ESTIMATING VARIANCE COMPONENTS IN COVARIANCE MODELS

BU-403-M by January, 1972

T. D. Mount and S. R. Searle
Department of Agricultural Economics and Biometrics Unit
Cornell University

Abstract

Variance component estimators are derived for the 1- and 2-way classification models in the presence of covariates.

Paper No. BU-403-M in the Biometrics Unit Mimeograph Series, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14850.

ESTIMATING VARIANCE COMPONENTS IN COVARIANCE MODELS

BU-403-M

by

T. D. Mount and S. R. Searle
Department of Agricultural Economics and Biometrics Unit
Cornell University

Summary

Variance component estimators are derived for the 1- and 2-way classification models in the presence of covariates.

1. Introduction

Methods for estimating variance components in linear models containing random effects are well-known (e.g., [1, 2, 3, 4, 5, 9, and 10]). Although at least one method provides opportunity for handling unbalanced data with covariates in the model, no specific results appear to be available for such data. We therefore derive some such results for the 1-way and 2-way classifications with covariates, at the same time commenting on alternative procedures that have been suggested for these models.

We begin with a general form of the mixed effects linear model,

$$\underline{y} = \underline{X}\underline{\beta} + \underline{Z}\underline{u} + \underline{e} \tag{1}$$

where

- $\underline{y}$ is a vector of N observations on a dependent variable;
- $\underline{X}$ is an $N \times (c + 1)$ matrix of known values;
- $\underline{\beta}$ is a vector of $c + 1$ unknown parameters;
- $\underline{u}$ is a vector of p unknown random effects;
- $\underline{Z}$ is an $N \times p$ matrix of known values; and
- $\underline{e}$ is a vector of N unobserved random residuals.

Paper No. BU-403-M in the Biometrics Unit Mimeograph Series, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14850.

In general we assume that $\underline{u}$ and $\underline{e}$ have zero means, variance-covariance matrices $\underline{D}$ and $\underline{R}\sigma_e^2$, respectively, and that the covariance matrix between them is null. That is, for E representing expectation over repeated sampling

$$E(\underline{u}) = \underline{0}, \quad E(\underline{u}\underline{u}') = \underline{D}, \quad E(\underline{u}\underline{e}') = \underline{0}, \quad E(\underline{e}) = \underline{0}, \quad E(\underline{e}\underline{e}') = \sigma_e^2\underline{R}$$

and hence

$$E(\underline{y}) = \underline{X}\underline{\beta}$$

and

$$\text{var}(\underline{y}) = \underline{Z}\underline{D}\underline{Z}' + \sigma_e^2\underline{R} = \underline{V}, \text{ say.} \tag{2}$$

This, with slight changes in notation, is the model considered in [4, 5, 6, 9, 10, and 11].

Models with both covariates and random effects factors have considerable application in economics. For example, the econometric models considered in [7, 8, and 13] can be specified in terms of (1) by defining

$$\underline{\beta} = \begin{bmatrix} \mu \\ \underline{\beta}_1 \end{bmatrix} \quad \text{with } \underline{X} = [\underline{1} \quad \underline{X}_1] \quad \text{and} \quad \underline{u} = \begin{bmatrix} \underline{u}_1 \\ \underline{u}_2 \end{bmatrix} \quad \text{with } \underline{Z} = [\underline{Z}_1 \quad \underline{Z}_2],$$

where

- μ is a general mean;
- $\underline{\beta}_1$ is a vector of c coefficients corresponding to c covariates;
- $\underline{1}$ is a vector of N unities;
- $\underline{X}_1$ is an $N \times c$ matrix of the N observations on each of the c covariates corresponding to the N observed y's;
- $\underline{u}_1$ is a vector of a random effects $u_{11} \cdots u_{1a}$ representing the a cross section effects in the data;

$\underset{-}{u}_2$ is a vector of b random effects $u_{21} \cdots u_{2b}$ representing the b time effects in the data; and

$\underset{-}{Z}_1$, $\underset{-}{Z}_2$ are the $N \times a$ and $N \times b$ design matrices corresponding to $\underset{-}{u}_1$ and $\underset{-}{u}_2$.

Thus for a particular observation y_{ij} the model is, for $i = 1, 2, \cdots, a$ and $j = 1, 2, \cdots, b$,

$$y_{ij} = \mu + \beta_1 x_{1ij} + \cdots + \beta_c x_{cij} + u_{1i} + u_{2j} + e_{ij} \tag{3}$$

the vector form of this being

$$\underset{-}{y} = \mu\underset{-}{1} + \underset{-}{X}_1\underset{-}{\beta}_1 + \underset{-}{Z}_1\underset{-}{u}_1 + \underset{-}{Z}_2\underset{-}{u}_2 + \underset{-}{e} , \tag{4}$$

which can also be written as

$$\underline{y} = \underline{X}\underline{\beta} + \underline{Z}_1\underline{u}_1 + \underline{Z}_2\underline{u}_2 + \underline{e}. \tag{5}$$

On assuming that the variance-covariance matrix of the elements of $\underline{u}_1$ is $\sigma_1^2\underline{I}_a$ and that of the elements of $\underline{u}_2$ is $\sigma_2^2\underline{I}_b$, as well as assuming that the covariance matrix of $\underline{u}_1$ with $\underline{u}_2$ is null, we have

$$\underline{D} = \text{var}(\underline{u}) = E(\underline{u}\underline{u}') = \begin{bmatrix} \sigma_1^2\underline{I}_a & \underline{0} \\ \underline{0} & \sigma_2^2\underline{I}_b \end{bmatrix}$$

where $\underline{I}_a$ and $\underline{I}_b$ are identity matrices of order a and b. Furthermore, as is customary, we assume that $\underline{R}$ of (2) is $\underline{R} = \underline{I}_N$ which therefore gives $\underline{V}$ of (2) as

$$\begin{aligned} \underline{V} &= \underline{Z}\underline{D}\underline{Z}' + \sigma_e^2\underline{R} \\ &= (\sigma_1^2\underline{Z}_1\underline{Z}_1' + \sigma_2^2\underline{Z}_2\underline{Z}_2' + \sigma_e^2\underline{I}). \end{aligned} \tag{6}$$

We note in passing that for the case of 1 observation in each of the ab cells of the model (4), $\underline{Z}_1\underline{Z}_1'$ and $\underline{Z}_2\underline{Z}_2'$ are, respectively, $\underline{A}$ and $\underline{B}$ of equation (5) of [13].

