

Covariance Structures of Linear Models

A. K. Gupta
Bowling Green State University
Bowling Green, Ohio, USA
gupta@bgsu.edu

D. G. Kabe
5971 Greensboro Dr., Mississauga
Canada
pkabe@rogers.com

Abstract

Young, Scariano, and Hallum (2005) study two univariate linear regression models, and establish a certain condition for the variances estimates to be equal; however, the condition is in error. They then apply this incorrect condition to some unknown completely randomized design model. The present paper records the correct condition for the equality of the context variances, and points out that this condition cannot be applied to experimental design models.

Keywords: Equality of regression parametric estimates; Equality of their covariances; Randomized Blocks design; Full rank generalized inverse; Generalized inverse; Moore-Penrose inverse.

1. Introduction

Given the linear regression model

$$y = X\beta + e; \quad e \sim N(0, \sigma^2 I), \quad (1)$$

where y is an n component (column) vectors, $X_{n \times q}$ is of rank $q < n$, and β has q components. Changing $V(y)$ to $A\sigma^2$, the two estimates of β are

$$\hat{\beta}_1 = (X'X)^{-1}X'y, \quad \hat{\beta}_2 = (X'A^{-1}X)^{-1}X'A^{-1}y. \quad (2)$$

The two estimates of σ^2 are

$$\hat{\sigma}_1^2 = y'(I - X(X'X)^{-1}X')y, \quad \hat{\sigma}_2^2 = y'(A^{-1} - A^{-1}X(X'A^{-1}X)^{-1}X'A^{-1})y, \quad (3)$$

and they cannot be equal unless $A = I$.

Now Young, Scariano, and Hallum (2005, p. 627, (3)), (YSH), establish the variances equality condition as

$$\text{rank}((I - X(X'X)^{-1}X')A) = \text{rank}(I - X(X'X)^{-1}X'). \quad (4)$$

However, we are at a loss to understand what exactly the expression in the left hand side brackets represents. Further (4) does not imply (3). Thus a condition for equality, if any exists, other than that of (3), must be examined. We now show that the equality of regression parameter estimates does not imply the equality of the variances.

The next section establishes a known condition, Rao and Toutenberg (1999, pp. 106-108), for regression parameter estimates equality. In section 3 we give two examples to show that there is no loss of generality in linear regression theory to using full rank matrices. Finally in section 4, it is shown that less than full rank matrices are not allowed in design of experiments.

2. Regression Parameters Estimates Equality

The minimum value problem

$$\min y' Ay, \text{ subject to } X'y = g, X(n \times q), \quad (5)$$

where $n \times n$ A is positive definite symmetric, yields the solution

$$\hat{y} = A^{-1} X(X'A^{-1}X)^{-1} g, \quad \hat{y}' A \hat{y} = g'(X'A^{-1}X)^{-1} g. \quad (6)$$

A solution representation to $X'y = g$ is

$$y = A^{-1} X(X'A^{-1}X)^{-1} X'y + C'(CAC')^{-1} CAy, X'C' = 0, \quad (7)$$

where $(n-q) \times n$ C is orthogonal to X . Now (7) implies that

$$A = X(X'A^{-1}X)^{-1} X' + AC'(CAC')^{-1} CA. \quad (8)$$

Note that (6) and (8) do not hold if A is deficient in rank.

Thus if $X'AC = 0$, then first multiplying (8) from the right by A^{-1} and then from the left by X' , we find that

$$(X'X)^{-1} X'y = (X'A^{-1}X)^{-1} X'A^{-1}y, \quad (9)$$

$$\text{i.e. (least squares)} \hat{\beta} = (\text{covariance structure}) \hat{\beta}. \quad (10)$$

It also follows from (7) that in this case

$$A = X(X'A^{-1}X)^{-1} X', \text{ if } CA = 0. \quad (11)$$

However, the equality (9) does not imply (3).

Moreover (YSH) (2005, p. 627, (3)), i. e. ,

$$(I - XX^+)A(I - XX^+) = (I - XX^+), \quad (12)$$

is wrong. If A is less than full rank matrix, then it is simply the expected value

$$E((I - XX^+)yy'(I - XX^+)) = \sigma^2(I - XX^+)A(I - XX^+), \quad (13)$$

and unless A is a full rank matrix, it cannot be the covariance matrix of $(I - XX^+)y$ as they claim. They do not supply a mathematical proof that this is actually the covariance matrix of $(I - XX^+)y$.

