2. Obtain the set of solutions $\boldsymbol{c}'=(c_1, c_2, \cdots, c_t)$ to the equations $\boldsymbol{l}'D=\boldsymbol{0}'$, where $\boldsymbol{l}=c_1\boldsymbol{b}_1+\cdots+c_t\boldsymbol{b}_t$.

Here we shall note that the equations $\boldsymbol{l}'D=\boldsymbol{0}'$ have t unknowns and $(p-r)$ equations. At this step, Theorem 2 and Theorem 3 can be applied. If the constraints are linearly independent, then Corollary 2 can also be applied at this step and we can find zero Lagrange multipliers.

Step 3. By making use of the result of Step 2, find a matrix $\bar{B}$ such that $\boldsymbol{l}$ must belong to $\mathscr{R}[\bar{B}']$, where each row of $\bar{B}$ must consist of a row of B, when $\boldsymbol{l}$ belongs to $\mathscr{R}[X']$. Clearly $\bar{B}$ is not unique, but the $\bar{B}$ with the smallest number of rows is preferable for the purpose of making the normal equations simpler.

If the constraints are not necessarily linearly independent, Theorem 1 can be applied at this step and we can find zero Lagrange multipliers.

Now we shall illustrate the method mentioned above by some examples.

EXAMPLE 1. The results given by Mann, H. B. [1] and Masuyama, M. [2].

Here we shall refer the result given by Masuyama. The assumptions adopted in it can be given in an equivalent form as follows: i) The $n\times p$ design matrix X has the properties (a) and (b). (a) The pth column consists solely of 1. (b) The sum of the first q columns of X is equal to the pth column of X. ii) In the linear constraints $B\boldsymbol{\beta}=\mathbf{0}$, the matrix B has the form such that

$$B=\left[\begin{array}{ccc:ccc:c} b_{11} & b_{12}\cdots & b_{1q} & b_{1q+1} & \cdots & b_{1p-1} & 0 \\ \hdashline & & & b_{2q+1} & \cdots & b_{2p-1} & \\ & 0 & & & & & 0 \\ & & & b_{tq+1} & \cdots & b_{tp-1} & \end{array}\right]\equiv\begin{bmatrix} \boldsymbol{b}_1' \\ \boldsymbol{b}_2' \\ \vdots \\ \boldsymbol{b}_t' \end{bmatrix}.$$

From the assumption (b) on X, we can choose a matrix D such that

$$D=\left[\begin{array}{c:c} \begin{matrix}1\\ \vdots \\ 1\end{matrix} & * \\ \hdashline \begin{matrix}0\\ \vdots \\ 0\end{matrix} & * \\ \hdashline -1 & * \end{array}\right]\begin{matrix}\}q\\ \\ \\ \end{matrix},$$

as a matrix which shows linear dependency among columns of X. Then, from the assumption on the matrix B, we have

$$\boldsymbol{l}'D=(c_1\boldsymbol{b}_1'+\cdots+c_t\boldsymbol{b}_t')D=\left(c_1\sum_j^q b_{1j}, *, \cdots, *\right).$$

Hence if $\sum_1^q b_{1j}\neq 0$, then $\boldsymbol{l}'D=\mathbf{0}'$ implies $c_1=0$. The last statement means that

$$\boldsymbol{l}\in\mathcal{R}[X'] \quad \text{implies} \quad c_1=0.$$

Thus, if $\sum_1^q b_{1j}\neq 0$, then $\lambda_1=0$ by Theorem 2. We note that the assumption (a) on X can be removed. If Theorem 1 is applied, we can conclude normal equations have a solution with $\lambda_1=0$.

EXAMPLE 2. The two-way layout with equal numbers of observations in the cells.

If y_{ijk} denotes the kth observation in the (i,j) cell, then y_{ijk} can be written as

$$y_{ijk}=\mu+\alpha_i+\beta_j+\gamma_{ij}+e_{ijk}, \qquad (i=1,2;\ j=1,2,3;\ k=1,2) \tag{3.1}$$

where μ represents the general mean, α_i the main effect of factor A, β_j the main effect of factor B, γ_{ij} the interaction between A and B and e_{ijk} the random error. Clearly the design matrix is not of full rank, and usually a set of linear constraints

$$\begin{aligned}&\sum_i \alpha_i=0\,, \qquad \sum_j \beta_j=0\\ &\sum_j \gamma_{ij}=0 \quad (i=1,2)\,, \qquad \sum_i \gamma_{ij}=0 \quad (j=1,2,3)\end{aligned} \tag{3. 2}$$

is introduced as a set of identifiability constraints. Let us consider the problem of obtaining the least squares estimates of parameters when $\alpha_1=\alpha_2=0$. This problem arises when a hypothesis H_0: $\alpha_1=\alpha_2=0$ is tested.

