IMPLEMENTATION OF MULTIPLE COMPARISON PROCEDURES IN A GENERALIZED LEAST SQUARES PROGRAM

by

MERVYN G. MARASINGHE
B.Sc., University of Sri Lanka, 1970

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas
1977

Approved by:

K. E. Kemp
Major Professor

TABLE OF CONTENTS

ILLEGIBLE DOCUMENT

THE FOLLOWING DOCUMENT(S) IS OF POOR LEGIBILITY IN THE ORIGINAL

THIS IS THE BEST COPY AVAILABLE

CHAPTER 1

INTRODUCTION

1.1 Statement of the Problem

The LEAST SQUARES ANALYSIS OF VARIANCE Program (LSQRS), which has been in consistent use for analysing designs with unbalanced data, could be more useful to experimenters and researchers if a procedure for multiple separations can be made available. This project has been undertaken to modify LSQRS to provide estimates of the standard errors and LSD's for differences of every pair of means.

There are two basic schools of thought regarding computational techniques to be adopted to analyse General Linear Models not of full rank. The classical way of attacking the problem has been the use of restrictions on the parameters to reparameterize the model to one of full rank, and apply the results of the full rank model. The problem of multiple separations in this case is not easy to handle and the generalization of the computing techniques is complex.

The more modern approach has been to use the results obtained for the General Linear Model of less than full rank directly in the computations. This involves the development of algorithms for the computation of the generalized inverse of matrices and also for conditions of estimability of parameters and testability of hypotheses that are easily computed and used in a program. These techniques are discussed in detail in BENTZ [1]. The method adopted herein is an attempt to use the reparameterized model for obtaining data necessary to use the results available for the model of less than full rank.

1.2 Contents and Goals of This Report

The goal of this report is to develop, program and implement algorithms which carry out multiple separations in the Least Squares program. The program uses restrictions on the parameters to reparameterize a less than full rank model to a full rank model. The proposed algorithm involves the construction of a matrix L, which is used to transform the non-full rank design matrix to the full rank design matrix obtained by the above mentioned restrictions.

The differences of the means, which are actually linear combinations of estimable functions of parameters, and their standard errors can then be estimated in terms of this matrix L and the inverse of the reduced sums of squares and cross-products matrix which is already available. This technique attempts to overcome the need for computing the generalized inverses and still use the results of the linear model of less than full rank.

Computation of estimates of estimable functions of the parameters requires the construction of vectors of constants. The proposed modifications contain routines for developing these vectors for every mean required to be analysed.

Chapter 2 contains a brief statement of the theory of the General Linear Model, in particular, the results which will be used in this report. Chapter 3 contains a description of the Least Squares Program and the reparameterization that is used in the program. The results used in the suggested modifications and how these are implemented in the program will be discussed in Chapter 4 while an example illustrating the computational techniques will be presented in Chapter 5.

CHAPTER 2

THE GENERAL LINEAR MODEL

This chapter consists of a review of the basic theorems and definitions concerning the General Linear Model. Most of the theory will be necessary for stating the problem in mathematical terms and for developing the results used in the computations. The detailed proofs and the required theory of matrices can be found in Graybill [2].

2.1 Notations and Definitions

In the foregoing statement of theorems, uppercase letters such as A,X,U, denote matrices while underlined uppercase letters such as $\underline{Y}$,$\underline{Z}$ denote random vectors. Underlined lowercase letters, such as $\underline{r}$,$\underline{a}$,$\underline{1}$, denote fixed column vectors. The transpose of matrix A is denoted by A'. Lower case letters which are not underlined denote scalars or constants.

The generalized inverse of a matrix A will be denoted by A^- and will be referred to as the g-inverse of A. A conditional inverse of A will be denoted by A^c and will be referred to as a c-inverse of A.

Definition 2.1.1 Graybill [2]

General Linear Model - Let $\underline{Y}$ be an $n \times 1$ observable vector of random variables, X be an $n \times p$ matrix $(n > p)$ of known fixed numbers, $\underline{\beta}$ be a $p \times 1$ vector of unknown parameters and $\underline{\varepsilon}$ be an $n \times 1$ unobservable vector of random variables where $\mathcal{E}[\underline{\varepsilon}] = \underline{0}$ and Cov $[\underline{\varepsilon}] = \Sigma$ where Σ is an $n \times n$ matrix of constants, then the general linear model is defined by

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} .$$

Throughout this report it will be assumed that $\underline{\varepsilon}$ is distributed normally with mean zero and covariance matrix $\sigma^2 I$, where $\sigma^2 > 0$ is an unknown parameter, and I an n x n identity matrix.

2.2 The General Linear Model of the Full Rank

Definition 2.2.1

In the model defined in Definition 2.1.1 if we assume that the rank of X is p, then it is called the General Linear Model of full rank.

Theorem 2.2.1 Graybill [2]

Let $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$ where $\underline{\varepsilon}$ is distributed $N(\underline{0}, \sigma^2 I)$, be as given by Def. 2.1.1. The results below follow:

1) $\underline{\hat{\beta}} = X^{-}\underline{Y}$ is the maximum likelihood estimator of $\underline{\beta}$.

2) $$\hat{\sigma}^2 = \frac{\underline{Y}'(I - X(X'X)^{-1}X')\underline{Y}}{n - p}$$ is the maximum likelihood estimator of σ^2 (adjusted for bias).

3) $\underline{\hat{\beta}}$ is distributed $N(\underline{\beta}, \sigma^2(X'X)^{-1})$.

4) $\frac{(n-p)\hat{\sigma}^2}{\sigma^2} = U$ is distributed $\chi^2(u; n-p)$

5) $\underline{\hat{\beta}}$ and $\hat{\sigma}^2$ are independent.

6) $\underline{\hat{\beta}}$ and $\hat{\sigma}^2$ are sufficient statistics for $\underline{\beta}$ and σ^2.

7) $\underline{\hat{\beta}}$ and $\hat{\sigma}^2$ are complete statistics.

Remark: Johnson [3]

$\underline{\ell}'\underline{\hat{\beta}}$ is distributed $N(\underline{\ell}'\underline{\beta}, \sigma^2\underline{\ell}'(X'X)^{-1}\underline{\ell})$

Since $\underline{\ell}'\underline{\hat{\beta}}$ is the unbiased estimator of $\underline{\ell}'\underline{\beta}$, and since $\underline{\ell}'\underline{\hat{\beta}}$ is a function of the complete sufficient statistics $\underline{\hat{\beta}}$ and $\hat{\sigma}^2$, we have that $\underline{\ell}'\underline{\hat{\beta}}$ is the UMVU estimator of $\underline{\ell}'\underline{\beta}$.

Theorem 2.2.2 Graybill [2]

In the general linear model $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$ where $\underline{\varepsilon}$ is distributed $N(\underline{0}, \sigma^2 I)$, W is a (function of the) generalized likelihood ratio test statistic for testing the hypothesis

$$H_o : H\underline{\beta} = \underline{h} \text{ vs } H_\alpha : H\underline{\beta} \neq \underline{h}$$

where H is a q x p matrix of rank q.

$$W = \frac{(H\hat{\underline{\beta}} - \underline{h})'(H(X'X)^{-1}H')^{-1}(H\hat{\underline{\beta}} - \underline{h})}{q\hat{\sigma}^2}$$

where $\hat{\underline{\beta}} = X^{-}\underline{Y}$ and $\hat{\sigma}^2 = \dfrac{\underline{Y}'(I - XX^{-})\underline{Y}}{n - p}$ are

UMVU estimators of $\underline{\beta}$ and σ^2, respectively.

Another form of W and the distributional properties of W are also included in this theorem, which are omitted here.

2.3 The Linear Model of less than full rank

Definition 2.3.1

$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$, $\underline{\varepsilon}$ is distributed $N(\underline{0}, \sigma^2 I)$

X has size n x p and rank k where $n > p > k$.

The parameter space is $\Omega = \{(\underline{\beta}, \sigma^2) : \underline{\beta} \in E_p, \sigma^2 > 0\}$

This model is referred to as a linear model of less than full rank.

Definition 2.3.2

Estimable Function Consider the design model defined above. A function, say $q(\underline{\beta}, \sigma^2)$, of the parameters $\underline{\beta}$, σ^2 is defined to be an estimable function if and only if there exists an unbiased estimator for $q(\underline{\beta}, \sigma^2)$.

Theorem 2.3.2 Graybill [2]

Consider the design model given in Definition 2.3.1. A specified linear function of $\underline{\beta}$, namely $\underline{\ell}'\underline{\beta}$ where $\underline{\ell}$ is a given p x 1 constant vector, is an estimable function if and only if any of the conditions below are satisfied.

1) $\underline{\ell}$ is a linear combination of the columns of X'

2) $\text{rank}\ (X', \underline{\ell}) = \text{rank}\ (X')$

3) $\text{rank}\ (X'X) = \text{rank}\ (X'X, \underline{\ell})$

4) a solution vector $\underline{r}$ exists for the equations $X'X\underline{r} = \underline{\ell}$

5) $\underline{\ell}'\ X^c X = \underline{\ell}'$ for any c-inverse of X

6) $X'\ (X')^c \underline{\ell} = \underline{\ell}$ for any c-inverse of X'

7) $(X'X)(X'X)^c \underline{\ell} = \underline{\ell}$ for any c-inverse of $X'X$

8) $\underline{\ell}'\ (X'X)^c (X'X) = \underline{\ell}'$ for any c-inverse of $X'X$

Definition 2.3.3 Graybill [2]

Linearly Independent Estimable Functions of $\underline{\beta}$

A set of m linear functions of $\underline{\beta}$, say $\underline{\ell}_1'\underline{\beta}$, $\underline{\ell}_2'\underline{\beta}$, ... $\underline{\ell}_m'\ \underline{\beta}$ is defined to be a set of m linearly independent estimable functions of $\underline{\beta}$ if and only if (1) each $\underline{\ell}_r'\ \underline{\beta}$ is an estimable function; (2) the p x 1 vectors $\underline{\ell}_1$, $\underline{\ell}_2$, ... $\underline{\ell}_m$ are linearly independent (or the rank of L is m where $L = [\underline{\ell}_1, \underline{\ell}_2, \ldots\ldots \underline{\ell}_m]$.

Theorem 2.3.3 Graybill [2].

The number of linearly independent estimable functions of $\underline{\beta}$ is equal to the rank of X, which is k.

Definition 2.3.4. Graybill [2]. Set, Full Set, and Basis Set of Estimable Functions.

Consider the p x m matrix of constants $L = [\underline{\ell}_1, \underline{\ell}_2, \ldots \underline{\ell}_m]$ where $\underline{\ell}_r' \underline{\beta}$ is an estimable function for each r = 1, 2, ..., m. Then $L' \underline{\beta}$ is defined to be

1) A Set of m estimable functions.
2) A Full Set of estimable function if L has rank k.
3) A Basis Set of estimable functions if m = k and L has rank k.

Theorem 2.3.4. Graybill [2].

Consider the design model in Def. 2.3.1. and the normal equations,

$$X'X\hat{\underline{\beta}} = X'\underline{Y}$$

1) If $\underline{\ell}'\underline{\beta}$ is an estimable function, then $\underline{\ell}'\hat{\underline{\beta}}$ is invariant for "any" solution $\hat{\underline{\beta}}$ of the normal equations, and
$$\underline{\ell}'\hat{\underline{\beta}} = \underline{\ell}'X^{-}\underline{Y}$$
2) If $\underline{\ell}'\underline{\beta}$ is an estimable function, then $\underline{\ell}'\hat{\underline{\beta}}$ is the UMVU estimator of $\underline{\ell}'\underline{\beta}$.
3) $\hat{\sigma}^2 = (n - k)^{-1} (\underline{Y}'\underline{Y} - \hat{\underline{\beta}}'X'\underline{Y})$ is invariant for any solution $\hat{\underline{\beta}}$ of the normal equations.
4) $\hat{\sigma}^2$ is the UMVU estimator of σ^2.