If the covariance matrix of y is deficient in rank, then use the full rank generalized inverse as follows

$$\min(y - X\beta)'A(y - X\beta), \text{ subject to } \beta, \text{ or with respect to } \beta. \quad (14)$$

Write (14) as

$$\min(y - X\beta)'(A + T'T)(y - X\beta), \text{ subject to } TA(y - X\beta) = 0, \quad (15)$$

and also subject to $E(T(y - X\hat{\beta})) = 0$, where T is such that $(A + T'T)$ is a full rank

matrix. Thus e. g. assume $X = J(n \times 1)$ vector of unities, $\beta = \mu$, $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Then our $t' = (0, 0, 1)$, $t'A = 0$, and the problem is

$$\min(y - J\mu)'(A + tt')(y - J\mu), \text{ subject to } \mu, \quad (16)$$

the solution is $y_3 = \hat{\mu} = (y_1 + y_2) / 2$.

Thus there is absolutely no loss of generality in assuming A to be a full rank matrix. Perhaps, (YSH) (2005) should have been careful to avoid the complications that result with less than full rank matrices.

(YSH) (2005, p. 627, equation(3)) should be

$$[A^{-1} - (I - XX^+)((I - XX^+)A(I - XX^+))^{-1}(I - XX^+)] = (I - XX^+). \quad (17)$$

In the next section, we illustrate two examples to show that there is no loss of generality in linear regression theory to using full rank matrices. In section 4, we show that less than full rank matrices are not at all allowed in Experimental Designs Theory.

3. Some Examples

Let the regression normal equation for the model (1) be

$$X'y = (X'X)\hat{\beta}. \quad (18)$$

Now if $a'\beta$ is estimable and has an unbiased estimator, then any solution of (18), will yield the same $a'\beta$. For example,

$$\begin{bmatrix} 1 & 2 & 4 & 1 \\ 2 & 0 & 4 & -2 \\ 1 & 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 22 & 16 & 15 \\ 16 & 24 & 14 \\ 15 & 14 & 16 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}, \quad (19)$$

where note that $X'X$ is of rank two, and hence set $\beta_1 = 0$, and let a' be obtained by adding first and third columns of X' , $a' = (5, 6, 4)$. The normal equations are

$$\begin{bmatrix} 1 & 2 & 4 & 1 \\ 2 & 0 & 4 & -2 \\ 1 & 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 0 & 16 & 15 \\ 0 & 24 & 14 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}, \quad (20)$$

i. e.

$$\begin{aligned} T_1 &= y_1 + 2y_2 + 4y_3 + y_4 = 16\beta_2 + 15\beta_3, \\ T_2 &= 2y_1 + 4y_2 - 2y_4 = 24\beta_2 + 14\beta_3, \end{aligned} \quad (21)$$

and hence

$$\hat{\beta}_2 = (16T_2 - 14T_1) / 136, \hat{\beta}_3 = (24T_1 - 16T_2) / 136, \quad (22)$$

$$a'\hat{\beta} = (6T_1 + 13T_2) / 68 = (32y_1 + 12y_2 + 76y_3 - 20y_4) / 68. \quad (23)$$

Thus (19) simply means

$$\begin{bmatrix} 1 & 2 & 4 & 1 \\ 2 & 0 & 4 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 16 & 15 \\ 24 & 14 \end{bmatrix} \begin{bmatrix} \beta_2 \\ \beta_3 \end{bmatrix}, \quad (24)$$

i. e. , we have the "reduced" model

$$y = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 4 & 4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} \beta_2 \\ \beta_3 \end{bmatrix} + e, \quad e \sim N(0, \sigma^2 I). \quad (25)$$

However, now (25) is a full rank model. Thus there is no loss of generality in assuming full rank models. We hope YSH (2005) might follow this suggestion of ours.

Now such a solution as in (20) is not at all allowed in Design and Analysis of Experiments.