2. Estimating Slopes when Variances are Known

When in (2) $\underline{D}$ and $\underline{R}$ are known, the best linear unbiased estimator (BLUE) of $\underline{\beta}$ in the model (1) is, as usual, the generalized least squares estimator

$$\hat{\underline{\beta}} = (\underline{X}'\underline{V}^{-1}\underline{X})^{-1}\underline{X}'\underline{V}^{-1}\underline{y} \; . \tag{7}$$

More particularly, for the specific model represented equivalently by (3) and (4) the BLUE of $\underline{\beta}$ is, when σ_1^2/σ_e^2 and σ_2^2/σ_e^2 are known, (7) using (6) for $\underline{V}$. This is the estimator discussed in [13]. Although that paper very ingenicusly gives the form of $\underline{V}^{-1}$ [for $\underline{V}$ of equation (6) with one observation per cell], it is pointed out in [6] that inversion of this N × N matrix can be avoided by a technique given in [4, 5, 10, and 11]. This amounts to solving the equations

$$\begin{bmatrix} \underline{X}'\underline{R}^{-1}\underline{X} & \underline{X}'\underline{R}^{-1}\underline{Z} \\ \underline{Z}'\underline{R}^{-1}\underline{X} & \underline{Z}'\underline{R}^{-1}\underline{Z} + \underline{D}^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\underline{\beta}} \\ \tilde{\underline{u}} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{R}^{-1}\underline{y} \\ \underline{Z}'\underline{R}^{-1}\underline{y} \end{bmatrix} . \tag{8}$$

The equivalence of $\tilde{\underline{\beta}}$ of (8) to $\hat{\underline{\beta}}$ of (7) is given in [4] and [5] and also in [11, p. 460], and the specific form of (8) for the model (4) with $\underline{V}$ as in (6) is given in [6].

3. Estimating Unknown Variances

Seldom do we know the variance components involved in $\underline{D}$ of (2), e.g., the variances σ_1^2 and σ_2^2 in (6). Before (7) or (8), or their simpler forms for the model (4) can be used, the required variance components must be estimated. Replacing the variance components in $\underline{V}$ by their estimators yields a consistent estimator $\tilde{\underline{V}}$ of $\underline{V}$. Then replacing $\underline{V}$ by $\tilde{\underline{V}}$ in (7) yields a consistent estimator $(\underline{X}'\tilde{\underline{V}}^{-1}\underline{X})^{-1}\underline{X}'\tilde{\underline{V}}^{-1}\underline{y}$ of $\underline{\beta}$.

To estimate $\underline{\beta}$ when the variance components are unknown, the first problem is therefore to estimate the variance components. This is easy when there are no covariates in the model and the data are balanced (meaning that each of the submost cells contains the same number of observations). In this case we obtain unbiased estimators of the variance components from the computed mean squares of a standard analysis of variance (see for example [11, Chapter 9] and many other statistics texts that deal with variance components). When there are no covariates and the data are unbalanced (unequal numbers of observations in the cells), estimation of the variance components is not so easy as with balanced data, but many methods are available for this case (e.g., [10] and [11, Chapters 10 and 11]). For each of these methods very few distributional properties are available for the resulting estimators, the only property which most of them possess being unbiasedness. When covariates are present, as in (4) above, and the data are balanced---the easy case when there are no covariates---it is tempting to use procedures similar to that easy case but, unfortunately, as is shown below, they do not yield unbiased estimators.

3.1. Ordinary Least Squares for the Fixed Effects

Were there no random effects in the model (1), the normal equations for the fixed effects, using ordinary least squares, would be $\underline{X}'\underline{X}\underline{\beta}^* = \underline{X}'\underline{y}$. To estimate the required variance components in $\underline{D}$, Wallace and Hussain [13, Section 5.A] suggest correcting the data according to this $\underline{\beta}^*$, to obtain $\underline{z} = \underline{y} - \underline{X}\underline{\beta}^*$, and they then calculate an analysis of variance of the z's assuming no fixed effects in the model. For example, in the case of the model (5) they derive $\underline{\beta}^*$ as $(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}$ and then calculate an analysis of variance of the z's based effectively on a model $z_{ij} = \mu_z + u_{1i} + u_{2j} + e_{ij}$, assuming throughout that the covariates no longer affect the estimation of the variance components. For example, at their equation (40) they estimate σ_e^2 by $\tilde{\sigma}_e^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}(z_{ij} - \bar{z}_{i.} - \bar{z}_{.j} + \bar{z}_{..})^2/(a - 1)(b - 1)$. It is worth noting that this estimator $\tilde{\sigma}^2$ is not unbiased because the model for $\underline{z}$ still includes $\underline{X}$, since

$$\begin{aligned}\underline{z} = \underline{y} - \underline{X}\underline{\beta}^* &= [\underline{I} - \underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}']\underline{y} \\ &= [\underline{I} - \underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}](\underline{X}\underline{\beta} + \underline{Z}\underline{u} + \underline{e}) \\ &= [\underline{I} - \underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'](\underline{Z}\underline{u} + \underline{e}).\end{aligned}$$

Clearly, then, expected values of mean squares in a traditional analysis of variance of the z's involves $\underline{X}$ and will not yield unbiased variance components estimators.

Estimating variance components in mixed models of the form (1) by means of correcting the data for the fixed effects through using $\underline{y} - \underline{X}\tilde{\underline{\beta}}$ for some $\tilde{\underline{\beta}}$ was first considered by Henderson [3]. One form of the method which has come to be known as Henderson's Method II was first given in [3] and has more recently been described in great detail by Searle [9]. Shown there are several forms of the method and their deficiencies, the main one being that the method is not uniquely defined. This is so because any form $\tilde{\underline{\beta}} = \underline{L}\underline{y}$ can be used so long as $\underline{L}$ is a generalized inverse of $\underline{X}$, meaning that $\underline{L}$ satisfies $\underline{XLX} = \underline{X}$, and hence can be one of an infinite number of matrices. In this connection, the specific procedure suggested by Wallace and Hussain is that described as Method 4 in [9, p. 765].