5) Every element of $X\underline{\beta}$ is an estimable function.

6) The UMVU estimator of any estimable function $\underline{\ell}'\underline{\beta}$ must be a linear combination of the UMVU estimators of every Basis Set of Estimable Functions (and also of every Full Set of Estimable Functions).

Note: Below are some equivalent expressions for $\hat{\sigma}^2$, where $\hat{\underline{\beta}}$ is any solution to the normal equations:

$$(n - k)^{-1}\hat{\sigma}^2 = \underline{Y}'\underline{Y} - \hat{\underline{\beta}}'X'\underline{Y} = \underline{Y}'(I - X'X^{-})\underline{Y}$$

$$= \underline{Y}'(I - X(X'X)^{c}X')\underline{Y}.$$

where $(X'X)^c$ is any c-inverse of $X'X$.

Theorem 2.3.5. Graybill [2].

Consider the design model in Def. 2.3.1. Let $L'\underline{\beta}$ be any set of estimable functions (where L' is $q \times p$ of rank m) and let $\hat{\underline{\beta}}$ denote any solution to the normal equations $X'X\hat{\underline{\beta}} = X'\underline{Y}$

1) $L'\hat{\underline{\beta}}$ is distributed as the q - variate normal distribution of rank m with mean $L'\underline{\beta}$ and covariance matrix $\sigma^2 L'(X'X)^c L$ where $(X'X)^c$ is any c-inverse of $(X'X)$.

2) $U = (n - k)^{-1}\hat{\sigma}^2/\sigma^2$ is distributed $\chi^2(n - k)$.

3) The random vector $L'\hat{\underline{\beta}}$ is independent of U.

2.4. Reparameterization.

In section 2.2, the results for the full rank model were presented while in section 2.3 the methods available for estimation of functions of parameters when the rank of X is less than p was discussed. Although, in the non-full rank case the estimators for $L'\underline{\beta}$ and σ^2 are unique for any solution of the normal equations, it is more difficult to standardise the methods for solution, so that a computer program can be used. At least a c-inverse of $X'X$ (or the g-inverse of $X'X$) has to be found.

If the matrix X, whose rank is less than the number of its columns can be transformed to a matrix U whose rank is equal to the number of its columns, and if we can obtain a transformed model involving U, which conforms to Def. 2.2.1., then all the results in Section 2.2 could be applied directly to this model.

The procedure will be to transform the $\underline{\beta}$ vector and the X matrix to a new vector $\underline{\theta}$ and a new matrix U of size n x k such that $X\underline{\beta} = U\underline{\theta}$. If U is n x k of rank k, then the transformed model $\underline{Y} = U\underline{\theta} + \underline{\varepsilon}$ satisfies all conditions of the general linear model of full rank. This transformation will be accomplished by transforming the parameter $\underline{\beta}$ to a new parameter $\underline{\theta}$ by a k x p matrix L' or rank k where the rows of L' form linearly independent estimable functions. The transformation $\underline{\theta} = L'\underline{\beta}$ of the parameter $\underline{\beta}$ to the parameter $\underline{\theta}$ is called <u>reparameterization.</u>

<u>Definition 2.4.1.</u> Graybill [2]. <u>Transformation and Reparameterization.</u>

Consider the design model $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$ in Def. 2.3.1., where X is n x p of rank k. Let L be any p x m matrix and let $\underline{\theta} = L'\underline{\beta}$. Denote the ith column of L by $\underline{\ell}_i$ so that $L = [\underline{\ell}_1, \underline{\ell}_2, \ldots\ldots, \underline{\ell}_m]$. Then

1) $L'\underline{\beta}$ is defined to be a "transformation" of the vector $\underline{\beta}$ to vector $\underline{\theta}$;

2) $L'\underline{\beta}$ is defined to be an "estimable transformation" of the vector $\underline{\beta}$ to the vector $\underline{\theta}$ if and only if each θ_i (each $\underline{\ell}_i' \underline{\beta}$) is estimable for $i = 1,2,\ldots\ldots, m$;

3) $L'\underline{\beta}$ is defined to be a "reparameterization" of the vector $\underline{\beta}$ to the vector $\underline{\theta}$ if and only if each θ_i (each $\underline{\ell}_i' \underline{\beta}$) is estimable for $i = 1,2,\ldots\ldots\ldots, m$ where L has rank k and $k = m$ i.e. $\underline{\theta}$ is a Basis Set of estimable functions of $\underline{\beta}$.

Let L' be a $k \times p$ matrix of rank k, such that $\underline{\theta} = L'\underline{\beta}$ is estimable. Then it can be shown that (Johnson [3]),

$$Y = X\underline{\beta} + \underline{\varepsilon} \text{ if and only if}$$

$$\begin{aligned} Y &= (XL'^{c})\,(L'\underline{\beta}) + \underline{\varepsilon} \\ &= U\underline{\theta} + \underline{\varepsilon} \quad \text{where } U = XL'^{c} \text{ is } n \times k \text{ and rank } (U) = k \\ &\qquad \text{and } \underline{\theta} = L'\underline{\beta} \end{aligned}$$

Thus the less than full rank model, $Y = X\underline{\beta} + \underline{\varepsilon}$ has been reparameterized to a full rank model $Y = U\underline{\theta} + \underline{\varepsilon}$. All the theorems in section 2.2 hold for this model. In particular $\hat{\underline{\theta}}$ and $\hat{\sigma}^2 = \dfrac{Y'(I - UU^-)\underline{Y}}{n - k}$ are complete sufficient statistics.

<u>Theorem 2.4.1</u> Johnson [3].

Let $Y = U\underline{\theta} + \underline{\varepsilon}$ be any reparameterization of $Y = X\underline{\beta} + \underline{\varepsilon}$. Then $\hat{\underline{\theta}} = L'\hat{\underline{\beta}}$ where $\hat{\underline{\beta}}$ is any solution to the normal equations $X'X\hat{\underline{\beta}} = X'\underline{Y}$.

Two theorems are included here which may be found useful later. One concerns the test of the hypothesis $H_0 : H\underline{\beta} = 0$ vs $H_a : H\underline{\beta} \neq 0$ where

$H\underline{\beta}$ is a set of linearly independent functions (referred to as a testable hypothesis) and the other, the distribution of $\underline{\ell}'\hat{\underline{\beta}}$.

Theorem 2.4.2. Graybill [2].

In the design model in Def. 2.3.1 let $H\underline{\beta}$ (H known) be a set of q independent estimable functions of $\underline{\beta}$. W is the generalized likelihood ratio test statistic for $H_0 : H\underline{\beta} = \underline{0}$ vs $H_a : H\underline{\beta} \neq \underline{0}$ where

$$W = \frac{(H\hat{\underline{\beta}})' \; (H \; (X'X)^c H')^{-1} (H\hat{\underline{\beta}})}{q \; \hat{\sigma}^2}$$

where $\hat{\underline{\beta}}$ is any solution of the normal equations $X'X\hat{\underline{\beta}} = X'\underline{Y}$ and $\hat{\sigma}^2 = (n - k)^{-1} \; (\underline{Y}'\underline{Y} - \hat{\underline{\beta}}'X'\underline{Y})$.

Theorem 2.4.3. Johnson [3].

Consider the design model given in Def. 2.3.1. Suppose $\underline{\ell}'\underline{\beta}$ is an estimable function. Then $\underline{\ell}'\hat{\underline{\beta}} \sim N(\underline{\ell}'\underline{\beta}, \; \sigma^2\underline{\ell}'(X'X)^c\underline{\ell} \;)$.

CHAPTER 3

THE LEAST SQUARES PROGRAM

Since most design matrices are of less than full rank, the ordinary least squares procedures cannot be used directly to analyse the models. In order to reduce the non-full rank model to a full rank model, the least squares program utilizes the technique of imposing restrictions on the parameters of the original model. This procedure as well as a method of obtaining the restricted normal equations by means of a reparameterization as introduced in Section 2.4., will be discussed in this chapter.

The techniques are illustrated using a two-way model with interaction. Extension to other models follow the same pattern.

3.1. General Description.

In the case of the two-way model with interaction and n_{ij} observation per cell,

$$Y_{ijk} = \mu + \alpha_i + \tau_j + \gamma_{ij} + e_{ijk}$$

where α_i = parameter for i^{th} level of treatment A, $i = 1,\ldots\ldots,a$

τ_j = parameter for j^{th} level of treatment B, $j = 1,\ldots\ldots,b$

γ_{ij} = parameter for interaction effect of i^{th} level of A and j^{th} level of B

Y_{ijk} = the k^{th} observation of the ij^{th} treatment combination.

e_{ijk} = the random errors, assumed to be $NID(0, \sigma^2)$.

With the usual notation, this model could be written in the form of the General Linear Model,

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$$

where $\underline{Y}$ = vector of observations (Y_{ij})

X = design matrix of zeros and ones

$\underline{\beta}$ = the vector of parameters where

$$\underline{\beta}' = (\mu\ \alpha_1\ \alpha_2 \cdots\cdots\ \alpha_a\ \tau_1\ \tau_2 \cdots\cdots\ \tau_b\ \gamma_{11}\ \gamma_{12} \cdots\cdots\ \gamma_{ab})$$

$\underline{\varepsilon}$ = the error vector ; $\underline{\varepsilon} \sim N(\underline{0}, \Sigma)$

where $\Sigma = \sigma^2 I$

The design matrix in this case will not be of full column rank and thus the normal equations given by

$$X'X\hat{\underline{\beta}} = X'\underline{Y}$$

cannot be used to obtain a unique estimate of $\underline{\beta}$ - vector, by applying the results of the GLM of full rank. Thus the results of the GLM of less than full rank must be used and the estimates of $\underline{\beta}$, σ^2 and the likelihood ratios used for hypothesis testing involve the use of the generalized inverse of the $X'X$ matrix (or the c-inverse of the $X'X$ matrix).

3.2. Estimation of Parameters

In the LSQRS program a set of restrictions on the parameters are used to reduce the $X'X$ matrix to a matrix of full rank, thus enabling the use of the results obtained for a full rank model.

The procedure used is to set a number of non-estimable functions of $\underline{\beta}$ to zero. If the rank of the X'X matrix of size p x p is k, then the number of such non-estimable conditions needed are given by p-k. In the documentation for the LSQRS program (KEMP [4]), the non-estimable conditions used are described as restrictions on the parameters. To clarify the foregoing the two-way model described in Section 3.1 is used here. In this case $\alpha_i (i = 1,......, a)$, $\tau_j (j = 1,......, b)$ and $\gamma_{ij} (i = 1,......, a; j = 1,......, b)$ are all non-estimable functions of $\underline{\beta}$. In addition, $\alpha_. = \sum_{i=1}^{a} \alpha_i$, $\tau_. = \sum_{j=1}^{b} \tau_j$, $\gamma_{.j} = \sum_{i=1}^{a} \gamma_{ij}$ $(j = 1,.....b)$, $\gamma_{i.} = \sum_{j=1}^{b} \gamma_{ij}$ $(i = 1,.....,a)$ are non-estimable. In LSQRS, the non-estimable conditions used as restrictions are these; i.e.

$$\alpha_. = 0$$

$$\tau_. = 0$$

$$\gamma_{i.} = 0 \quad ; i = 1,......,a$$

$$\gamma_{.j} = 0 \quad ; j = 1,......,b$$

It has to be pointed out that the sums of squares for testing the hypotheses $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_a$ and $H_0 : \tau_1 = \tau_2 = ... = \tau_b$ can be obtained by using the R( ) - notation, from the restricted model by employing the above restrictions (SPEED and HOCKING [5]). The method of obtaining these sums of squares is described in KEMP [4].