Let now

$$y = A\theta + e, \quad e \sim N(0, \sigma^2 I) \quad (26)$$

be a one way classification model, and the normal equations are

$$\hat{\theta} = \begin{bmatrix} 0 & 0 & 0 \\ A'A & 0 & JJ' \\ 0 & 0 & JJ' \end{bmatrix}^{-1} \begin{bmatrix} A'y \\ J'y \end{bmatrix}, \quad J'\beta = 0. \quad (27)$$

However, A is of rank, say two, $\theta' = (\mu, \alpha_1, \alpha_2)$, and hence

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \end{bmatrix} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad (28)$$

where we have set the condition that $J'\alpha = \alpha_1 + \alpha_2 = 0$, and modified $A'A$ in (27) as

$$\hat{\theta} = \begin{bmatrix} & 0 & 0 & 0 \\ & 0 & 1 & 1 \\ & 0 & 1 & 1 \\ A'A & + & & \end{bmatrix}^{-1} A'y. \quad (29)$$

Hence, we have the normal equation (28). The solutions are

$$\begin{aligned} y_{..} = T_1 &= y_{11} + y_{12} + y_{21} + y_{22} = 4\mu + 2\alpha_1 + 2\alpha_2, \\ y_{1.} = T_2 &= y_{11} + y_{12} = 2\mu + 2\alpha_1, \\ y_{2.} = T_3 &= y_{21} + y_{22} = 2\mu + 2\alpha_2, \end{aligned} \quad (30)$$

giving us

$$\begin{aligned} \hat{\mu} &= \frac{T_1}{4} = \frac{y_{..}}{4} = \bar{y}_{..}, \\ \hat{\alpha}_1 &= \frac{y_{1.}}{2} - \bar{y}_{..} = \bar{y}_{1.} - \bar{y}_{..}, \\ \hat{\alpha}_2 &= \frac{y_{2.}}{2} - \bar{y}_{..} = \bar{y}_{2.} - \bar{y}_{..}. \end{aligned} \quad (31)$$

We want solutions in exactly this format, no other format will do. Thus one must use full rank generalized inverse. This full rank generalized inverse is always obtained in Experimental Designs theory, by placing, if necessary, a linear restriction on one or more of a set of parameters, as in (27).

4. Randomized Blocks Design

The following analysis of randomized blocks design or a two way classification design is adapted from Chakravarti (1962). This is an orthogonal design.

Here pq observations are arranged in p rows and q columns, rows are A classification (A_1, \dots, A_p) , and columns are B classification (B_1, \dots, B_q) .

| A \ B | B | | | | | | Totals |
|----------|----------|----------|---------|----------|---------|----------|----------|
| | B_1 | B_2 | \dots | B_j | \dots | B_q | |
| A_1 | y_{11} | y_{12} | \dots | y_{1j} | \dots | y_{1q} | $y_{1.}$ |
| \vdots | \dots | \dots | \dots | \dots | \dots | \dots | \vdots |
| A_i | y_{i1} | y_{i2} | \dots | y_{ij} | \dots | y_{iq} | $y_{i.}$ |
| \vdots | \dots | \dots | \dots | \dots | \dots | \dots | \vdots |
| A_p | y_{p1} | y_{p2} | \dots | y_{pj} | \dots | y_{pq} | $y_{p.}$ |
| Totals | $y_{.1}$ | $y_{.2}$ | \dots | $y_{.j}$ | \dots | $y_{.q}$ | $y_{..}$ |

Let $\alpha_1, \dots, \alpha_p$ be block effects parameters, and β_1, \dots, β_q the treatment effects parameters. Then the model is

$$y_{ij} = \alpha_i + \beta_j + e_{ij}, \quad e_{ij} \sim N(0, \sigma^2 I).$$

The two hypotheses to be tested are

$$H_1: \alpha_1 = \alpha_2 = \dots = \alpha_p = \alpha \text{ (unspecified)},$$

$$H_2: \beta_1 = \beta_2 = \dots = \beta_q = \beta \text{ (unspecified)}.$$

We now set

$$y' = (y_{11}, y_{12}, \dots, y_{1q}; y_{21}, \dots, y_{2q}; \dots; y_{p1}, \dots, y_{pq}),$$