3.2. Using Least Squares Solutions for Random Effects

When the u-vector in (1) is assumed to be a vector of fixed effects, ordinary least squares solutions for $\underline{\beta}$ and $\underline{u}$ are the solutions $\underline{\beta}^o$ and $\underline{u}^o$ to

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} \end{bmatrix} \begin{bmatrix} \underline{\beta}^o \\ \underline{u}^o \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \end{bmatrix} \tag{9}$$

Nerlove [8] has suggested using $\underline{u}^o$ to estimate the variance components. For example, with the model (4) with the u_2-effects omitted, he suggests [at equation 2.15] using $\tilde{\sigma}_1^2 = \sum_{i=1}^{a} (u_{1i}^o - \bar{u}_{1.}^o)^2/a$ as an estimator of σ_1^2. In Nerlove's model a lagged dependent variable is present as one of the covariates, but even if this is excluded from his model $\tilde{\sigma}_1^2$ is not an unbiased estimator of σ_1^2. Consider (9): its solution is [e.g., 11, pp. 341-342]

$$\underline{\beta}^o = (\underline{X}'\underline{X})^{-1}\underline{X}'(\underline{y} - \underline{Z}\underline{u}^o) \tag{10}$$

and

$$\underline{u}^o = (\underline{Z}'\underline{P}\underline{Z})^{-1}\underline{Z}'\underline{P}\underline{y} \tag{11}$$

where

$$\underline{P} = \underline{I} - \underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'$$

which is symmetric, and idempotent. Substituting (1) into (11) gives

$$\begin{aligned}\underline{u}^o &= (\underline{Z}'\underline{P}\underline{Z})^{-1}\underline{Z}'\underline{P}(\underline{X}\beta + \underline{Z}\underline{u} + \underline{e}) \\ &= \underline{u} + (\underline{Z}'\underline{P}\underline{Z})^{-1}\underline{Z}'\underline{P}\underline{e} \; . \end{aligned} \tag{12}$$

Sums of squares of elements of $\underline{u}^o$, as used by Nerlove [8], therefore have expected values that are not functions of the variance components alone but involve the covariates through the occurrence of $\underline{X}$ in $\underline{P}$ in (12).

The special case considered by Nerlove [8] is (5) without the u_2-classification. In this case there are, in each of the a levels of the u_1-factor, b observations on the y-variable and on each of the c cofactors. Furthermore,

$$\underline{Z}_1 = \sum_{i=1}^{a}{}^{+} \underline{1}_b, \quad \underline{Z}_1'\underline{Z}_1 = b\underline{I}_a, \quad \text{and} \quad \underline{Z}_2 = 0 \tag{13}$$

where $\underline{1}_b$ is a vector of b unities and Σ^+ denotes the operation of the direct sum of matrices, and $\underline{u}^o$ of (12) becomes

$$\underline{u}_1^o = \underline{u}_1 + \underline{S}\underline{e} \quad \text{where} \quad \underline{S} = (\underline{Z}_1'\underline{P}\underline{Z}_1)^{-1}\underline{Z}_1'\underline{P}. \tag{14}$$

The estimator of σ_1^2 suggested by Nerlove [8] is then

$$\tilde{\sigma}_1^2 = (1/a) \sum_{i=1}^{a} (u_{1i}^o - \bar{u}_{1.}^o)^2 = \underline{u}_1^o{}'\underline{A}\underline{u}_1^o$$

where

$$\underset{\sim}{A} = (1/a^2)(a\mathbf{I}_a - \mathbf{1}_a\mathbf{1}_a') \ . \tag{15}$$

Therefore, on using the general rule for the expected value of a quadratic form [e.g., 11, p. 55], we have

$$E(\tilde{\sigma}_1^2) = \mathrm{tr}[\mathbf{A}\ \mathrm{var}(\mathbf{u}_1 + \mathbf{S}\mathbf{e})]$$

and, as shown in the appendix at equation (A11), this reduces to

$$E(\tilde{\sigma}_1^2) = (a - 1)(b\sigma_1^2 + \sigma_e^2)/ab + a^{-1}\mathrm{tr}(\mathbf{W}^{-1}\mathbf{B})\sigma_e^2$$

where $\mathbf{W}$ and $\mathbf{B}$ are, respectively, the $c \times c$ matrices of within-level and between-level sums of squares and products of the covariates, as shown in (A8) and (A10), respectively. Thus we have shown that Nerlove's $\tilde{\sigma}_1^2$ is not an unbiased estimator of σ_1^2.

3.3. Fitting Different Models

The symbol $R(\boldsymbol{\beta}_1)$ is reasonably standard for the reduction in sum of squares due to fitting the model $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{e}$, its value being $R(\boldsymbol{\beta}_1) = \mathbf{y}'\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-}\ \mathbf{X}_1'\mathbf{y}$. Similarly, for fitting

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{e} \tag{16}$$

the reduction in sum of squares is

$$R(\beta_1,\beta_2) = \mathbf{y}'[\mathbf{X}_1 \quad \mathbf{X}_2]\begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 \end{bmatrix}^{-}\begin{bmatrix} \mathbf{X}_1' \\ \mathbf{X}_2' \end{bmatrix}\mathbf{y} \ .$$

Furthermore, in fitting (16), the reduction in sum of squares due to fitting $\boldsymbol{\beta}_2$ over and above fitting $\boldsymbol{\beta}_1$ is

$$R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1) = R(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2) - R(\boldsymbol{\beta}_1).$$

As given in [9, 10, 11], the expected value of $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$ in the model (16) is

$$E\left[R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)\right] = \text{tr}\left\{\underset{\sim}{X}_2'[\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-}\underset{\sim}{X}_1']\underset{\sim}{X}_2E(\underset{\sim}{\beta}_2\underset{\sim}{\beta}_2')\right\} + \sigma_e^2\left\{r[\underset{\sim}{X}_1 \quad \underset{\sim}{X}_2] - r(\underset{\sim}{X}_1)\right\}. \quad (17)$$

where $r(\underset{\sim}{X})$ is the rank of $\underset{\sim}{X}$. This is quite a general result which has several ramifications, discussed in the afore-mentioned references. The most important is that its right-hand side does not involve $\underset{\sim}{\beta}_1$. Hence, if a mixed model such as (1) is written in any variety of ways in the form of (16), but always with the fixed effects included in $\underset{\sim}{\beta}_1$, then (17) provides the expectation of corresponding values of $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$, unaffected by the fixed effects. From these and the residual sum of squares due to fitting (16), which provides an unbiased estimator of σ_e^2, we can unbiasedly estimate variance components in a mixed model. This is the procedure of Henderson's Method III, first described in [3] and more recently [9, 10, 11] presented in matrix notation. We now show how (17) can be used for models that include covariates.