Using these restrictions, the original X'X matrix, which is less than full rank is converted to a full rank matrix by deleting the columns corresponding to the last class of each main effect and adjusting the columns corresponding to the other classes. The elements of columns due to interaction effects in the adjusted matrix are obtained by the products of the appropriate elements of the adjusted main effect columns. This procedure is treated completely in KEMP [4].

To exemplify the foregoing, a simple model is used.

Let $Y_{ijk} = \mu + \alpha_i + \tau_j + \gamma_{ij} + e_{ijk}$ $\qquad i = 1,2$; $j = 1,2,3$; $k = 1,2$

be the model under consideration. Each cell corresponding to each ij combination contains 2 observations.

Then the model could be written as

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$$

where

$$\underline{Y}' = (Y_{111} \; Y_{112} \; Y_{121} \; Y_{122} \; Y_{131} \; Y_{132} \; Y_{211} \; Y_{212} \; Y_{221} \; Y_{222} \; Y_{231} \; Y_{232})$$

$$\underline{\beta}' = (\mu \;\; \alpha_1 \;\; \alpha_2 \;\; \tau_1 \;\; \tau_2 \;\; \tau_3 \;\; \gamma_{11} \;\; \gamma_{12} \;\; \gamma_{13} \;\; \gamma_{21} \;\; \gamma_{22} \;\; \gamma_{23}) \quad 12 \times 1$$

and

$$X = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad 12 \times 12$$

Note that X'X will be of rank 6 and thus 6 restrictions on the parameters are needed to transform X'X to a full rank matrix. If the restrictions $\alpha_{.} = \tau_{.} = \gamma_{i.} = \gamma_{.j} = 0$ are imposed and the matrix X is adjusted using these restrictions, then the matrix U is obtained where

$$U = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix}_{12 \times 6}$$

and the corresponding vector of parameters $\underline{\theta}$ is

$$\underline{\theta}' = (\mu^* \; \alpha^*_1 \; \tau^*_1 \; \tau^*_2 \; \gamma_{11}^* \; \gamma_{12}^*).$$

Now U'U can be computed which is of rank 6 and thus the normal equations $U'U\hat{\underline{\theta}} = U'\underline{Y}$ can be solved to obtain the UMVU estimate of $\underline{\theta}$.

3.3 A reparameterization to obtain a model equivalent to the model obtained by the imposing of above restrictions.

It is observed that by post-multiplying X by a 12 x 6 matrix of full column rank, U may be obtained directly. The matrix used for this purpose is denoted by P.

$$U = \underset{12 \times 12}{\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}} \underset{12 \times 6}{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}} = XP \text{ (say)}$$

That is U = XP where P is given by the 12 x 6 matrix shown above. Thus, the procedure described above can be looked upon as a reparameterization of the model $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$ to the model $\underline{Y} = U\underline{\theta} + \underline{\varepsilon}$, as introduced in Section 2.4.

Consider the matrix L',

$$L' = \begin{bmatrix} 1 & 1/2 & 1/2 & 1/3 & 1/3 & 1/3 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 1/2 & -1/2 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & -1/6 & -1/6 & -1/6 \\ 0 & 0 & 0 & 2/3 & -1/3 & -1/3 & 1/3 & -1/6 & -1/6 & 1/3 & -1/6 & -1/6 \\ 0 & 0 & 0 & -1/3 & 2/3 & -1/3 & -1/6 & 1/3 & -1/6 & -1/6 & 1/3 & -1/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & -1/6 & -1/6 & -1/3 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1/6 & 1/3 & -1/6 & 1/6 & -1/3 & 1/6 \end{bmatrix}_{6 \times 12}$$

This matrix has been arrived at by constructing rows of L' such that $L'\underline{\beta} = \underline{\theta}$ where $\underline{\theta}$ is a set of estimable linear combinations of the parameters. If the $\underline{\beta}$ vector is pre-multiplied by L' we obtain a vector $\underline{\theta}$ where

$$\underline{\theta}_{6 \times 1} = \begin{bmatrix} \mu + \bar{\alpha}_{\cdot} + \tau_{\cdot} + \bar{\gamma}_{\cdot\cdot} \\ \alpha_1 - \bar{\alpha}_{\cdot} + \bar{\gamma}_{1\cdot} - \bar{\gamma}_{\cdot\cdot} \\ \tau_1 - \bar{\tau}_{\cdot} + \bar{\gamma}_{\cdot 1} - \bar{\gamma}_{\cdot\cdot} \\ \tau_2 - \bar{\tau}_{\cdot} + \bar{\gamma}_{\cdot 2} - \bar{\gamma}_{\cdot\cdot} \\ \gamma_{11} - \bar{\gamma}_{1\cdot} - \bar{\gamma}_{\cdot 1} + \bar{\gamma}_{\cdot\cdot} \\ \gamma_{12} - \bar{\gamma}_{1\cdot} - \bar{\gamma}_{\cdot 2} + \bar{\gamma}_{\cdot\cdot} \end{bmatrix}$$

Let elements of $\underline{\theta}$ be denoted by $\underline{\theta}' = (\mu \;\; \alpha_1^* \;\; \tau_1^* \;\; \tau_2^* \;\; \gamma_{11}^* \;\; \gamma_{12}^*)$. By the theory of the two-way design model with interaction, it can be shown that $\mu^*, \alpha_1^*, \tau_1^*, \tau_2^*, \gamma_{11}^*, \gamma_{12}^*$ are all estimable functions of the parameters of the model given above, namely $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$.

Now consider the matrix product L'PL'. This is shown to be equal to L', which indicates that actually P is a c-inverse of the matrix L'. Therefore let $L'^{c} = P$.

So using the results of Section 2.4, $L'\underline{\beta} = \underline{\theta}$ is a reparameterization of the model $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$ to the model $\underline{Y} = U\underline{\theta} + \underline{\varepsilon}$ where L' is as given above. Since $\hat{\underline{\theta}}$ is unique, $\hat{\underline{\theta}} = (U'U)^{-1}\, U'\underline{Y}$, it follows that whatever L' chosen such that it satisfies the above conditions, will give the same $\hat{\underline{\theta}}$. Thus imposing the above restrictions on the parameters is equivalent to reparameterization using the L' matrix indicated. It is observed that the matrix L'^{c} is a matrix possessing a certain pattern; in this case

$$L'^{c} = \left[\begin{array}{ccc|c} 1 & 0 & & \\ 0 & \begin{bmatrix} I_1 \\ -1 \end{bmatrix}_{2\times 1} & 0_{3\times 2} & \\ 0 & & & 0_{6\times 2} \\ & 0_{3\times 2} & \begin{bmatrix} I_2 \\ -\underline{j}' \end{bmatrix}_{3\times 2} & \\ \hline & 0_{6\times 4} & & \begin{bmatrix} I_2 \\ -\underline{j}' \end{bmatrix}_{3\times 2} \otimes \begin{bmatrix} I_1 \\ -1 \end{bmatrix}_{2\times 1} \end{array}\right]$$

Thus L'^{c} is a matrix that can be constructed for a particular design given the levels of each effect. The L'^{c} that is constructed this way, will be a c-inverse of a matrix L' which transforms the $\underline{\beta}$-vector to a vector of estimable functions of the parameters. Since $(U'U)^{-1}$ and $\hat{\underline{\theta}}$ are already computed in the LSQRS program, the results of Section 2.4, used in conjuction with L'^{c}, $(U'U)^{-1}$ and $\hat{\underline{\theta}}$ may be used for computation of standard errors of linear combinations of $\hat{\underline{\beta}}$.

CHAPTER 4

THE MODIFICATIONS TO THE LEAST SQUARES PROGRAM

This chapter describes the computations that are to be carried out in the mean separation routine. Sec. 4.1 contains the derivation of the results that will be used for the calculation required in the program. Sec. 4.2 gives a brief outline of the organization of routines within the program which performs these computations.

4.1. Results used in the computations.

In Section 2.4, it is shown that $\underline{\theta} = L'\underline{\beta}$ is a reparameterization of the model

$$Y = X\underline{\beta} + \underline{\varepsilon} \text{ where } X \text{ is } n \times p \text{ and rank } (X) = k \text{ to}$$

$$Y = U\underline{\theta} + \underline{\varepsilon} \text{ where } U = XL'^{c} \text{ is } n \times k \text{ of rank } k.$$

From the theorems in Section 2.2 the UMVU estimator of $\underline{\theta}$ is given by $\hat{\underline{\theta}} = U^{-}\underline{Y} = (U'U)^{-1}U'\underline{Y}$. From the theorems in Section 2.4, $\hat{\underline{\theta}} = L'\hat{\underline{\beta}}$ where $\hat{\underline{\beta}}$ is any solution to the normal equations $X'X\hat{\underline{\beta}} = X'\underline{Y}$.

4.1.1. Estimation of $\underline{\ell}'\underline{\beta}$.

Let $\underline{\ell}'\underline{\beta}$ be any estimable function of the parameters of the model $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$. Then from the results in Section 2.2, $\underline{\ell}'\hat{\underline{\beta}}$ is the UMVU estimator of $\underline{\ell}'\underline{\beta}$.

Now consider $\underline{\ell}'\, L'^{c}\, \hat{\underline{\theta}}$ where $\hat{\underline{\theta}}$ and L'^{c} are as defined above

$$\begin{aligned} \underline{\ell}'\, L'^{c}\, \hat{\underline{\theta}} &= \underline{\ell}'\, L'^{c}\, U^{-}\underline{Y} \\ &= \underline{a}'\, XL'^{c}\, U^{-}\underline{Y} \end{aligned}$$

Since $\underline{\ell}'\underline{\beta}$ is estimable, there exists a vector $\underline{a}$ such that $\underline{a}'\,X = \underline{\ell}'$.

$$\therefore\ \underline{\ell}'\,L'^{c}\underline{\hat{\theta}} = \underline{a}'\,UU^{-}\underline{Y}$$
$$= \underline{a}'\,XX^{-}\underline{Y}$$

since it can be proved that $UU^{-} = XX^{-}$ as shown below;

X is a n x p matrix of rank k

L' is a k x p matrix of rank k

U is a n x k matrix of rank k where $U = XL'^{c}$

Since L' is a k x p matrix of rank k, it follows that $L'L'^{-} = I_k$ (By a theorem concerning generalized inverses.)

Therefore $UU^{-} = UL'L'^{-}U^{-}$

Since U is n x k of rank k and L' is k x p of rank k,

$(UL')^{-} = L'^{-}U^{-}$ and thus

$$UU^{-} = UL'(UL')^{-}$$
$$= XX^{-}$$

Thus it follows that

$$\underline{\ell}'\,L'^{c}\,\underline{\hat{\theta}} = \underline{a}'\,XX^{-}\underline{Y}$$
$$= \underline{\ell}'\,X^{-}\underline{Y} = \underline{\ell}'\underline{\hat{\beta}}$$

So the UMVU estimator of $\underline{\ell}'\underline{\beta}$ is equal to $\underline{\ell}'L'^{c}\underline{\hat{\theta}}$ where $\underline{\hat{\theta}}$ is the solution to the normal equations $U'U\underline{\hat{\theta}} = U\underline{Y}$.