When $\alpha_1=\alpha_2=0$, the model (3. 1) turns out to be

$$y_{ijk}=\mu+\beta_j+\gamma_{ij}+e_{ijk}\,, \tag{3. 3}$$

and a set of constraints (3. 2) turns out to be

$$\sum_j \beta_j=0\,, \qquad \sum_j \gamma_{ij}=0 \quad (i=1,2)\,, \qquad \sum_i \gamma_{ij}=0 \quad (j=1,2,3)\,. \tag{3. 4}$$

If we consider the linearly independent constraints only, then the constraints (3. 4) can be written in our notations as

$$B=\left[\begin{array}{c|ccc|ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1\\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right]\equiv\begin{bmatrix} \boldsymbol{b}_1' \\ \boldsymbol{b}_2' \\ \boldsymbol{b}_3' \\ \boldsymbol{b}_4' \\ \boldsymbol{b}_5' \end{bmatrix},$$

$$\boldsymbol{\beta}'=(\mu,\ \beta_1,\ \beta_2,\ \beta_3,\ \gamma_{11},\ \gamma_{12},\ \gamma_{13},\ \gamma_{21},\ \gamma_{22},\ \gamma_{23})\,.$$

The design matrix X has 12 rows and 10 columns and has rank 6. We can choose a matrix D such that

$$D=\left[\begin{array}{cccc} 1 & 1 & 0 & 0\\ \hline -1 & 0 & 1 & 0\\ -1 & 0 & 0 & 1\\ -1 & 0 & 0 & 0\\ \hline 0 & -1 & -1 & 0\\ 0 & -1 & 0 & -1\\ 0 & -1 & 0 & 0\\ \hline 0 & -1 & -1 & 0\\ 0 & -1 & 0 & -1\\ 0 & -1 & 0 & 0 \end{array}\right].$$

From the equations $\boldsymbol{l}'D=(c_1\boldsymbol{b}_1'+\cdots+c_5\boldsymbol{b}_5')D=\boldsymbol{0}'$, we have

$$c_1=c_4=c_5=0 \quad \text{and} \quad c_2+c_3=0\,.$$

Since the constraints are linearly independent, Corollary 2 can be applied and we get

$$\lambda_1=\lambda_4=\lambda_5=0\,, \quad \lambda_2\neq 0 \quad \text{and} \quad \lambda_3\neq 0\,.$$

Furthermore, by Theorem 3, we can see there exists a relationship such that $\lambda_2+\lambda_3=0$.

EXAMPLE 3. The two-way layout with one observation per cell.

If y_{ij} denotes the observation in the (i, j) cell, then y_{ij} can be written as

$$y_{ij}=\mu+\alpha_i+\beta_j+e_{ij}\,, \qquad (i=1, 2\,;\ j=1, 2, 3)$$

where μ, α_i, and β_j are as in Example 2 and e_{ij} the random error. The design matrix is not of full rank, and usually a set of linear constraints

$$\sum_i \alpha_i=0\,, \qquad \sum_j \beta_j=0 \tag{3. 5}$$

is introduced as a set of identifiability constraints. Let us consider the problem of obtaining the least squares estimates of parameters subject to an additional constraint

$$2\beta_1-\beta_2=0\,. \tag{3. 6}$$

In order to illustrate our method in the case where the linear constraints are not necessarily linearly independent, suppose that there exists a further linear constraint

$$3\beta_1+\beta_3=0\,,$$

which is derived from (3. 5) and (3. 6). After all, the given constraints can be written in our notations as

$$B=\left[\begin{array}{c:cc:ccc} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 \end{array}\right] \equiv \begin{bmatrix} \boldsymbol{b}_1' \\ \boldsymbol{b}_2' \\ \boldsymbol{b}_3' \\ \boldsymbol{b}_4' \end{bmatrix}, \tag{3. 8}$$

$$\boldsymbol{\beta}'=(\mu, \alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3)\,.$$

The design matrix X has rank 4 and we can choose a matrix D such that

$$D=\begin{bmatrix} 1 & 1 \\ -1 & 0 \\ -1 & 0 \\ 0 & -1 \\ 0 & -1 \\ 0 & -1 \end{bmatrix}.$$

Thus the equations $\boldsymbol{l}'D=\boldsymbol{0}'$ are

$$\begin{aligned} &2c_1=0\,, \\ &3c_2+c_3+4c_4=0\,, \end{aligned} \tag{3. 9}$$

and the general solution to (3. 9) is given by

$$c_1=0\,, \qquad c_2=k_1\,, \qquad c_3=k_2\,, \qquad c_4=-\frac{3}{4}k_1-\frac{1}{4}k_2\,, \tag{3. 10}$$

where k_1 and k_2 are arbitrary constants. Then, when $\boldsymbol{l}\in\mathscr{R}[X']$, $\boldsymbol{l}$ can be witten as

$$\boldsymbol{l}'=k_3(0,0,0,-5,4,1)$$

$$=k_3(\boldsymbol{b}_2'-3\boldsymbol{b}_3') \tag{3. 11}$$

$$=k_3(-4\boldsymbol{b}_3'+\boldsymbol{b}_4')\,, \tag{3. 12}$$

where k_3 is an arbitrary constant.