4. The 1-Way Classification

Similar to (4) we take the model for the 1-way classification as being

$$\underset{\sim}{y} = \mu\underset{\sim}{1} + \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{Z}_1\underset{\sim}{u}_1 + \underset{\sim}{e} \quad (18)$$

corresponding to

$$y_{ij} = \mu + \beta_1 x_{1ij} + \cdots + \beta_c x_{cij} + u_{1i} + e_{ij} \; .$$

In (18) $\underset{\sim}{\beta}_1$ is the $1 \times c$ vector of elements $\beta_1, \beta_2, \cdots, \beta_c$; $\underset{\sim}{X}_1$ is the $N \times c$ matrix of covariates; and $\underset{\sim}{u}_1$ is the vector of a elements u_{1i}, $i = 1, 2, \cdots, a$, corresponding to the a levels of the random factor. For increased generality compared to the model considered previously we now let the number of observations in the i^{th} level of the random effect $\underset{\sim}{u}_1$ be n_i rather than b. Then

$$\underset{\sim}{Z}_1 = \sum_{i=1}^{a}{}^{+} \underset{\sim}{1}_{n_i} \; .$$

The variance components to be estimated are σ_1^2 corresponding to the random effects in $\mathbf{u}_1$ and σ_e^2. The latter is easily estimated as the residual mean square due to fitting (18). Since this is the standard covariance model for a 1-way classification (with unequal numbers of observations, n_i, in the subclasses however), $\hat{\sigma}_e^2$ is

$$\hat{\sigma}_e^2 = \Big(\sum_{i=1}^{a} \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a} n_i \bar{y}_{i.}^2 - \mathbf{w}'\mathbf{W}^{-1}\mathbf{w} \Big)/(N - a - c) \tag{19}$$

where $N = n_. = \sum_{i=1}^{a} n_i$ and $\mathbf{w}$ is the $c \times 1$ vector of within-group sums of products of the covariates with the y_{ij}'s, i.e.

$$\mathbf{w} = \Big\{ \sum_{i=1}^{a} \sum_{j=1}^{n_i} (x_{tij} - \bar{x}_{ti.})(y_{ij} - \bar{y}_{i.}) \Big\} \quad \text{for } t = 1, 2, \cdots, c;$$

and $\mathbf{W}$ is the $c \times c$ matrix of within-group sums of squares and products of the covariates given in (A8) of the Appendix.

Use is now made of (17) to derive $E\, R(\mathbf{u}_1|\mu,\boldsymbol{\beta}_1)$ for the model (18). It will be a linear function of σ_1^2 and σ_e^2, and together with $\hat{\sigma}_e^2$ will lead to an unbiased estimator of σ_1^2. Correspondence between the general model (16) and the model (18) for which we want $E\big[R(\mathbf{u}_1|\mu,\boldsymbol{\beta}_1)\big]$ is as follows:

General case, (16) and (17):	$\mathbf{X}_1$	$\mathbf{X}_2$	$\boldsymbol{\beta}_1$	$\boldsymbol{\beta}_2$
Model (18):	$[\mathbf{1}\ \ \mathbf{X}_1]$	$\mathbf{Z}_1$	$\begin{bmatrix} \mu \\ \boldsymbol{\beta}_1 \end{bmatrix}$	$\mathbf{u}_1$

Thus from (17)

$$E\big[R(\mathbf{u}_1|\mu,\boldsymbol{\beta}_1)\big] = \mathrm{tr}\Big\{\mathbf{Z}_1' \Big[\mathbf{I} - [\mathbf{1}\ \ \mathbf{X}_1] \begin{bmatrix} N & \mathbf{1}'\mathbf{X}_1 \\ \mathbf{X}_1'\mathbf{1} & \mathbf{X}_1'\mathbf{X}_1 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}' \\ \mathbf{X}_1' \end{bmatrix} \Big] \mathbf{Z}_1 E(\mathbf{u}_1\mathbf{u}_1')\Big\} \tag{20}$$

$$+ \sigma_e^2\big(r[\mathbf{1}\ \ \mathbf{X}_1\ \ \mathbf{Z}_1] - r[\mathbf{1}\ \ \mathbf{X}_1]\big) .$$

Obtaining the matrix inverse required here by the familiar result for the inverse of a partitioned matrix (see [11], page 84 for this example), and using $E(\underline{u}_1\underline{u}_1') = \sigma_1^2\underline{I}_a$, we find that

$$E\left[R(\underline{u}_1|\mu,\underline{\beta}_1)\right] = \sigma_1^2\ \mathrm{tr}(\underline{Z}_1'\underline{Z}_1)$$

$$- \sigma_1^2\ \mathrm{tr}\left\{\underline{Z}_1'[\underline{1}\quad \underline{X}_1]\left[\begin{bmatrix} N^{-1} & \underline{0}' \\ \underline{0} & \underline{0} \end{bmatrix} + \begin{bmatrix} -\bar{\underline{x}}' \\ \underline{I} \end{bmatrix} \underline{T}^{-1}[-\bar{\underline{x}}\quad \underline{I}]\right]\begin{bmatrix} \underline{1}' \\ \underline{X}_1' \end{bmatrix}\underline{Z}_1\right\}$$

$$+ \sigma_e^2(c + a - c - 1)$$

where $\bar{\underline{x}}$ is the $c \times 1$ vector of means of the covariates $\bar{x}_{t..}$, for $t = 1, \cdots, c$,

and $\underline{T}$ is the $c \times c$ matrix of total sums of squares and products of the covariates corrected for their means;

i.e., $\bar{\underline{x}} = \{\bar{x}_{t..}\}$ for $t = 1, \cdots, c$

and $\underline{T} = \left\{\sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{tij} - \bar{x}_{t..})(x_{t'ij} - \bar{x}_{t'..})\right\}$ for $t,t' = 1, 2, \cdots, c.$

Simplifying further gives

$$E\left[R(\underline{u}_1|\mu,\underline{\beta}_1)\right] = \sigma_1^2\left\{\mathrm{tr}(\underline{Z}_1'\underline{Z}_1) - N^{-1}\mathrm{tr}[(\underline{1}'\underline{Z}_1)'\underline{1}'\underline{Z}_1]\right.$$

$$\left. - \mathrm{tr}[\underline{Z}_1'(\underline{X}_1 - \underline{1}\bar{\underline{x}}')\underline{T}^{-1}(\underline{X}_1' - \bar{\underline{x}}\underline{1}')\underline{Z}_1]\right\} + (a - 1)\sigma_e^2 .$$

Now by the definition of $\underline{Z}_1$, $\underline{Z}_1'\underline{Z}_1$ is a diagonal matrix of the n_i's and $\underline{1}'\underline{Z}_1$ is a vector of the n_i's so that

$$\mathrm{tr}(\underline{Z}_1'\underline{Z}_1) = N = n_. \quad \text{and} \quad \mathrm{tr}(\underline{1}'\underline{Z}_1)'\underline{1}'\underline{Z}_1 = \sum_{i=1}^{a} n_i^2 .$$