4.1.2. Estimation of the variance of $\underline{\ell}'\underline{\hat{\beta}}$

From the theorems in Section 2.4 it follows that the variance of $\underline{\ell}'\underline{\hat{\beta}}$ can be estimated by $\hat{\sigma}^2\underline{\ell}'(X'X)^{c}\underline{\ell}$ (or by $\hat{\sigma}^2\underline{\ell}'(X'X)^{-}\underline{\ell}$ since $\underline{\ell}'(X'X)^{c}\underline{\ell}$ is invariant to the choice of the pseudo-inverse) where $\hat{\sigma}^2$ is the UMVU estimator of σ^2 which is given by

$$\hat{\sigma}^2 = \frac{\underline{Y}'(I - XX^-)\underline{Y}}{n - k} = \frac{\underline{Y}'(I - X(X'X)^-X')\underline{Y}}{n - k}$$

or by

$$\hat{\sigma}^2 = \frac{\underline{Y}'(I - UU^-)\underline{Y}}{n - k} = \frac{\underline{Y}'(I - U(U'U)^{-1}U')\underline{Y}}{n - k}$$

In LSQRS $\hat{\sigma}^2$ is estimated by

$$= \frac{\underline{Y}'\underline{Y} - \underline{\hat{\theta}}'U'\underline{Y}}{n - k}$$

which is equivalent to above statements.

Now consider

$$\underline{\ell}'L'^c\ (U'U)^{-1}\ (L'^c)'\ \underline{\ell} \quad \text{where } \underline{\ell}'\underline{\beta} \text{ is estimable as before.}$$

$$= \underline{\ell}'\ L'^c\ (U'U)^-\ (L'^c)'\ \underline{\ell}$$

$$= \underline{\ell}'\ L'^c\ U^-\ U'^-\ (L'^c)'\ \underline{\ell}$$

$$= \underline{a}'\ X\ L'^c\ U^-\ U'^-\ (L'^c)'\ X'\ \underline{a} \quad \text{since } \underline{\ell}'\underline{\beta} \text{ is estimable}$$

$$= \underline{a}'\ U\ U^-\ U'^-\ U'\ \underline{a} = \underline{a}'\ U\ U^-\ U\ U^-\ \underline{a}$$

$$= \underline{a}'\ U\ U^-\ \underline{a} = \underline{a}'\ X\ X^-\ \underline{a}$$

$$= \underline{a}'\ X(X'X)^-\ X'\underline{a}$$

$$= \underline{\ell}'\ (X'X)^-\ \underline{\ell}$$

Thus the variance of $\underline{\ell}'\underline{\beta}$ can be estimated by $\hat{\sigma}^2\ \underline{\ell}'\ L'^c(U'U)^{-1}(L'^c)'\underline{\ell}$ where $\hat{\sigma}^2$ is known. Thus it can be seen that only the matrix L'^c is required additionally to compute the estimate of a given estimable function of the $\underline{\beta}$ vector and its variance. Ofcourse, the appropriate $\underline{\ell}$-vector for the estimable function of $\underline{\beta}$ required has to be built within the program.

4.1.3. Comparison of Means.

This procedure is illustrated here using the model introduced in Section 3.2 i.e.

$$Y_{ijk} = \mu + \alpha_i + \tau_j + \gamma_{ij} + \varepsilon_{ijk} \qquad \begin{array}{l} i = 1,2 \\ j = 1,2,3 \\ k = 1,2 \end{array}$$

It can be derived that the best unbiased estimator of

$$\alpha_i - \bar{\alpha}_{.} + \bar{\gamma}_{i.} - \bar{\gamma}_{..} \quad \text{which is estimable}$$

is $$\bar{Z}_{i.} - \bar{Z}_{..} \quad \text{where}$$

$$Z_{ij} = \bar{Y}_{ij.} \quad Z_{i.} = \sum_{j=1}^{3} Z_{ij} \quad \bar{Z}_{i.} = 1/3\ Z_{i.}$$

$$Z_{..} = \sum_{i=1}^{2} Z_{i.}$$

Similarly $\alpha_{i'} - \bar{\alpha}_{.} + \bar{\gamma}_{i'.} - \bar{\gamma}_{..}$ is estimated by

$$\bar{Z}_{i'.} - \bar{Z}_{..}$$

Therefore $\bar{Z}_{i.} - \bar{Z}_{i'.}$ estimates $\alpha_i - \alpha_{i'} + \bar{\gamma}_{i.} - \bar{\gamma}_{i'.}$

Since $\underline{\ell}'\underline{\beta}$ is estimated by $\underline{\ell}'\hat{\underline{\beta}}$, $\bar{Z}_{i.} - \bar{Z}_{i'.} = \underline{\ell}'\hat{\underline{\beta}}$ where $\underline{\ell}'$ is chosen depending on, i, and i',

For instance $\bar{Z}_{1.} - \bar{Z}_{2.} = \underline{\ell}_1'\hat{\underline{\beta}}$ where

$$\underline{\ell}_1' = \underset{1 \times 12}{(0\ 1\ -1\ 0\ 0\ 0\ 1/3\ 1/3\ 1/3\ -1/3\ -1/3\ -1/3)}$$

Similarly $\bar{Z}_{.j} - \bar{Z}_{.j'}$ estimates $\tau_j - \tau_{j'} + \bar{\gamma}_{.j} - \bar{\gamma}_{.j'}$

For example, $\bar{Z}_{.1} - \bar{Z}_{.3} = \underline{\ell}_2' \, \underline{\hat{\beta}}$ where

$$\underline{\ell}_2' = (0 \;\; 0 \;\; 0 \;\; 1 \;\; 0 \;\; -1 \;\; \tfrac{1}{2} \;\; 0 \;\; -\tfrac{1}{2} \;\; \tfrac{1}{2} \;\; 0 \;\; -\tfrac{1}{2})_{1 \times 12}$$

Also $Z_{ij} - Z_{i'j'}$ estimates

$$\alpha_i - \alpha_{i'} + \tau_j - \tau_{j'} + \gamma_{ij} - \gamma_{i'j'}$$

For instance, $Z_{11} - Z_{12} = \underline{\ell}_3' \, \underline{\hat{\beta}}$ where

$$\underline{\ell}_3' = (\; 0 \;\; 0 \;\; 0 \;\; 1 \;\; -1 \;\; 0 \;\; 1 \;\; -1 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;)_{1 \times 12}$$

and $Z_{11} - Z_{23} = \underline{\ell}_4' \, \underline{\hat{\beta}}$ where

$$\underline{\ell}_4' = (\; 0 \;\; 1 \;\; -1 \;\; 1 \;\; 0 \;\; -1 \;\; 1 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; -1 \;)_{1 \times 12}$$

4.2. Description of the Program

The proposed modifications to LSQRS, carryout 3 main functions. These are to construct the L'^c matrix internally, to construct the appropriate $\underline{\ell}$-vectors required for each mean and lastly to carryout the final computations required. As shown in Section 3.3, L'^c is a patterned matrix and given the number of levels of each effect, it can be constructed easily. The only complexity arises when constructing the portion of the L'^c matrix which corresponds to interactions. This involves a routine to form the direct product of up to 3 matrices. The portions of L'^c corresponding to the main effects are constructed in subroutine MAINF and those corresponding to interactions in subroutine INTF. The subroutines

VECTOR and INTVEC are the routines which construct the $\underline{\ell}$-vectors required for estimating the differences of the means. Subroutine LSDCAL completes the analysis by carrying out the calculations given in Section 4.1, to obtain the required estimates and the LSD's.

All the subroutines required to construct the L'^c matrix and the $\underline{\ell}$-vectors for a given model (MAINF, INTF, VECTOR, INTVEC) are activated by a driver subroutine called SEPAR. This routine is called from subroutine MEANSE in LSQRS if the parameter card activating means separation procedures is present. SEPAR calls the appropriate routines depending on the model and causes the L'^c to be constructed. SEPAR also calls in the routines which construct all the $\underline{\ell}$-vectors for all possible differences of means and writes on disk with appropriate identifiers.

Subroutine MEANSE later calls in LSDCAL to calculate the estimates of differences of means, their standard errors and the LSD's using L'^c and the $\underline{\ell}$-vectors and output LSD tables, for those effects which are significant in the AOV at the given level.

The LSQRS program supplies the details regarding the model to the above routines through a common block named STUFF2 which consists of arrays containing information such as the depth of each effect (i.e. main effect, 2-way interaction or 3-way interaction), the name of each effect (i.e. the number the effect is identified with, in the model parameter) the number of levels of each effect and the address of each effect. These are computed in subroutine MEANSE and are in appropriate form to be easily accessible to the routines mentioned above.

4.3. Testing of Hypotheses using the Reparameterization

As a further development, the c-inverse of the reparameterization matrix L which is constructed for the purpose of obtaining the estimates of $\underline{\ell}'\underline{\beta}$ and it's variance where $\underline{\ell}'\underline{\beta}$ is estimable, can be used to obtain the test statistic for testing a given 'testable' hypothesis on the parameters.

Suppose the following hypothesis has to be tested.

$$H_o : H\underline{\beta} = \underline{0} \quad \text{vs.} \quad H_a : H\underline{\beta} \neq \underline{0} ,$$

where H is q x p and of rank q. The program has to ensure first that this is a testable hypothesis and then use Theorem 2.4.2 to obtain the test statistic or the sum of squares given by the numerator of the test statistic divided by the rank of H. The hypothesis is testable if and only if each row of $H\underline{\beta}$ is estimable and thus the condition of 'testability' can be stated as:

$H_o : H\underline{\beta} = \underline{0}$ vs. $H_a : H\underline{\beta} \neq \underline{0}$ is 'testable' if and only if $H\,(X'X)^c(X'X) = H$.

$H\,(X'X)^c(X'X) = H$ if and only if

$H\,(I - (X'X)^cX'X) = 0$ if and only if

$tr\,[H\,(I - (X'X)^cX'X)H'] = 0$ where $tr\,(A)$ indicates the trace of the matrix A.

Now using the notations and results used in Section 4.1, $U = XL'^c$ and $UL' = X$,

$$X'X = LU'UL' .$$

Now let the c-inverse of $X'X$ be given by $L'^c(U'U)^{-1}L^c$. If this relation holds $X'X(X'X)^cX'X$ should be equal to $X'X$. To verify this fact, substituting for $X'X$ and $(X'X)^c$,

$$X'X(X'X)^{c}X'X = LU'UL'L'^{c}(U'U)^{-1}L^{c}LU'UL'$$

$$= LL^{c}LU'UL' \text{ since } L'L'^{c} = I \text{ as shown below.}$$

$$= LU'UL'$$

$$= X'X.$$

Thus, the c-inverse of $X'X$ is given by $L'^{c}(U'U)^{-1}L^{c}$.

Therefore,

$$(X'X)^{c}X'X = L'^{c}(U'U)^{-1}L^{c}X'X$$

$$= L'^{c}(U'U)^{-1}L^{c}LU'UL' \ .$$

Now, $(L'^{c})' = L^{c}$ since

$$(L\ (L'^{c})'\ L)' = L'L'^{c}L'$$

$$= L'$$

implying that $L(L'^{c})'L = L$.

Thus, $L^{c} = (L'^{c})'$.

Therefore, $(L^{c}L)' = L'L'^{c} = I$,

implying that $L^{c}L = I$.

Therefore, $(X'X)^{c}X'X = L'^{c}L'$

Thus the condition for testability reduces to

$$\operatorname{tr}\ H(I - L'^{c}L')H' = 0.$$

To evaluate this L must be derived.