If Theorem 3 is applied to (3. 9) or (3. 10), then we have

$$\lambda_1=0 \qquad \text{and} \qquad 3\lambda_2+\lambda_3+4\lambda_4=0\,. \tag{3. 13}$$

If Theorem 1 is applied to (3. 11), then we can see that the normal equations have a solution which satisfies

$$\lambda_1=\lambda_4=0\,, \tag{3. 14}$$

$$\lambda_2\neq0\,, \qquad \lambda_3\neq0 \tag{3. 15}$$

and

$$2\lambda_2+\lambda_3=0\,. \tag{3. 16}$$

(3. 14) and (3. 15) follow from Theorem 1 by noticing that $\boldsymbol{l}\in\mathscr{R}[\boldsymbol{b}_2\,\vdots\,\boldsymbol{b}_3]$ and that $\boldsymbol{b}_2$ and $\boldsymbol{b}_3$ are linearly independent. (3. 16) follows from (3. 13).

On the other hand, if Theorem 1 is applied to (3. 12), then we can see that the normal equations have a solution which satisfies

$$\lambda_1=\lambda_2=0\,,$$

$$\lambda_3\neq0\,, \qquad \lambda_4\neq0$$

and

$$\lambda_3+4\lambda_4=0\,.$$

4. Some Remarks

As mentioned before, if the given constrains consist of identifiability constraints only, then the Lagrange multipliers corresponding to these are equal to zero (or can be taken to be equal to zero). If there exist additional constraints in addition to identifiability constraints, however, then the Lagrange multipliers corresponding to identifiability constraints may not be equal to zero.

As an example, we shall refer Example 3. Suppose that the constraints under consideration consist of the identifiability constraints (3. 5) and an additional constraint (3. 6), and denote the Lagrange multipliers corresponding to these by λ_1, λ_2 and λ_3 respectively. Then we have $\lambda_1=0$, $\lambda_2\neq 0$ and $\lambda_3\neq 0$.

In fact, the values of Lagrange multipliers depend on the expressional form of additional constraints. Denote the identifiability constraints and the additional constraints by

$$H\boldsymbol{\beta}=\mathbf{0} \tag{4. 1}$$

and

$$F\boldsymbol{\beta}=\mathbf{0}\,, \tag{4. 2}$$

respectively. Since $H\boldsymbol{\beta}=\mathbf{0}$ are identifiability constraints, $F=G\begin{bmatrix} X \\ H \end{bmatrix}$ for some G. Thus (4. 2) can be transformed into

$$GX\boldsymbol{\beta}=\mathbf{0}\,. \tag{4. 3}$$

If (4. 3) is given as additional constraints instead of (4. 2) and all the constraints are linearly independent, then we can conclude from Corollary 2 that the Lagrange multipliers corresponding to identifiability constraints are all equal to zero and that each of the Lagrange multipliers corresponding to additional constraints is not equal to zero.

In general, a given constraints $B\boldsymbol{\beta}=\mathbf{0}$ can be transformed into the equivalent constraints $\begin{bmatrix} \tilde{B}_0 \\ \hdashline \tilde{B}_1 \end{bmatrix}\boldsymbol{\beta}=\mathbf{0}$ which satisfy the following properties: i) The rows of $\begin{bmatrix} \tilde{B}_0 \\ \hdashline \tilde{B}_1 \end{bmatrix}$ are linearly independent. ii) No linear combination of the rows of $\tilde{B}_0$ is a linear combination of the rows of X except $\mathbf{0}'$, i.e., $\tilde{B}_0\boldsymbol{\beta}=\mathbf{0}$ are identifiability constraints or a subset of them. iii) Each column of $\tilde{B}_1'$ belongs to $\mathcal{R}[X']$. If $\begin{bmatrix} \tilde{B}_0 \\ \hdashline \tilde{B}_1 \end{bmatrix}\boldsymbol{\beta}=\mathbf{0}$ is adopted as constraints instead of $B\boldsymbol{\beta}=\mathbf{0}$, then the Lagrange multipliers corresponding to $\tilde{B}_0\boldsymbol{\beta}=\mathbf{0}$ are equal to zero and each of the Lagrange multipliers corresponding to $\tilde{B}_1\boldsymbol{\beta}=\mathbf{0}$ is not equal to zero.

References

[1] Mann, H. B. (1949): *Analysis and design of experiments.* Dover Publications, Inc., New York.

[2] Masuyama, M. (1967-1968): *Mathematics of experimental designs.* Lecture note at the Catholic Univ. of America.

[3] Plackett, R. L. (1960): *Principles of regression analysis.* Oxford.

[4] Reiersφl, O. (1963): Identifiability, estimability, phenorestricting specifications, and zero Lagrange multipliers in the analysis of variance. *Skand. Aktuarietidskrift, Häft.* 3-4, 131-142.

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

[6] Yamamoto, S. and Fujikoshi, Y. (1967): The linear hypotheses and constraints. *J. Sci. Hiroshima Univ. Ser.* AI 31, 211-219.