Furthermore,

$$(\underset{\sim}{X}_1' - \bar{\underset{\sim}{x}}\underset{\sim}{1}')\underset{\sim}{Z}_1 = (\underset{\sim}{X}_1' - \bar{\underset{\sim}{x}}\underset{\sim}{1}') \sum_{i=1}^{a}{}^{+} \underset{\sim}{1}_{n_i}$$

$$= \left\{n_i(\bar{x}_{ti.} - \bar{x}_{t..})\right\} \quad \text{for } t = 1, 2, \cdots, c; \; i = 1, 2, \cdots, a,$$

$$= \underset{\sim}{U}_1 \quad \text{say,}$$

and so, using these results and the commutative property of matrix products under the trace operation,

$$E\left[R(\underset{\sim}{u}_1|\mu,\underset{\sim}{\beta}_1)\right] = \sigma_1^2\left[(N - \Sigma n_i^2/N) - \mathrm{tr}(\underset{\sim}{T}^{-1}\underset{\sim}{U}_1\underset{\sim}{U}_1')\right] + (a - 1)\sigma_e^2 \tag{21}$$

where

$$\underset{\sim}{U}_1\underset{\sim}{U}_1' = \left\{\sum_{i=1}^{a} n_i^2(\bar{x}_{ti.} - \bar{x}_{t..})(\bar{x}_{t'i.} - \bar{x}_{t'..})\right\} \quad \text{for } t,t' = 1, 2, \cdots, c.$$

Hence $\hat{\sigma}_e^2$ is given by (19) and

$$\hat{\sigma}_1^2 = \frac{R(\underset{\sim}{u}_1|\mu,\underset{\sim}{\beta}_1) - (a - 1)\hat{\sigma}_e^2}{N - \Sigma n_i^2/N - \mathrm{tr}(\underset{\sim}{T}^{-1}\underset{\sim}{U}_1\underset{\sim}{U}_1')} .$$

An alternative computing procedure for the coefficient of σ_1^2 in (21) can be derived from (20). Since

$$E(\underset{\sim}{u}_1\underset{\sim}{u}_1') = \sigma_1^2\underset{\sim}{I}_a$$

(20) simplifies to

$$E\left[R(\underset{\sim}{u}_1|\mu,\underset{\sim}{\beta}_1)\right] = \sigma_1^2 \,\mathrm{tr}\left\{\underset{\sim}{Z}_1'[\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}']\underset{\sim}{Z}_1\right\} + \sigma_e^2(a - 1)$$

where

$$\underset{\sim}{X} = [\underset{\sim}{1} \;\; \underset{\sim}{X}_1].$$

The coefficient of σ_1^2 in the above expression can be further simplified by noting that $\underline{M} = \underline{I} - \underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'$ is symmetric and idempotent to give

$$\begin{aligned} \mathrm{tr}\{\underline{Z}_1'\underline{M}\underline{Z}_1\} &= \mathrm{tr}\{(\underline{M}\underline{Z}_1)'(\underline{M}\underline{Z}_1)\} \\ &= \mathrm{tr}\{(\underline{Z}_1 - \hat{\underline{Z}}_1)'(\underline{Z}_1 - \hat{\underline{Z}}_1)\} \end{aligned}$$

where $\hat{\underline{Z}}_1 = \underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Z}_1$ is the ordinary least squares predicted value of $\underline{Z}_1$ derived by regressing $\underline{Z}_1$ on $\underline{X}$. Hence, the coefficient of σ_1^2 in (21) is in fact equal to the sum of the a sums of squares of the estimated residuals computed by regressing each column of $\underline{Z}_1$ on $\underline{X}$.

5. 2-Way Crossed Classification, 1 Observation Per Cell

We now use the model in (4),

$$\underline{y} = \mu\underline{1} + \underline{X}_1\underline{\beta}_1 + \underline{Z}_1\underline{u}_1 + \underline{Z}_2\underline{u}_2 + \underline{e} \, , \tag{22}$$

with

$$\underline{Z}_1 = \sum_{i=1}^{a}{}^{+} \underline{1}_b \quad \text{and} \quad \underline{Z}_2 = \underline{1}_a * \underline{I}_b$$

where $*$ is the operation of direct (Kronecker) multiplication. To estimate σ_e^2 we have

$$E\left[\underline{y}'\underline{y} - R(\mu,\underline{\beta}_1,\underline{u}_1,\underline{u}_2)\right] = \sigma_e^2\left[(a - 1)(b - 1) - c\right] . \tag{23}$$

The reductions in sums of squares that can be used for estimating the other variance components, σ_1^2 and σ_2^2 corresponding to the random effects represented by $\underline{u}_1$ and $\underline{u}_2$, are $R(\underline{u}_1,\underline{u}_2|\mu,\underline{\beta}_1)$, $R(\underline{u}_1|\mu,\underline{\beta}_1,\underline{u}_2)$, and $R(\underline{u}_2|\mu,\underline{\beta}_1,\underline{u}_1)$. Since μ and $\underline{\beta}_1$ contain all the fixed effects of the model, expected values of these 3 reductions contain, as a consequence of (17), no terms in those fixed effects and are linear functions of σ_1^2, σ_2^2, and σ_e^2, of σ_1^2 and σ_e^2, and of σ_2^2 and σ_e^2, respectively. Derivation of these expected values is shown in the appendix. The results follow.

From (A13) we have

$$E\left[R(\underline{u}_1,\underline{u}_2|\mu,\underline{\beta}_1)\right] = b\left[(a - 1) - b\ \mathrm{tr}(\underline{T}^{-1}\underline{B}_1)\right]\sigma_1^2$$

$$+ a\left[(b - 1) - a\ \mathrm{tr}(\underline{T}^{-1}\underline{B}_2)\right]\sigma_2^2 + (a + b - 2)\sigma_e^2, \qquad (24)$$

where, as before, $\underline{T}$ is the $c \times c$ matrix of total sums of squares and products of the covariates, corrected for their means; and $\underline{B}_1$ and $\underline{B}_2$ are $c \times c$ matrices of between-levels (of the u_1- and u_2-factors, respectively) sums of squares and products of the covariates, as shown in (A14) and (A15) of the appendix. From (A16) and (A18) we also have

$$E\left[R(\underline{u}_1|\mu,\underline{\beta}_1,\underline{u}_2)\right] = b\left[(a - 1) - b\ \mathrm{tr}(\underline{W}_2^{-1}\underline{B}_2)\right]\sigma_1^2 + (b - 1)\sigma_e^2 \qquad (25)$$

and

$$E\left[R(\underline{u}_2|\mu,\underline{\beta}_2,\underline{u}_1)\right] = a\left[(b - 1) - a\ \mathrm{tr}(\underline{W}_1^{-1}\underline{B}_1)\right]\sigma_2^2 + (a - 1)\sigma_e^2 \qquad (26)$$

where $\underline{B}_1$ and $\underline{B}_2$ are as above, and $\underline{W}_1$ and $\underline{W}_2$ are $c \times c$ matrices of within-levels (of the u_1- and u_2-factors, respectively) sums of squares and cross-products of the covariates, as shown in (A17) and (A19).