Notice that $(L'^{c})^{c} = L'$ since

$$L'^{c}(L'^{c})^{c}L'^{c} = L'^{c}L'L'^{c}$$

$$= L'^{c} \ .$$

Since $A^c = (A'A)^{-1}A'$ where A is of full column rank,

$$(L'^c)^c = [(L'^c)' \, L'^c]^{-1} \, (L'^c)',$$

$$L' = (L^c \, L'^c)^{-1} L^c .$$

$$\text{Therefore,} L'^c \, L' = L'^c \, (L^c \, L'^c)^{-1} \, L^c .$$

Thus the condition for testability can be obtained in terms of H and the L'^c matrix which is available. Although this involves the evaluation of $(L^c \, L'^c)^{-1}$, the condition of testability ultimately reduces to checking whether the trace of $[H(I - L'^c(L^cL'^c)^{-1}L^c)H']$ is equal to zero or not.

Once testability of the hypothesis is established, the hypothesis can be tested using Theorem 2.4.2. It might also be verified that rank of H is actually q, by deriving the rank of H by some technique.

The sum of squares due to the hypothesis H_o is given by SS_{H_o}, where

$$SS_{H_o} = \frac{(H\hat{\underline{\beta}})' \, [H(X'X)^c \, H']^{-1} \, (H\hat{\underline{\beta}})}{q}$$

$$= \frac{(HL'^c \, \hat{\underline{\theta}})' \, [HL'^c(U'U)^{-1} \, L^cH']^{-1} \, HL'^c \, \hat{\underline{\theta}}}{q} .$$

This result can be derived easily using similar methods indicated in Section 4.1.

CHAPTER 5

EXAMPLE ILLUSTRATING THE TECHNIQUES

This chapter provides an example illustrating the internal computations of the program. The example used is from Winer (6) and is a 2-way design with interaction. Factor A represents two levels of calibrating a dial and the four levels of B are background illumination. The response variable (Y) is the accuracy score obtained from a series of readings.

The data:

	B_1	B_2	B_3	B_4
A_1	3,4,6,7	5,6,6,7,7	4,6,8,8	8,10,10,7,11
A_2	2,3,4	3,5,6,3	9,12,12,8	9,7,12,11

The model:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \qquad \begin{array}{l} i = 1,2 \\ j = 1,2,3,4 \end{array}$$

with the usual notation.

The design matrix

$$X = \begin{pmatrix}
1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix} \qquad Y = \begin{pmatrix} 3 \\ 4 \\ 6 \\ 7 \\ 5 \\ 6 \\ 6 \\ 7 \\ 7 \end{pmatrix}$$

$$
\begin{pmatrix}
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\qquad
\begin{pmatrix}
4 \\ 6 \\ 8 \\ 8 \\ 8 \\ 10 \\ 10 \\ 7 \\ 11 \\ 2 \\ 3 \\ 4 \\ 3 \\ 5 \\ 6 \\ 3 \\ 9 \\ 12 \\ 12 \\ 8 \\ 9 \\ 7 \\ 12 \\ 11
\end{pmatrix}
$$

33 x 15

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

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 & 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 & 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 & 1 & 0 & 0 & -1 \\ 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 \end{bmatrix} \quad 33 \times 8$$

U'U matrix, U'$\underline{Y}$ vector and $(U'U)^{-1}$ are:

$$(U'U) = \begin{bmatrix} 33 & 3 & -2 & 0 & -1 & 0 & 0 & -1 \\ 3 & 33 & 0 & 0 & -1 & -2 & 0 & -1 \\ -2 & 0 & 16 & 9 & 9 & 2 & 1 & 1 \\ 0 & 0 & 9 & 18 & 9 & 1 & 2 & 1 \\ -1 & -1 & 9 & 9 & 17 & 1 & 1 & 1 \\ 0 & -2 & 2 & 1 & 1 & 16 & 9 & 9 \\ 0 & 0 & 1 & 2 & 1 & 9 & 18 & 9 \\ -1 & -1 & 1 & 1 & 1 & 9 & 9 & 17 \end{bmatrix} \quad 8 \times 8 \qquad U'\underline{Y} = \begin{bmatrix} 229 \\ 17 \\ -56 \\ -37 \\ -18 \\ 4 \\ 7 \\ -22 \end{bmatrix}$$

$$(U'U)^{-1} = \begin{bmatrix} 0.03099 & & & & & & & \\ -0.00286 & 0.03099 & & & & & & \\ 0.00547 & -0.00234 & 0.10390 & & & & & \\ -0.00286 & -0.00026 & -0.03359 & 0.08724 & & & & \\ 0.00026 & 0.00286 & -0.03672 & -0.02839 & 0.09349 & & & \\ -0.00234 & 0.00547 & -0.01328 & 0.00547 & 0.00234 & 0.10390 & & \\ -0.00026 & -0.000286 & 0.00547 & -0.00911 & 0.00026 & -0.03359 & 0.08724 & \\ 0.00286 & 0.00026 & 0.00234 & 0.00026 & -0.00286 & -0.03672 & -0.02839 & 0.09349 \end{bmatrix}$$

Thus $\underline{\hat{\theta}}$ is given by:

$$\underline{\hat{\theta}} = [6.769 \quad -0.044 \quad -2.769 \quad -1.544 \quad 1.606 \quad 1.044 \quad 1.019 \quad -1.831]$$

All the above results are computed by LSQRS for the purpose of obtaining the AOV.

On entry to the means separation routine, the matrix denoted by L'^{c} in the previous chapters, is constructed first. In this case L'^{c} is given by:

$$L'^{c} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}_{15 \times 8}$$

It can be verified that $U = XL'^{c}$. Next the $\underline{\ell}$-vectors which are required for the comparison of means are formed and written on disk, sequentially. In this case, effect A is found to be not significant and the comparison of means for effect A are not made. For effect B, since there are 4 levels, six comparisons can be made.

With the notation developed in Section 4.1.3 the means for effect B can be denoted by $\overline{Z}_{.j}$ where $j = 1,2,3$.

Recall that $Z_{ij} = \overline{Y}_{ij.}$ and $Z_{.j} = \sum_i \overline{Y}_{ij.}$

$$\widehat{\beta_j - \beta_{j'} + (\overline{\alpha\beta})_{.j} - (\overline{\alpha\beta})_{.j'}} = \overline{Z}_{.j} - \overline{Z}_{.j'}$$

Thus, $\underline{\ell}_1' \hat{\underline{\beta}} = \bar{Z}_{.1} - \bar{Z}_{.2}$,

where $\underline{\ell}_1' = [0\ 0\ 0\ 1\ -1\ 0\ 0\ \tfrac{1}{2}\ -\tfrac{1}{2}\ 0\ 0\ \tfrac{1}{2}\ -\tfrac{1}{2}\ 0\ 0]_{1 \times 15}$

Note that $\underline{\beta}' = [\mu\ \alpha_1\ \alpha_2\ \beta_1\ \beta_2\ \beta_3\ \beta_4\ (\alpha\beta)_{11}\ (\alpha\beta)_{12}\ (\alpha\beta)_{13}\ (\alpha\beta)_{14}\ (\alpha\beta)_{21}\ (\alpha\beta)_{22}\ (\alpha\beta)_{23}\ (\alpha\beta)_{24}]$

and $\underline{\ell}_2' \hat{\underline{\beta}} = \bar{Z}_{.1} - \bar{Z}_{.3}$,

where $\underline{\ell}_2' = [0\ 0\ 0\ 1\ 0\ -1\ 0\ \tfrac{1}{2}\ 0\ -\tfrac{1}{2}\ 0\ \tfrac{1}{2}\ 0\ -\tfrac{1}{2}\ 0]$.

The $\underline{\ell}$-vectors for estimating the differences of all pairs of means of B are given below:

Comparison	$\underline{\ell}'$														
$\bar{Z}_{.1} - \bar{Z}_{.2}$	0	0	0	1	-1	0	0	½	-½	0	0	½	-½	0	0
$\bar{Z}_{.1} - \bar{Z}_{.3}$	0	0	0	1	0	-1	0	½	0	-½	0	½	0	-½	0
$\bar{Z}_{.1} - \bar{Z}_{.4}$	0	0	0	1	0	0	-1	½	0	0	-½	½	0	0	-½
$\bar{Z}_{.2} - \bar{Z}_{.3}$	0	0	0	0	1	-1	0	0	½	-½	0	0	½	-½	0
$\bar{Z}_{.2} - \bar{Z}_{.4}$	0	0	0	0	1	0	-1	0	½	0	-½	0	½	0	-½
$\bar{Z}_{.3} - \bar{Z}_{.4}$	0	0	0	0	0	1	-1	0	0	½	-½	0	0	½	-½

For the interaction effect, since it turns out to be significant, 28 comparisons can be made. In general,

$$Z_{ij} - Z_{i'j'} \text{ is estimated by}$$

$$\alpha_i - \alpha_{i'} + \beta_j - \beta_{j'} + \alpha\beta_{ij} - \alpha\beta_{i'j'}$$

Thus $\underline{\ell}_1' \hat{\underline{\beta}} = Z_{11} - Z_{12}$

where $\underline{\ell}_1' = [0\ 0\ 0\ 1\ -1\ 0\ 0\ 1\ -1\ 0\ 0\ 0\ 0\ 0\ 0]$

and $\underline{\ell}' \hat{\underline{\beta}} = Z_{11} - Z_{23}$

where $\underline{\ell}_2' = [0\ 1\ -1\ 1\ 0\ -1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ -1\ 0]$

Because it is unnecessary to tabulate the appropriate $\underline{\ell}$-vectors for the set of 28 comparisons of the interaction means, it is omitted here.

Thus all the $\underline{\ell}$-vectors required to compute the estimates of the differences and their standard errors, are made available at this stage. All requirements necessary to make the mean comparisons as indicated in Section 4.1, are supplied to the routine which carries out the computations indicated in Section 4.1.

The results of the calculations for the comparisons of means of effect B are tabulated below:

Comparison	$\underline{\ell}' L'^{c} \underline{\hat{\theta}}$	$\sqrt{\hat{\sigma}^2 \underline{\ell}' L'^{c}(U'U)^{-1}(L'^{c})'\underline{\ell}}$	LSD
$\bar{Z}_{.1} - \bar{Z}_{.2}$	-1.22499	0.855646	1.76263
$\bar{Z}_{.1} - \bar{Z}_{.3}$	-4.37499	0.876103	1.804771
$\bar{Z}_{.1} - \bar{Z}_{.4}$	-5.474998	0.855646	1.762630
$\bar{Z}_{.2} - \bar{Z}_{.3}$	-3.14999	0.820419	1.644986
$\bar{Z}_{.2} - \bar{Z}_{.4}$	-4.24999	0.798537	1.644986
$\bar{Z}_{.3} - \bar{Z}_{.4}$	-1.10000	0.820419	1.690063

The second column gives the estimated differences and the third column, the standard errors. A similar tabulation for the comparisons of the interaction effect of this problem is included in the appendix.

CHAPTER 6

CONCLUSIONS

As demonstrated in previous chapters, it has been possible to make means comparisons and test hypotheses about the parameters of the unrestricted model. This goal has been achieved without the actual computing of a pseudo-inverse of the X'X matrix. Although the procedure involved the construction of the L'^{c} matrix, the major task proved to be the development of the algorithms necessary to construct the vectors corresponding to the comparisons of means.

The portions of the program to carry out the above procedures were developed independent of the LSQRS program and implemented in the program at a later stage. The means separation routines in LSQRS may be made active by using a single parameter card, in which the user specifies a significance level at which comparisons have to be made and optionally, indicates a selection of effects of which the user requires the means separated. If the latter is omitted means of all effects significant at the given level, will be separated.

