The Philippine Statistician
Volume XVI, Nos. 3/4, 1967.

ON THE MATRIX APPROACH FOR TESTING LINEAR HYPOTHESES IN RANDOMIZED COMPLETE BLOCK DESIGNS

by

SANTIAGO STA. INES[1]

1. **Preliminaries.** One purpose of an experiment is to determine which treatment is the best. In a randomized complete block design the classical analysis involves a method in which the total sum of squares is partitioned into different sums of squares attributed to treatments, block and error. The classical method of analysis of variance utilizes an F-test to detect the significance of these treatments. However, one could isolate the desired sum of squares for any linear contrast of treatments with the use of matrices instead of the usual way of splitting the total sum of squares.

The usual hypotheses to be tested are:

(a) All treatment effects equal zero;

(b) The linear, quadratic, cubic and quartic contrasts, depending on the design, are each equal to zero. Both hypotheses can be expressed as $C\lambda = O$ where C is a given matrix consisting of s independent row vectors and λ is a column vector of unknown parameters to be estimated.

Roy[1] presented an expression for the F-test criterion for the univariate analysis of variance in terms of matrices as follows:

$$F = \frac{X'A_1 (A'_1A_1)^{-1} C'_1 [C_1 (A'_1A_1)^{-1} C'_1]^{-1} C_1 (A'_1A_1)^{-1} A'_1X / s}{X' [I(n) - A_1 (A'_1A_1)^{-1} A'_1] X / (n-r)}$$

[1] Major, Philippine Military Academy.

with s and (n-r) degrees of freedom, where

$X = (X_1, X_2, \ldots, X_n)$ is a 1 by n vector of n independent stochastic variates with a common unknown variance σ^2,

A_1 is a basis of the incidence matrix A with elements

0 and 1 for experiments that do not involve regression,

C_1 is the submatrix of C with elements that depend on the hypothesis,

s is the number of independent rows in the matrix C,

n is the number of independent variates or observations,

I(n) is the n by n identity matrix.

For brevity, all matrices in the expression (1) between X' and X will be designated by an equivalent expression for the numerator with

$$M'M = A_1 (A'_1A_1)^{-1} C'_1 [C_1 (A'_1A_1)^{-1} C'_1]^{-1} C_1 (A'_1A_1)^{-1} A'_1 \tag{2}$$

and the denominator with

$$L'L = I(n) - A_1 (A'_1A_1)^{-1} A'_1. \tag{3}$$

Thus, the expression (1) can now be written in its equivalent form

$$F = \frac{X'M'MX / s}{X'L'LX / (n\text{-}r)} \tag{4}$$

Both expression (1) and its equivalent expression (4) are for the variance ratio F in the analysis of variance and are expressed in terms of quantities involving raw data directly obtained from the experiment and the hypothesis to be tested.

In a randomized complete block design the linear estimation model is

$$E\ (X_{ij}) = \mu + \beta_i + \tau_j \tag{5}$$

where

μ = the true mean,

β_i = the effect of the i-th block (i = 1, 2, ..., q),

τ_j = the effect of the j-th experiment (j = 1, 2, ..., p).

in matrix form, the expression (5) is

$$E(X) = A\lambda \qquad (6)$$

where A is n by m, λ is an m by 1 unknown matrix of parameters to be estimated and m = p + q + 1 is the total number of column vectors of the matrix A and is also the total number of unknown parameters to be estimated.

2. **Testing the Hypotheses.** In testing a linear hypothesis, the matrix expression C = O is used. For the null hypothesis that all treatment effects equal zero, H_0: i = O (i = 1, ..., p) the expression is

$$\left[\begin{array}{cccc|ccccc} 0 & 0 & \ldots & 0 & 1 & -1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 1 & 0 & -1 & \ldots & 0 \\ . & . & \ldots & . & . & . & . & \ldots & . \\ 0 & 0 & \ldots & 0 & 1 & 0 & 0 & \ldots & -1 \end{array}\right] \begin{bmatrix} \lambda_1 \\ \lambda^2 \\ . \\ \vdots \\ \lambda_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ . \\ \vdots \\ 0 \end{bmatrix}$$

For the hypothesis that a particular contrast equals zero the expression is

$$(0 \quad 0 \ldots 0 \mid C_1) \begin{bmatrix} \lambda_1 \\ \lambda^2 \\ . \\ \vdots \\ \lambda_n \end{bmatrix} = 0 \text{ (a scalar)}$$

where C_1 is 1 by p and the left partition of zeroes is 1 by q-1. In both (7) and (8), $\lambda' = (\mu \ \beta_1 \ \beta_1 \ \ldots \ \beta_q \ \tau_1 \ \ldots \ \tau_q)$ and the elements for C_1 are the coefficients representing the particular contrast to be tested.