Equations (23) - (26) are linear in σ_1^2, σ_2^2, and σ_e^2. When their left-hand sides are replaced by calculated values of the corresponding reductions in sums of squares and their right-hand sides have σ_1^2, σ_2^2, and σ_e^2 replaced by estimators $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, and $\hat{\sigma}_e^2$, the resulting equations represent 4 equations in these three unbiased estimators. This presents a problem of over-identifiability, of having more equations than unknowns, a problem which is quite customary with unbalanced data without covariates, and to which there is at present no general solution. One possibility suggested in Searle [9, 10], that maintains unbiasedness of the estimators, is to obtain least squares estimators from the four equations. In matrix notation if $\underline{r}$ is the vector of reductions and if $E(\underline{r}) = \underline{A}\underline{\sigma}^2$ is their expected values, then

$$\hat{\underline{\sigma}}^2 = (\underline{A}'\underline{A})^{-1}\underline{A}'\underline{r} \tag{27}$$

is a "least squares" estimator of $\underline{\sigma}^2$. It should be noted that if there are no covariates, (24) is the sum of (25) and (26) and the variances are not over-identified. However, when covariates are present, this is not the case as $\underline{T} \neq \underline{W}_1 \neq \underline{W}_2$.

Using procedures similar to those described at the end of the preceding section for the 1-way classification, the coefficients of σ_1^2 and σ_2^2 in (24) - (26) can be computed in an alternative way. The alternative expressions for the coefficients of σ_1^2 and σ_2^2 in (24) are given in (A20) and (A21) of the appendix; and derivation of the alternative expressions for the coefficients of σ_1^2 in (25) and of σ_2^2 in (26) is equivalent to the procedure used in the 1-way classification. In all cases, the alternative expression is a sum of sums of squares of estimated residuals derived from a series of regressions; thus we regress

$$[\underset{\sim}{Z}_1 \;\; \underset{\sim}{Z}_2] \quad \text{on} \quad [\underset{\sim}{1} \;\; \underset{\sim}{X}_1] \quad \text{for } (24),$$

$$\underset{\sim}{Z}_1 \quad \text{on} \quad [\underset{\sim}{1} \;\; \underset{\sim}{X}_1 \;\; \underset{\sim}{Z}_2] \quad \text{for } (25), \tag{28}$$

and

$$\underset{\sim}{Z}_2 \quad \text{on} \quad [\underset{\sim}{1} \;\; \underset{\sim}{X}_1 \;\; \underset{\sim}{Z}_1] \quad \text{for } (26).$$

Thus in (24) the coefficient of σ_1^2 is

$$b\left[(a - 1) - b \, \mathrm{tr}(\underset{\sim}{T}^{-1}\underset{\sim}{B}_1)\right] = \sum_{i=1}^{a} \hat{\underset{\sim}{e}}'_{1i}\hat{\underset{\sim}{e}}_{1i} \tag{29}$$

where

$$\left[\hat{\underset{\sim}{e}}_{11} \; \hat{\underset{\sim}{e}}_{12} \cdots \hat{\underset{\sim}{e}}_{1a}\right] = \underset{\sim}{Z}_1 - (\underset{\sim}{1} \;\; \underset{\sim}{X}_1)\left[(\underset{\sim}{1} \;\; \underset{\sim}{X}_1)'(\underset{\sim}{1} \;\; \underset{\sim}{X}_1)\right]^{-1}(\underset{\sim}{1} \;\; \underset{\sim}{X}_1)'\underset{\sim}{Z}_1$$

is an $N \times a$ matrix of estimated residuals. Equivalently the coefficient of σ_2^2 in (24) is

$$a\left[(b - 1) - a \, \mathrm{tr}(\underset{\sim}{T}^{-1}\underset{\sim}{B}_2)\right] = \sum_{j=1}^{b} \hat{\underset{\sim}{e}}'_{2j}\hat{\underset{\sim}{e}}_{2j} \tag{30}$$

where

$$\left[\hat{\underset{\sim}{e}}_{21} \; \hat{\underset{\sim}{e}}_{22} \cdots \hat{\underset{\sim}{e}}_{2b}\right] = \underset{\sim}{Z}_2 - (\underset{\sim}{1} \;\; \underset{\sim}{X}_1)\left[(\underset{\sim}{1} \;\; \underset{\sim}{X}_1)'(\underset{\sim}{1} \;\; \underset{\sim}{X}_1)\right]^{-1}(\underset{\sim}{1} \;\; \underset{\sim}{X}_1)'\underset{\sim}{Z}_2$$

is an $N \times b$ matrix of estimated residuals. Similarly in (25) σ_1^2 has the coefficient

$$b\left[(a - 1) - b \, \mathrm{tr}(\underset{\sim}{W}_2^{-1}\underset{\sim}{B}_2)\right] = \sum_{i=1}^{a} \tilde{\underset{\sim}{e}}'_{1i}\tilde{\underset{\sim}{e}}_{1i} \tag{31}$$

where

$$\left[\tilde{\underset{\sim}{e}}_{11} \; \tilde{\underset{\sim}{e}}_{12} \cdots \tilde{\underset{\sim}{e}}_{1a}\right] = \underset{\sim}{Z}_1 - (\underset{\sim}{1} \;\; \underset{\sim}{X}_1 \;\; \underset{\sim}{Z}_2)\left[(\underset{\sim}{1} \;\; \underset{\sim}{X}_1 \;\; \underset{\sim}{Z}_2)'(\underset{\sim}{1} \;\; \underset{\sim}{X}_1 \;\; \underset{\sim}{Z}_2)\right]^{-1}(\underset{\sim}{1} \;\; \underset{\sim}{X}_1 \; \underset{\sim}{Z}_2)'\underset{\sim}{Z}_1$$

is an $N \times a$ matrix of estimated residuals. And finally, in (26) the coefficient of σ_2^2 is

$$a\left[(b - 1) - a\,\mathrm{tr}(\underline{W}_1^{-1}\underline{B}_1)\right] = \sum_{j=1}^{b} \tilde{\underline{e}}_{2j}'\tilde{\underline{e}}_{2j} \tag{32}$$

where

$$\left[\tilde{\underline{e}}_{21}\ \tilde{\underline{e}}_{22} \cdots \tilde{\underline{e}}_{2b}\right] = \underline{Z}_2 - (\underline{1}\ \ \underline{X}_1\ \ \underline{Z}_1)\left[(\underline{1}\ \ \underline{X}_1\ \ \underline{Z}_1)'(\underline{1}\ \ \underline{X}_1\ \ \underline{Z}_1)\right]^{-1}(\underline{1}\ \ \underline{X}_1\ \ \underline{Z}_1)'\underline{Z}_2$$

is an $N \times b$ matrix of estimated residuals.