$$\theta' = (\alpha_1, \alpha_2, \dots, \alpha_p; \beta_1, \beta_2, \dots, \beta_q),$$

$$E(y) = A\theta, \quad A = \begin{bmatrix} J & 0 & \dots & 0 & I \\ 0 & J & \dots & 0 & I \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & J & I \end{bmatrix},$$

where J is a column vector of q unities, and I is $q \times q$ identity matrix. The rank of A is $(p+q-1)$, and hence we set $\beta_{\cdot} = J'\beta = 0$, and write the normal equations as

$$\begin{bmatrix} 0 & 0 & 0 \\ A'A & + & 0 & JJ' \end{bmatrix} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = A'y,$$

and the solutions are the same as those obtained by minimizing

$$L = \sum_{i=1}^p \sum_{j=1}^q (y_{ij} - \alpha_i - \beta_j)^2.$$

w. r. t. α and β .

The normal equations are

$$y_{i.} = q\alpha_i + \beta, \quad y_{.j} = \alpha + p\beta_j,$$

and the solutions are

$$\hat{\alpha}_i = \bar{y}_{i.} = y_{i.} / q, \quad \hat{\alpha} = y_{..} / q = \bar{y}_{..},$$

$$\hat{\beta}_j = (y_{.j} - y_{..} / q) / p = \bar{y}_{.j} - \bar{y}_{..}.$$

We get

$$\begin{aligned} \min L &= \text{Error SS} \\ &= \sum_{i=1}^p \sum_{j=1}^q (y_{ij} - \hat{\alpha}_i - \hat{\beta}_j)^2 \\ &= \sum_{i=1}^p \sum_{j=1}^q (y_{ij} - \hat{\alpha}_i - \hat{\beta}_j) y_{ij} \\ &= \left(\sum_{i=1}^p \sum_{j=1}^q y_{ij}^2 - \frac{y_{..}^2}{pq} \right) - \left(\sum_i \frac{y_{i.}^2}{q} - \frac{y_{..}^2}{pq} \right) - \left(\sum_j \frac{y_{.j}^2}{p} - \frac{y_{..}^2}{pq} \right). \end{aligned}$$

ANOVA Table

| Source | ss | d. f. | mss | F-tests |
|------------|--|------------------|-------|-----------|
| Blocks | $\sum_i \frac{y_{i.}^2}{q} - \frac{y_{..}^2}{pq}$ | $p - 1$ | V_A | V_A/V_E |
| Treatments | $\sum_j \frac{y_{.j}^2}{p} - \frac{y_{..}^2}{pq}$ | $q - 1$ | V_B | V_B/V_E |
| Error | substraction | $(p - 1)(q - 1)$ | V_E | |
| Totals | $\sum_{i=1}^p \sum_{j=1}^q y_{ij}^2 - \frac{y_{..}^2}{pq}$ | $pq - 1$ | | |

$$V(\hat{\alpha}_i - \hat{\alpha}_j) = 2\sigma^2 / q, \quad V(\hat{\beta}_i - \hat{\beta}_j) = 2\sigma^2 / p,$$

$$t = \frac{\hat{\alpha}_i - \hat{\alpha}_j}{\sqrt{2V_E / q}}, \text{ with } (p-1)(q-1) \text{ d.f.},$$

$$t = \frac{\hat{\beta}_i - \hat{\beta}_j}{\sqrt{2V_E / p}}, \text{ with } (p-1)(q-1) \text{ d.f.}.$$

It seems (YSH) (2005) are not acquainted with Design and Analysis of Experiments. They mention (2005, p. 628, equation (6)) a one way classification design, which has an arbitrary covariance matrix, which is deficient in rank. There is no such design. It would be impossible to analyze it. All experimental designs have identity covariance matrices.

References

1. Chakravarti, M. C. (1962). "Design and Analysis of Experiments." *Asia Publishing House*, Bombay, India.
2. Rao, C. R. , and Toutenberg, H. (1999). "Linear Models." *Springer-Verlag*, New York.
3. Young, Dean M., Scariano, Stephen M., and Hallum Cecil R. (2005). "Estimation-Equivalent Covariance Structures for the Least Squares and Unique Estimators of the Linear Model Variance." *Commun. Statist. Theor. Math.* 34, 625-629.