3. **The Generalized Matrices.** The matrices M'M and L'L have been generalized for p treatments and q replications. For the null hypothesis that all treatment effects equal zero, i.e. H_0: $\tau_i = 0$ (i = 1, 2, ..., p),

$$M'M = \frac{1}{pq} (R_{ij}) \quad (i, j = 1, 2, \ldots, q) \tag{9}$$

where the submatrices R_{ij}'s are all identically given by the expression (a p by p matrix)

$$R_{ij} = \begin{pmatrix} p\text{-}1 & \text{-}1 & \ldots & \text{-}1 \\ \text{-}1 & p\text{-}1 & \ldots & \text{-}1 \\ . & . & \ldots & . \\ \text{-}1 & \text{-}1 & \ldots & p\text{-}1 \end{pmatrix} \tag{10}$$

For the null hypothesis that a particular contrast equals zero, i.e. H_o: $\sum_{i=1}^{p} K_i \tau_i = 0$ where K_j is the orthogonal coefficient of the i-th treatment as taken from the tables on orthogonal coefficients by Fisher and Yates. [2]

$$M'M = (1/ q \sum_{i=1}^{p} K^2_i)(R_{ij}) \qquad (i, j = 1, 2, \ldots, q) \tag{11}$$

where the submatrices R_{ij} are all identical, i.e.

$$R_{ij} = \begin{pmatrix} K_1 & K_1K_2 & \ldots & K_1K_p \\ K_2K_1 & K_2 & \ldots & K_2K_p \\ . & . & \ldots & . \\ K_pK_1 & K_pK_2 & \ldots & K^2 \end{pmatrix} \tag{12}$$

For all linear hypotheses to be tested, the matrix L'L remains the same for p treatments and q replications. The generalized matrix L'L is given by

$$L'L = \frac{1}{pq} ({}^*R_{ij}), \qquad (i, j = 1, 2, \ldots, q) \tag{1}$$

where the submatrix ${}^*R_{ij}$ on the main diagonal is (the p by p)

$${}^*R_{ij} = \begin{pmatrix} (p\text{-}1)(q\text{-}1) & (1\text{-}q) & \ldots & (1\text{-}q) \\ (1\text{-}q) & (p\text{-}1)(q\text{-}1) & \ldots & (1\text{-}q) \\ . & . & \ldots & . \\ (1\text{-}q) & (1\text{-}q) & \ldots & (p\text{-}1)(q\text{-}1) \end{pmatrix} \quad (i = j)$$

which is the submatrix (10) multiplied by (q-1) and the p by p submatrix submatrix $^*R_{ij}$ not on the main diagnoal is

$$^*R_{ij} = \begin{pmatrix} (1-p) & 1 & \cdots & 1 \\ 1 & (1-p) & \cdots & 1 \\ . & . & \cdots & . \\ 1 & 1 & \cdots & (1-p) \end{pmatrix} \quad (i \neq j)$$

which is the negative of the matrix (13).

4. **Conclusion.** Within the generalization of the matrices M′M and L′L for use in testing linear hypotheses, an experimenter utilizing the randomized complete block design of p treatments and q replications has at his disposal a tool for directly isolating the desired sum of squares in testing a particular hypothesis.

For a single variate the linear estimation model is

$$E \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{12} & a_{22} & \cdots & a_{2m} \\ . & . & \cdots & . \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}$$

and for p-variates the linear estimation model is

$$E \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ . & . & \cdots & . \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ . & . & \cdots & . \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1p} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2p} \\ . & . & \cdots & . \\ \lambda_{1m} & \lambda_{m2} & \cdots & \lambda_{mp} \end{pmatrix}$$

It is seen from the expressions (16) and (17) that the matrix A remains unchanged. Hence, the same matrices as presented can be used for both the univariate and multivariate analysis.

REFERENCES

[1] ROY, S. N. *Some Aspects of Multivariate Analysis.* New York: John Wiley and Sons, 1957, p. 81.

[2] FISHER, R. A. and F. YATES. *Statistical Tables for Biological, Agricultural, and Medical Research.* New York: Hafner Publishing Company, 1957, pp. 80-90.