5.1. An Empirical Example

Production data from a cross section of $a = 16$ firms (the u_1-factor) and a time series of $b = 18$ periods (the u_2-factor) are used to illustrate the estimation of variance components for a 2-way crossed classification with one observation per cell. Using a Cobb-Douglas production function with $c = 6$ input variables as covariates, the model is identical to (22) with the logarithms of the output and input variables comprising $\underline{y}$ and $\underline{X}_1$, respectively.

The computed values of the reductions in sums of squares involved in (23) - (26) are

$$R(\mu,\underline{\beta}_1,\underline{u}_1,\underline{u}_2) = 15059.468620 \qquad R(\mu,\underline{\beta}_1) = 15055.819043$$

$$R(\mu,\underline{\beta}_1,\underline{u}_1) = 15058.149475 \qquad R(\mu,\underline{\beta}_1,\underline{u}_2) = 15056.942590$$

so that with $\underline{y}'\underline{y}$ being 15061.39,

$$\begin{aligned} \underline{y}'\underline{y} - R(\mu,\underline{\beta}_1,\underline{u}_1,\underline{u}_2) &= 1.921380, \\ R(\underline{u}_1,\underline{u}_2|\mu,\underline{\beta}_1) = R(\mu,\underline{\beta}_1,\underline{u}_1,\underline{u}_2) - R(\mu,\underline{\beta}_1) &= 3.649577, \\ R(\underline{u}_1|\mu,\underline{\beta}_1,\underline{u}_2) = R(\mu,\underline{\beta}_1,\underline{u}_1,\underline{u}_2) - R(\mu,\underline{\beta}_1,\underline{u}_2) &= 2.526030, \\ \text{and} \quad R(\underline{u}_2|\mu,\underline{\beta}_1,\underline{u}_1) = R(\mu,\underline{\beta}_1,\underline{u}_1,\underline{u}_2) - R(\mu,\underline{\beta}_1,\underline{u}_1) &= 1.319145. \end{aligned} \tag{33}$$

Coefficients of σ_1^2 and σ_2^2 in the expected values of (33) are, by equations (29) - (32), for these data

(28):	189.48		(29):	269.56
(30):	187.99	and	(31):	261.89

so that equations (23) - (26) become for the estimators

$$\begin{bmatrix} 1.921380 \\ 3.649577 \\ 2.526030 \\ 1.319145 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 249 \\ 189.48 & 269.56 & 32 \\ 187.99 & 0 & 15 \\ 0 & 261.89 & 17 \end{bmatrix} \begin{bmatrix} \hat{\sigma}_1^2 \\ \hat{\sigma}_2^2 \\ \hat{\sigma}_e^2 \end{bmatrix} . \tag{34}$$

Solving these in the manner suggested by (27) gives the unbiased estimates $\hat{\sigma}_1^2 = 0.0124$, $\hat{\sigma}_2^2 = 0.0044$, and $\hat{\sigma}_e^2 = 0.0077$.

REFERENCES

[1] Crump, S. L., "The Estimation of Variance Components in Analysis of Variance", Biometrics Bulletin 2, 1946, pp. 7-11.

[2] Eisenhart, C., "The Assumptions Underlying the Analysis of Variance", Biometrics 3, 1947, pp. 1-21.

[3] Henderson, C. R., "Estimation of Variance and Covariance Components", Biometrics 9, June 1953, pp. 226-252.

[4] Henderson, C. R., "Selection Index and Expected Genetic Advances", Statistical Genetics and Plant Breeding, National Academy of Sciences, National Research Council Publication 982, Washington, D. C., 1963, pp. 141-163.

[5] Henderson, C. R., Kempthorne, O., Searle, S. R., and Von Krosigk, C. N., "Estimation of Environmental and Genetic Trends from Records Sub to Culling", Biometrics 15, 1959, pp. 192-218.

[6] Henderson, C. R., Jr., "Comment on the Use of Error Components Models in Combining Cross Section with Time Series Data", Econometrica 39, March 1971, pp. 397-401.

[7] Maddala, G. S., "The Use of Variance Components Models in Pooling Cross Section and Time Series Data", Econometrica 39, March 1971, pp. 341-358.

[8] Nerlove, M., "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections", Econometrica 39, March 1971, pp. 359-382.

[9] Searle, S. R., "Another Look at Henderson's Methods of Estimating Variance Components", Biometrics 24, December 1968, pp. 749-787.

[10] Searle, S. R., "Topics in Variance Components Estimation", Biometrics 27, March 1971, pp. 1-76.

[11] Searle, S. R., Linear Models, Wiley, New York, 1971.

[12] Searle, S. R., Matrix Algebra for Business and Economics , Wiley, New York, 1970.

[13] Wallace, T. D. and Ashig Hussain, "The Use of Error Components Models in Combining Cross Section With Time Series Data", Econometrica 37, January 1969, pp. 55-72.

6. Appendix

6.1. Nerlove's Estimator

With $\underline{S}$ of (14) and $\underline{A}$ of (15) we have, in Section 3.2

$$E(\tilde{\sigma}_1^2) = \mathrm{tr}\left[\underline{A}\ \mathrm{var}(\underline{u}_1 + \underline{S}\underline{e})\right].$$

Hence

$$E(\tilde{\sigma}_1^2) = \mathrm{tr}\left[\underline{A}(\sigma_1^2\underline{I}_a + \sigma_e^2\underline{S}\underline{S}')\right]$$

$$= \sigma_1^2\ \mathrm{tr}(\underline{A}) + \sigma_e^2\ \mathrm{tr}(\underline{A}\underline{S}\underline{S}'). \tag{A1}$$

Now $$\mathrm{tr}(\underline{A}) = \mathrm{tr}\left[(1/a^2)(a\underline{I}_a - \underline{1}_a\underline{1}_a')\right] = (1/a^2)(a^2 - a) = (a - 1)/a. \tag{A2}$$

Also,

$$\underline{S}\underline{S}' = (\underline{Z}_1'\underline{P}\underline{Z}_1)^{-1}\underline{Z}_1'\underline{P}\underline{P}'\underline{Z}_1(\underline{Z}_1'\underline{P}\underline{Z}_1)^{-1}$$

$$= (\underline{Z}_1'\underline{P}\underline{Z}_1)^{-1} \quad \text{because} \quad \underline{P}' = \underline{P} = \underline{P}^2$$

$$= \left(\sum_{i=1}^{a}{}^{+} \underline{1}_b'[\underline{I} - \underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'] \sum_{i=1}^{a}{}^{+} \underline{1}_b\right)^{-1}$$

$$= [b\underline{I}_a - \underline{M}(\underline{X}'\underline{X})^{-1}\underline{M}']^{-1} \tag{A3}$$

where

$$\underline{M}' = \underline{X}' \sum_{i=1}^{a}{}^{+} \underline{1}_b = \begin{bmatrix} b & b & \cdots & b \\ x_{11\cdot} & x_{12\cdot} & \cdots & x_{1a\cdot} \\ x_{21\cdot} & x_{22\cdot} & \cdots & x_{2a\cdot} \\ \vdots & \vdots & \vdots & \vdots \\ x_{c1\cdot} & x_{c2\cdot} & \cdots & x_{ca\cdot} \end{bmatrix}. \tag{A4}$$