APPENDIX

COMPUTER PROGRAM AND SAMPLE OUTPUT

```
C
C     THIS SUBROUTINE ACTIVATES THE MEANS SEPARATION PROCEDURES
C
      SUBROUTINE SEPAR( IEFCT,IEDF,MRT)
      INTEGER*2 MINT(3)/3*0/,A(120,60),INTL(3)/3*1/,B(7200)
      INTEGER CRDR,PPTR,DISK
      REAL*8 LABELS(50),TITLE(10)
      COMMON/ALPHA/TITLE,LABELS,CRDR,PRTR,DISK,ACCT(12)
      COMMON/STUFF2/MINTAL(25),KLTS(25),LENTH(15),KADR(25)
      COMMON/STUFF/M2,M3,NEDF
      COMMON/STUFF3/ B
      EQUIVALENCE (A,B)
      IF(MRT.EQ.0)RETURN
      REWIND DISK
      NEDF = IEDF
      DO  6  I=1,7200
    6 B(I) = 0
C
C     MATRIX L' C-INVERSE IS BUILT IN THE 2 DIMENSIONAL ARRAY 'A'
C     ACCORDING TO THE LEVELS OF DIFFERENT EFFECTS THAT ARE IN THE
C     MODEL.CALLS ARE MADE TO MAINF OR INTF ACCORDING TO AS  WHETHER
C      THE EFFECT IS A MAIN EFFECT OR AN INTERACTION.
C
      A(1,1) = 1
      M2 = 1
      M3 = 1
      NADDR = 1
      N = 0
      ND = 0
      NINT = 0
      L1 = 1
      DO  1  I=1,IEFCT
      M = LENTH(I)
      IF(M.GT.50)GO TO 9
   11 IF(M.GT.1) GO TO 2
      IF(MINTAL(L1).LT.0)GO TO 9
      M1 = KLTS(L1)
      IF(ND.NE.0)GO TO 12
      NADDR = NADDR + M1
   12 M2 = M2 + 1
      M3 = M3 + 1
      N = N + 1
      L1 = L1 + 1
      CALL MAINF(M1,M2,M3)
      GO TO 1
    9 IF(ND.NE.0)GO TO 10
    8 ND = N
      NDINT = NINT
      N = 0
      NINT = 0
      IF(M.LT.50 ) GO TO 5
   10 M =      M - (M/50)*50
      GO TO 11
    2 DO  4  J=1,M
      K = 3-M+J
      INTL(K) = KLTS(L1)
    4 L1 = L1 + 1
      CALL INTF(INTL,M2,M3,M)
```

```
    7 NINT = NINT + 1
    1 CONTINUE
      GO TO 8
    5 CONTINUE
C
C     SUBROUTINE VECTOR CONSTRUCTS THE L VECTORS CORRESPONDING TO
C     THE MAIN EFFECT MEAN COMPARISONS
C
      NC = N + NINT
      CALL VECTOR(NADDR,ND,NDINT,NC)
      IF(NDINT.EQ.0)GO TO 25
C
C     SUBROUTINE INTVEC COSTRUCTS L VECTORS FOR INTERACTION MEAN COMPARISONS
C
      CALL INTVEC(JCT,NADDR,NDINT,ND)
   25 END FILE DISK
      RETURN
      END
```

```
C
C     THIS SUBROUTINE COMPLETES  THE PORTIONS OF THE L * C-INVERSE
C     MATRIX  THAT CORRESPOND TO THE MAIN EFFECTS
C
      SUBROUTINE MAINE(M1,M2,M3  )
      INTEGER*2 A(120,60)
      COMMON/STUFF3/A
      IF(M1.NE.1) GO TO  60
      A(M2,M3) = 1
      GO TO 61
   60 K = M2 + M1 - 2
      L = M3 + M1 - 2
      DO  1  I=M2,K
      K1  = I-M2+1
      DO  1  J=M3,L
      IF(K1.EQ.(J-M3+1))A(I,J)= 1
    1 CONTINUE
      DO  2  J=M3,L
    2 A(K+1,J)  = -1
      M2  = K+ 1
      M3  =  L
   61 RETURN
      END
```

```
C
C      THIS SUBROUTINE COMPLETES THOSE PARTS OF THE L'C-INVERSE
C      MATRIX THAT CORRESPOND TO THE INTERACTIONS.
C
       SUBROUTINE INTF(INTL  ,MI,MJ,M)
       INTEGER*2 A(120,60) ,INTL(3)
       COMMON/STUFF3/A
       M1 = INTL(1)
       M2 = INTL(2)
       M3 = INTL(3)
       N1 = M1-1
       N2 = M2-1
       N3 = M3-1
       KC1 =(M1-1)*M3*M2
       KC11 = KC1 + MI
       KC12= KC11 + M3
       KC2 = (M2-1)*M3
       K3  = KC11 + KC2
       K33= KC12+ KC2
       KC21 = KC2 + MI
       KC22 = KC21 + M3
       I = 1
     2 KA1 = (I-1)*M2*M3
       KB1 =(I-1)*N2*N3
       KC23 = KA1 + KC21
       K23 = KA1 + KC22
       DO    1  J =1,N2
       KA2 = (J-1)*M3
       KB2 = (J-1)*N3
       K1    = KA1 + KA2 + MI
       K2   = KB1 + KB2 + MJ
       K11 = K1 + M3
       KC13 = KA2 + KC11
       K13 = KA2 + KC12
       DO   1  K=1,N3
       N  =  K2 + K
       A(K1+K,N)   = 1
       A(K11,N)    =-1
       A(KC23+K,N)=-1
       A(K23,N)    = 1
       IF(M.EQ. 2) GO TO 1
       A(KC13+K,N)=-1
       A(K13 ,N)   = 1
       A(K3+K,N)   = 1
       A(K33,N)    =-1
     1 CONTINUE
       I = I+ 1
       IF(I.LE.N1) GO TO 2
       MI = K33
       MJ = N
       RETURN
       END
```

```
C
C      THIS SUBROUTINE CONSTRUCTS THE L VECTORS CORRESPONDING  TO
C      THE MAIN EFFECT MEAN COMPARISONS USING MODEL DATA PASSED FROM
C       THE MAIN PROGRAM THROUGH COMMON/STUFF2/
C
       SUBROUTINE VECTOR(NADDR, N,NINT,NC)
       INTEGER*4 MDEV(3),INTLTS(3)/3*1/
       INTEGER CRDR,PRTR,DISK
       INTEGER*4 INTL(3),LM1(5),LM2(5),MINT(3),ISUB(3)
       REAL*4 L(120)/120*0./,MM
       REAL*8 LABELS(50),TITLE(10)
       COMMON/ALPHA/TITLE,LABELS,CRDR,PRTR,DISK,ACCT(12)
       COMMON/STUFF2/MINTAL(25),KLTS(25),LENTH(15),KADR(25)
       COMMON/MRTS/MRT,IALPHA,ISW(25),LOSTA,NDEFCT,KONT(28),NKONT
       EQUIVALENCE (INTL(1),M1),(INTL(2),M2),(INTL(3),M3),(INTLTS(1),JM)
       EQUIVALENCE(LM1(1),J),(LM2(1),JJ),(ISUB(1),JD)
       EQUIVALENCE (MDEV(1),MD1),(MDEV(2),MD2),(MDEV(3),MD3)
C
C      IN THIS LOOP SUCCESSIVELY INSERT  PAIRS OF +1 AND -1 IN SLOTS
C      ALLOTED FOR EACH OF THE MAIN EFFECTS IN THE L VECTOR; ADJUST
C      ADDRESSING THRO' MADDR
C
       KJ = 1
       MADDR = 1
       DO  1  I=1,N
       JM1 = KLTS(I) - 1
       JM  = KLTS(I)
       IF(ISW(I).NE.0)GO TO 5
       JD  = MINTAL(I)
       IF(JD.LT.0)GO TO 5
       NSFT = 1
       DO  3  J=1,JM1
       K = J + 1
       DO  3  JJ=K,JM
       L(J+MADDR) = 1
       L(JJ + MADDR)  = -1
       MADR1 = MADDR
       L1 = N+1
       IF(NINT.EQ.0)GO TO 4
C
C      IN THIS LOOP  ADJUST REST OF THE L VECTOR(INTERACTIONS) DEPENDING
C      ON THE COMPARISON AS INDICATED BY THE MAIN EFFECT PORTION OF L
C      ; ADDRESSING ACCOMPLISHED BY MEANS OF THREE CO-ORDINATES
C      CORRESPONDING TO THE EFFECT SUBSCRIPT
C
       DO  11  IK=1,NINT
       M = LENTH(N+IK)
       IP = 0
       M1 = 1
       M2 = 1
       M3 = 1
       DO  12  II=1,M
       MINT(II) = MINTAL(L1)
       INTL(II) = KLTS(L1)
       IF(JD.EQ.MINT(II))IP = II
    12 L1 = L1 + 1
       IF(M.EQ.1)GO TO 11
       IF(IP.EQ.0) GO TO 11
```

```
      MPROD = M1*M2*M3
      II = IP + 1
      IF(II.GT.3)II = 1
      MM2 = INTL(II)
      MM3 = MPROD/(INTL(IP)*MM2)
      MM = JM
      MM = MM/MPROD
      DO  7  II=1,3
      MDEV(II) = 1
      DO  8  IJ=1,IP
    8 MDEV(II) = MDEV(II)*INTL(IJ)
      IP = IP + 1
      IF(IP.GT.3)IP=1
    7 MDEV(II) = MPROD/MDEV(II)
      K1 = (J-1)*MD1
      J1 = (JJ-1)*MD1
      DO  9  KA=1,MM2
      K2 = (KA-1)*MD2
      K3 = NADR1 + K1 + K2
      K4 = NADR1 + J1 + K2
      DO  9  KB=1,MM3
      K5 = (KB-1)*MD3
      L(K3+K5+1) = L(K3+K5+1) + MM
    9 L(K4+K5+1) = L(K4+K5+1) - MM
   11 NADR1= NADR1 + M1*M2*M3
C
C      WRITE L ON DISC  ALONG WITH  COMPARISON SUBSCRIPTS, EFFECT
C     CONSIDERED AND NSET INDICATING  WHETHER THIS EFFECT IS DIFFERENT
C     FROM THE PREVIOUS ONE (FOR PAGING PURPOSES )
C
    4 WRITE(DISK)KJ,ISUB,NSET,LM1,LM2,L,INTLTS,MADDR
      DO  2  K=1,120
    2 L(K) = 0.
      NSET = 0
    3 CONTINUE
    5 MADDR  = MADDR + JM
    1 CONTINUE
      RETURN
      END
```

```
C
C     THIS SUBROUTINE CONSTRUCTS L VECTORS CORRESPONDING TO THE
C     COMPARISONS  OF INTERACTION MEANS USING MODEL DATA PASSED FROM
C      THE MAIN PROGRAM THRO' COMMON/STUFF2/.
C
      SUBROUTINE  INTVEC(JCT,NADDR,NINT,N)
      INTEGER*4 LM1(5),LM2(5),INTAD(5),INTLTS(5),MINT(3),INTL(4)
      INTEGER*4 IVECT1( 5),IVECT2( 5),ISUB(3),ILTS(3)
      INTEGER CRDR,PRTR,DISK
      REAL*8 LABELS(50),TITLE(10)
      REAL*4 L(120)/120*0./,MM
      LOGICAL SW(3),SW1,SW2,SW3
      COMMON/ALPHA/TITLE,LABELS,CRDR,PRTR,DISK,ACCT(12)
      COMMON/STUFF2/MINTAL(25),KLTS(25),LENTH(15),KADR(25)
      COMMON/MRTS/MRT,IALPHA,ISW(25),LDSTA,NDEFCT,CONT(28),NCONT
      EQUIVALENCE(SW(1),SW1),(SW(2),SW2),(SW(3),SW3)  ,
     1 (ILTS(1),INTLTS(1))
      EQUIVALENCE (M1,INTL(1)),(M2,INTL(2)),(M3,INTL(3)),(M4,INTL(4))
      IB = N+1
C     THIS PORTION OF THE SUBROUTINE INSERTS PAIRS OF +1 AND -1
C     SUCCESSIVELY  IN THE LVECTOR  IN THE SLOTS ALLOTTED  FOR
C     MAIN EFFECTS  USING  EACH OF THE INTERACTIONS  TO OBTAIN
C     APPROPRIATE ADDRESSES
C
      MADDR = NADDR
      DO  30  I1=1,NINT
      DO  1  I=1,5
      IVECT1(I)=0
      IVECT2(I)=0
      INTLTS(I) = 1
      LM1(I) = 1
    1 LM2(I) = 1
      IP = 1
      I = LENTH(N+I1)
      DO  31  I2=1,I
      ISUB(I2) = MINTAL(IB)
      IF(ISUB(I2).LT.0)GO TO 30
      INTLTS(I2) = KLTS(IB)
      INTAD(I2) = KADR(IB)
      IB = IB+1
   31 IP = IP*INTLTS(I2)
      IP = IP*I
      IF(ISW(I1+N).NE.0)GO TO 30
      NSET
      SW1= .TRUE.
      SW2 = .TRUE.
      SW3 =.TRUE.
      L1 = 1
   10 IVECT1(L1) = IVECT1(L1) + 1
    5 IF(IVECT1(L1).LE.INTLTS(L1)) GO TO 4
      IVECT1(L1)    = 0
      L1 = L1 -1
      IF(L1.EQ.0)GO TO 30
      SW(L1) = .TRUE.
      GO TO 5
    4 L2 = L1
    8 IVECT2(L2) = IVECT2(L2) + 1
      IF(IVECT2(L2).LE.INTLTS(L1)) GO TO 6
```

```
      IVECT2(L2) = 0
      SW(L2) = .TRUE.
      GO TO 10
    6 IF(IVECT2(L2) - IVECT1(L2))11,63,61
   61 SW(L2) = .FALSE.
   63 IF(L2.GE.I) GO TO 11
      L1 = L2 + 1
      GO TO 10
   11 IF(SW1.AND.SW2.AND.SW3) GO TO 8
   81 J1 = 0
      DO  13  K=1,I
      L3 = I   -K+1
      L4 = I-K+1
      J1 = INTAD(L3)
      J2 = IVECT1(L3)
      J3 = IVECT2(L3)
      LM1(L4) = J2
      LM2(L4) = J3
      L(J1+J2) = L(J1+J2) + 1
      L(J1+J3) = L(J1+J3) - 1
   13 CONTINUE
      IF( J1.EQ.0) GO TO 8
C
C     THIS LOOP USES   THE SUBSCRIPT OF THE MAIN EFFECT SET ABOVE
C     TO ADJUST THE REST OF L (INTERACTIONS) ACCORDINGLY
C
      NADR1 = NADDR
      LL = N+1
      DO  14  IK=1,NINT
      M = LENTH(N+IK)
      M1 = 1
      M2 = 1
      M3 = 1
      M4 = 1
      IP = 1
      DO  15  II=1,M
      MINT(II) = MINTAL(LL)
      INTL(II) = KLTS(LL)
   15 LL = LL+1
      MPROD = M1*M2*M3
      MP = 1
      K1 = 0
      J1 = 0
      MD2 = 0
      DO  2  II=1,M
      MP = MP* INTL(II)
      DO  3  IJ=1,I
      IF(ISUB(IJ).NE.MINT(II)) GO TO 3
      K1 = K1 + (LM1(IJ) - 1)*(MPROD/MP)
      J1 = J1 + (LM2(IJ) - 1)*(MPROD/MP)
      IP = IP*INTL(II)
      MD2 = MD2*INTL(II)
      GO TO 2
    3 CONTINUE
      MD2 = 1
    2 CONTINUE
      MM2 = MPROD/IP
      MM = IP
```

```
      MM = MM/MPROD
      DO  9  KA=1,MM2
      K2 = (KA-1)*MD2
      K3 = NADR1 + K1 + K2
      K4 = NADR1 + J1 + K2
      L(K3+1) = L(K3+1) + MM
    9 L(K4+1) = L(K4+1) - MM
   14 NADR1 = NADR1 + MPROD
C
C     WRITE  L ON DISC ALONG WITH SUBSCRIPT VALUES,EFFECT CONSIDERED
C     AND NSET INDICATING WHETHER THIS IS IS A DIFFERENT EFFECT
C     FROM THE PREVIOUS ONE (FOR PAGING PURPOSES )
C
      WRITE(DISK)I,ISUB,NSET,LM1,LM2,L,ILTS,MADDR
      NSET = 0
      DO  7  I2=1,120
    7 L(I2) = 0.
      GO TO 8
   30 MADDR = MADDR+ IR
      CONTINUE
      RETURN
      END
```

```
C
C     CALCULATE  THE DIFFERANCES  AND THE STD. ERRORS USING THE L'C
C     MATRIX AND THE LVECTORS CONSTUCTED FOR  COMPARISONS
C
      SUBROUTINE LSDCAL(NSUB,JJ,NEST)
      INTEGER CRDR,PRTR,DISK
      INTEGER*4 LM1(5),LM2(5),ISUB(3),INTLTS(3),ACTDST(150)
      REAL*8 ERRMS,DSQRT,DABS
      REAL*8 XL(120),ESTD,ESTS,FLSD
      REAL*4  L(120)
      INTEGER*2   A(120,60)
      REAL*8 TITLE(10),LABELS(50),  XXIN,RHMS,EMS(25),TSS,VECTT(60),VPXX
     1V(500),INAM(5)
      REAL SP/'    '/,STAR(2)/' ** ','  * '/,SIGN(2)/' 1% ',' 5% '/
      COMMON NF,NOCLAS,KENSKP(2),NSTSWT,NSTMIS(9),AVG(100),BETA(1386)
      COMMON/ALPHA/TITLE,LABELS,CRDR,PRTR,DISK,ACCT(12)
      COMMON /MATRIX/ INTNAM(25),INTLOC(25),IRHM(25),IDSTAA(26),IVAR(8),N
     1,MUSW,IEFCT,NRHM,IAHDF,TSS(25),  XXIN(3300),IDSTA(200),COVECT(60),
     2X(47),LSTCLS(5),INTSTR,NN,NNAM,LEVELS(4),JK
      COMMON/STUFF/MI,MJ,IEDF
      COMMON/STUFF3/A
      COMMON/MRTS/MRT,IALPHA,ISW(25),LDSTA,NDEFCT,KONT(28),NKONT
      EQUIVALENCE  (EMS(1),XXIN(2456)),(ACTDST(1),XXIN(3042))
      MI = LDSTA
      MJ = N
      F = SIGN(IALPHA )
      FSTAR = STAR(IALPHA)
      TT= T(IEDF,IALPHA)
C
C     READ EACH L VECTOR
C
      REWIND DISK
    8 READ(DISK,END=9)M,ISUB,NSET,LM1,LM2,L,INTLTS,MADDR
C
C      MULTIPLY THE LVECTOR BY L' C-INVRSE
C
      DO  11  J=1,MJ
      XL(J) = 0.
      DO  11  I=1,MI
   11 XL(J) = XL(J) + L(I) * A(I,J)
      ESTD = 0.00
      ESTS = 0.00
C
C     FORM THE DIFFERENCE BY MULTIPLYING  IT BY BETA HAT
C
      DO  12  J=1,MJ
   12 ESTD = ESTD + XL(J) * BETA(NSUB+J)
C
C      POST- AND PREMULTIPLY X'X INVERSE BY IT TO GET THE STD. ERROR
C      USING  THE ERRMS AS SIGMA HAT
C
      DO  13  I=1,MJ
      KK = I
      DO  13  J=1,MJ
      ESTS = ESTS + XL(I)*XL(J)*XXIN(KK)
      IF(J-I)14,15,15
   15 K = 1
      GO TO 13
```

```
     14 K = MJ - J
     13 KK = KK + K
        IF(NEST.NE.0)GO TO 7
        ESTS = DSQRT(ESTS*EMS(JJ))
        FLSD =TT*ESTS
        FMARK = SP
        IF(DABS(FSTD).GT.FLSD) FMARK = STAR(IALPHA)
        IF(NSET.EQ.0) GO TO 10
C
C        PRINT EACH COMPARISON
C
        CALL XTITLE
        WRITE(PRTR,201)IRHM(JJ),LABELS(IRHM(JJ)),FSTAR,F
    201 FORMAT('-',40X,'LEAST SIGNIFICANT DIFFERENCES FOR VAR. ',I3,
       1 ' - ',A8// 48X,A4,'DENOTES SIGNIFICANCE AT',A4,'LEVEL'//)
        WRITE(PRTR,202)
    202 FORMAT( 8X,'-------L E V E L S  O F---------',8X,'---------L E V E
       1L S  O F-------',     5X,'DIFFERENCE',8X,'STD.ERROR',8X,'L.S.D.'/)
        GO TO 5
      7 IF(NSET.EQ.0) GO TO 10
        CALL XTITLE
        WRITE(PRTR,205)IRHM(JJ),LABELS(IRHM(JJ))
    205 FORMAT('-',20X,'USE THE MULTIPLIER AND A PROPER ERROR MEAN SQUARE
       1TO CALCULATE THE STD.ERRORS OF EACH COMPARISON '/40X,' STD.ERROR
       2= SQRT ( MULTIPLIER * ERROR MEAN SQUARE )'/55X,'VAR.NO. ',I3,
       3' - ',A8//)
        WRITE(PRTR,206)
    206 FORMAT( 8X,'-------L E V E L S  O F---------',8X,'---------L E V E
       1L S  O F-------',5X,'DIFFERENCE',8X,'MULTIPLIER',8X,'STD.ERR'/)
      5 WRITE(PRTR,203)(LABELS(ISUB(K)),K=1,NSET)
        WRITE(PRTR,204)(LABELS(ISUB(K)),K=1,NSET)
    203 FORMAT(/ 8X,3(A8,2X))
    204 FORMAT('+',46X,3(A8,2X))
C
C       THIS ROUTINE  CHANGES DISCRETE LEVEL NOS. USED TILL NOW TO
C       THEIR ACTUAL NAMES AS GIVEN IN THE DISCRETE CONTROL CARDS
C
     10 KM = 1
        JP = 0
        JQ = 0
        DO  2  I=1,M
        L3 = M-I+1
        JP = JP + (LM1(L3)-1)*KM
        JQ = JQ + (LM2(L3)-1)*KM
      2 KM = KM*LNTLTS(L3)
        JP = JP*M + MADDR - 1
        JQ = JQ*M + MADDR - 1
        DO   3  I=1,M
        LM1(I) = ACTDST(JP+I)
      3 LM2(I) = ACTDST(JQ+I)
        WRITE(PRTR,208)(LM1(I),I=1,M)
    208 FORMAT(10X,3(I2,9X))
        WRITE(6,209)(LM2(I),I=1,M)
    209 FORMAT('+',48X,3(I2,8X))
        IF(NEST.NE.0)GO TO 6
        WRITE(6,210)ESTD,ESTS,FLSD ,FMARK
    210 FORMAT('+', 78X,F15.6,F17.6,F14.6,A4)
        GO TO 8
```

```
FORTRAN IV G LEVEL  21              LSDCAL              DATE = 77060          15/14/52              PAGE 0003

              6 WRITE(6,210)ESTD,ESTS
                GO TO 8
              9 RETURN
                END
```

```
DEPARTMENT OF STATISTICS AND THE STATISTICAL LABORATORY - UNEQUAL SUBCLASS ANALYSIS OF VARIANCE DATE OF RUN   3/09/77

                    JOB -- RCL14384                              USER -- MERVYN
TITLE,EXAMPLE FROM 'STATISTICAL PRINCIPLES IN EXPERIMENTAL DESIGN' BY B.J.WINER                      (STAT CONTROL CARD)
DATA,3.                                                                                              (STAT CONTROL CARD)
LABEL(1)=A,B,Y                                                                                       (STAT CONTROL CARD)
RHM,3.                                                                                               (STAT CONTROL CARD)
DISCRETE,1=1-2.                                                                                      (STAT CONTROL CARD)
DISCRETE,2=1-4.                                                                                      (STAT CONTROL CARD)
OUTPUT,MEANS                                                                                         (STAT CONTROL CARD)
MRT,LSD,0.05                                                                                         (STAT CONTROL CARD)
MODEL,Y=U+1+2+1*2.                                                                                   (STAT CONTROL CARD)
END                                            PARAMETER COMPILE TIME =   0.37 SECONDS - VERSION VII  (STAT CONTROL CARD)
                    FORMAT(2F2.0,F3.0)
```

DEPARTMENT OF STATISTICS AND THE STATISTICAL LABORATORY - UNEQUAL SUBCLASS ANALYSIS OF VARIANCE DATE OF RUN 3/09/77

EXAMPLE FROM 'STATISTICAL PRINCIPLES IN EXPERIMENTAL DESIGN' BY B.J.WINER

JOB -- RCL14384 USER -- MERVYN

BASIC STATISTICS ON INPUT DATA

33 OBSERVATIONS READ IN

33 OBSERVATIONS USED

0 OBSERVATIONS DROPPED

FIRST OBSERVATION

1	2	3
A	B	Y
1.000000	1.000000	3.000000

DEPARTMENT OF STATISTICS AND THE STATISTICAL LABORATORY - UNEQUAL SUBCLASS ANALYSIS OF VARIANCE DATE OF RUN 3/09/77

EXAMPLE FROM 'STATISTICAL PRINCIPLES IN EXPERIMENTAL DESIGN' BY B.J.WINER

JOB -- RCL14384 USER -- MERVYN

MEANS AND VARIANCES

VAR. NO.	VAR. NAME	MEAN	VARIANCE	STD. DEV.
1	A	1.4545	0.25568182	0.50565
2	B	2.5758	1.25189394	1.11888
3	Y	6.9394	8.30871212	2.88248

DATA INPUT AND COMPUTATION TIME 0.3 SECONDS
SSCP TIME = 0.41

DEPARTMENT OF STATISTICS AND THE STATISTICAL LABORATORY - UNEQUAL SUBCLASS ANALYSIS OF VARIANCE DATE OF RUN 3/09/77

EXAMPLE FROM 'STATISTICAL PRINCIPLES IN EXPERIMENTAL DESIGN' BY B.J.WINER

JOB -- RCL14384 USER -- MERVYN

SUBMATRIX DISTRIBUTION

```
         A
   18     15

         B
    7      9      8      9

         A         * B
    4      5      4      5      3      4      4      4
```

DEPARTMENT OF STATISTICS AND THE STATISTICAL LABORATORY - UNEQUAL SUBCLASS ANALYSIS OF VARIANCE DATE OF RUN 3/09/77

EXAMPLE FROM 'STATISTICAL PRINCIPLES IN EXPERIMENTAL DESIGN' BY B.J.WINER

JOB -- RCL14384 USER -- MERVYN

ANALYSIS OF VARIANCE FOR VARIABLE NO. 3 - Y

SOURCE	D.F.	SUMS OF SQUARES	MEAN SQUARES	F-RATIO	PROB.
A	1	0.06176451	0.06176451	0.022	0.8838
B	3	158.95093728	52.98364258	18.695	0.0000
A * B	3	43.99129191	14.66376305	5.174	0.0064
RESIDUAL	25	70.85140991	2.83405590		
TOTAL	32	265.87866211			

BETA VECTOR

1	2	3	4	5	6	7	8
6.76874542	-0.04374993	-2.76874828	-1.54374886	1.60624790	1.04374981	1.01874924	-1.83124924

ANOVA EXEC TIME = 0.1 SECONDS

DEPARTMENT OF STATISTICS AND THE STATISTICAL LABORATORY - UNEQUAL SUBCLASS ANALYSIS OF VARIANCE DATE OF RUN 3/09/77

EXAMPLE FROM 'STATISTICAL PRINCIPLES IN EXPERIMENTAL DESIGN' BY B.J.WINER

JOB -- RCL14384 USER -- MERVYN

MEANS AND STANDARD ERRORS FOR VARIABLE NO. 3 - Y

SOURCE	SUBCLASS LEVELS		MEAN	STANDARD ERROR
	A			
A	1		6.7249947	0.399269
A	2		6.8124952	0.438052
	B			
B	1		3.9999971	0.642884
B	2		5.2249966	0.564651
B	3		8.3749933	0.595195
B	4		9.4749947	0.564651
	A	B		
A * B	1	1	4.9999962	0.841733
A * B	1	2	6.1999950	0.752869
A * B	1	3	6.4999933	0.841733
A * B	1	4	9.1999941	0.752869
A * B	2	1	2.9999971	0.971949
A * B	2	2	4.2499971	0.841733
A * B	2	3	10.2499924	0.841733
A * B	2	4	9.7499943	0.841733

DEPARTMENT OF STATISTICS AND THE STATISTICAL LABORATORY - UNEQUAL SUBCLASS ANALYSIS OF VARIANCE DATE OF RUN 3/09/77

EXAMPLE FROM 'STATISTICAL PRINCIPLES IN EXPERIMENTAL DESIGN' BY B.J.WINER

JOB -- RCL14384 USER -- MERVYN

LEAST SIGNIFICANT DIFFERENCES FOR VAR. 3 - Y

* DENOTES SIGNIFICANCE AT 5% LEVEL

-------L E V E L S O F--------	---------L E V E L S O F-------	DIFFERENCE	STD.ERROR	L.S.D.	
B	B				
1	2	-1.224999	0.855647	1.762632	
1	3	-4.374996	0.876103	1.804772	*
1	4	-5.474998	0.855647	1.762632	*
2	3	-3.149997	0.820420	1.690064	*
2	4	-4.249998	0.798538	1.644987	*
3	4	-1.100001	0.820420	1.690064	

DEPARTMENT OF STATISTICS AND THE STATISTICAL LABORATORY - UNEQUAL SUBCLASS ANALYSIS OF VARIANCE DATE OF RUN 3/09/77

EXAMPLE FROM 'STATISTICAL PRINCIPLES IN EXPERIMENTAL DESIGN' BY B.J.WINER

JOB -- RCL14384 USER -- MERVYN

LEAST SIGNIFICANT DIFFERENCES FOR VAR. 3 - Y

* DENOTES SIGNIFICANCE AT 5% LEVEL

-------L E V E L S O F--------		---------L E V E L S O F-------		DIFFERENCE	STD.ERROR	L.S.D.	
A	B	A	B				
1	1	1	2	-1.199999	1.129303	2.326363	
1	1	1	3	-1.499997	1.190390	2.452202	
1	1	1	4	-4.199998	1.129303	2.326363	*
1	2	1	3	-0.299998	1.129303	2.326363	
1	2	1	4	-2.999999	1.064717	2.193317	*
1	3	1	4	-2.700001	1.129303	2.326363	*
1	1	2	1	2.000000	1.285768	2.648681	
1	1	2	2	0.750000	1.190390	2.452202	
1	1	2	3	-5.249995	1.190390	2.452202	*
1	1	2	4	-4.749997	1.190390	2.452202	*
1	2	2	1	3.199999	1.229429	2.532624	*
1	2	2	2	1.949999	1.129303	2.326363	
1	2	2	3	-4.049997	1.129303	2.326363	*
1	2	2	4	-3.549999	1.129303	2.326363	*
1	3	2	1	3.499997	1.285769	2.648681	*
1	3	2	2	2.249997	1.190390	2.452202	
1	3	2	3	-3.749998	1.190390	2.452202	*
1	3	2	4	-3.250000	1.190390	2.452202	*
1	4	2	1	6.199998	1.229429	2.532624	*
1	4	2	2	4.949998	1.129303	2.326363	*
1	4	2	3	-1.049998	1.129303	2.326363	
1	4	2	4	-0.549999	1.129303	2.326363	
2	1	2	2	-1.250000	1.285769	2.648681	
2	1	2	3	-7.249995	1.285768	2.648681	*
2	1	2	4	-6.749997	1.285768	2.648681	*
2	2	2	3	-5.999995	1.190390	2.452202	*
2	2	2	4	-5.499997	1.190390	2.452202	*
2	3	2	4	0.499998	1.190390	2.452202	

MEANS EXECUTION TIME = 0.86 SECONDS

REFERENCES

[1] BENTZ, CAROL ANN (1972) Computational Techniques for the analysis of the General Linear Model. A Master's Report Department of Statistics, KANSAS STATE UNIVERSITY, Manhattan.

[2] GRAYBILL, FRANKLIN. A. (1976) Applications and Theory of the Linear Model.

[3] JOHNSON, DALLAS. E. (1976) Linear Models I and II class notes. Department of Statistics, KANSAS STATE UNIVERSITY.

[4] KEMP, KENNETH. E. (1972) Least Squares Analysis of Variance; A procedure, A program, and Examples of their use. Part One and Two. Contribution 168 Department of Statistics and the Statistical Laboratory, KANSAS STATE UNIVERSITY.

[5] SPEED F. M. and HOCKING R. R. (1976) The Use of the R( )-notation with Unbalanced Data. The American Statistician February 1976. Vol. 30, No.1.

[6] WINER B. J. (1962) Statistical Principles in Experimental Design. McGraw Hill Publishing Company, New York.

ACKNOWLEDGEMENTS

The author wishes to express his appreciation to Dr. Kenneth E. Kemp for his guidance and encouragement during this project and also for making the necessary modifications to the LSQRS program during the implementation procedure.

The author is also grateful to Dr. George A. Milliken and Dr. Dallas E. Johnson for the valuable suggestions made during the development of the theoretical material used in the report. They are also appreciated for serving on the author's committee.

The author is also indebted to his beloved wife Sumi for the superb typing of this report.

IMPLEMENTATION OF MULTIPLE COMPARISON PROCEDURES
IN A GENERALIZED LEAST SQUARES PROGRAM

by

MERVYN G. MARASINGHE
B.S., University of Sri Lanka, 1970

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1977

An algorithm for the computation of differences of means and their standard errors is implemented in the LEAST SQUARES ANALYSIS OF VARIANCE (LSQRS) program, which is a program developed and maintained by the Statistical Laboratory, Department of Statistics of the Kansas State University.

The basic function of the algorithm is the construction of a transformation matrix for the purpose of converting a non-full rank linear model to one of full rank, using information regarding the design model to be analyzed. This matrix and the availability of the inverse of the transformed (or reduced) sums of squares matrix, make possible the calculation of estimates of estimable functions of the parameters of the original design model along with estimates of their variances.

In addition, this matrix could be used to obtain the sums of squares for testing any testable hypothesis of the original model.