Now observe that (A3) is $\mathbf{H}$ as defined in the partitioned inverse

$$\begin{bmatrix} b\mathbf{I}_a & \mathbf{M} \\ \mathbf{M}' & \mathbf{X}'\mathbf{X} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{H} & \mathbf{K} \\ \mathbf{K}' & \mathbf{L} \end{bmatrix}$$

[12, p. 212], and therefore also has the form [11, p. 27]

$$\mathbf{H} = \mathbf{S}\mathbf{S}' = (b\mathbf{I}_a)^{-1} + (b\mathbf{I}_a)^{-1}\mathbf{M}[\mathbf{X}'\mathbf{X} - \mathbf{M}'(b\mathbf{I}_a)^{-1}\mathbf{M}]^{-}\mathbf{M}'(b\mathbf{I}_a)^{-1}$$

$$= b^{-1}\mathbf{I}_a + b^{-2}\mathbf{M}(\mathbf{X}'\mathbf{X} - b^{-1}\mathbf{M}'\mathbf{M})^{-}\mathbf{M}'$$

where the superscript minus denotes a generalized inverse, that of any matrix $\mathbf{Q}$, say, being $\mathbf{Q}^{-}$ such that $\mathbf{Q}\mathbf{Q}^{-}\mathbf{Q} = \mathbf{Q}$. Therefore

$$\mathrm{tr}(\mathbf{A}\mathbf{S}\mathbf{S}') = b^{-1}\mathrm{tr}(\mathbf{A}) + b^{-2}\mathrm{tr}[\mathbf{A}\mathbf{M}(\mathbf{X}'\mathbf{X} - b^{-1}\mathbf{M}'\mathbf{M})^{-}\mathbf{M}'] \tag{A5}$$

and through $a\mathbf{A}$ being symmetric and idempotent the cyclical commutative property of matrix products under the trace operation gives

$$\mathrm{tr}(\mathbf{A}\mathbf{S}\mathbf{S}') = b^{-1}\mathrm{tr}(\mathbf{A}) + a^{-1}b^{-2}\mathrm{tr}\Big[(\mathbf{X}'\mathbf{X} - b^{-1}\mathbf{M}'\mathbf{M})^{-}\mathbf{M}'\mathbf{A}a(\mathbf{M}'\mathbf{A}a)'\Big]$$

$$= (a - 1)/ab + ab^{-2}\mathrm{tr}\Big[(\mathbf{X}'\mathbf{X} - b^{-1}\mathbf{M}'\mathbf{M})^{-}\mathbf{M}'\mathbf{A}(\mathbf{M}'\mathbf{A})'\Big]. \tag{A6}$$

Now because of the form of $\mathbf{M}'$ shown in (A4), $\mathbf{X}'\mathbf{X} - b^{-1}\mathbf{M}'\mathbf{M}$ is the $(c + 1)$-order matrix of within-group sums of squares and products of the elements of the columns of $\mathbf{X}$. In addition, because the first column of $\mathbf{X}$ is $\mathbf{1}$, this becomes

$$(\mathbf{X}'\mathbf{X} - b^{-1}\mathbf{M}'\mathbf{M})^{-} = \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \mathbf{W} \end{bmatrix}^{-} = \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \mathbf{W}^{-1} \end{bmatrix} \tag{A7}$$

where

$$\mathbf{W} = \left\{ \sum_{i=1}^{a} \sum_{j=1}^{b} (x_{tij} - \bar{x}_{ti.})(x_{t'ij} - \bar{x}_{t'i.}) \right\} \quad \text{for } t,t' = 1, 2, \cdots, c \tag{A8}$$

is the $c \times c$ matrix of within-level sums of squares and products of the covariates. Also, from (A4) and the definition of $\underline{A}$ given in (15),

$$\underline{M}'\underline{A} = a^{-1}\underline{M}' - a^{-2}(\underline{M}'\underline{1}_a)\underline{1}_a'$$

$$= (b/a)\begin{bmatrix} 0' \\ \{\bar{x}_{ti.} - \bar{x}_{t..}\} \end{bmatrix} \quad \text{for } t = 1, 2, \cdots, c \text{ and } i = 1, 2, \cdots, a.$$

Hence for (A6)

$$(\underline{M}'\underline{A})(\underline{M}'\underline{A})' = (b^2/a^2)\begin{bmatrix} 0 & \underline{0}' \\ \underline{0} & \underline{B} \end{bmatrix} \tag{A9}$$

where

$$\underline{B} = \left\{\sum_{i=1}^{a}(\bar{x}_{ti.} - \bar{x}_{t..})(\bar{x}_{t'i.} - \bar{x}_{t'..})\right\} \quad \text{for } t,t' = 1, 2, \cdots, c \tag{A10}$$

is the $c \times c$ matrix of between-level sums of squares and products of the covariates. Therefore (A6) becomes, on substituting from (A7) and (A9),

$$\mathrm{tr}(\underline{ASS}') = (a - 1)/ab + ab^{-2}(b^2/a^2)\mathrm{tr}(\underline{W}^{-1}\underline{B})$$

so that, with (A2), we have $E(\tilde{\sigma}_1^2)$ as

$$E(\tilde{\sigma}_1^2) = \Big[(a-1)/a\Big]\sigma_1^2 + \Big[(a - 1)/ab + a^{-1}\mathrm{tr}(\underline{W}^{-1}\underline{B})\Big]\sigma_e^2$$

$$= (a - 1)(b\sigma_1^2 + \sigma_e^2)/ab + a^{-1}\mathrm{tr}(\underline{W}^{-1}\underline{B})\sigma_e^2 \tag{A11}$$

6.2. The 2-Way Classification

For applying the general result (17) to the 3 reductions in sums of squares discussed below equation (23), we have the correspondences of notation shown in Table 1.

Table 1: Correspondences of notation between the general model (16) and (17) and sub-models of (22).

General case, (16) and (17):	$\underline{X}_1$	$\underline{X}_2$	$\underline{\beta}_1$	$\underline{\beta}_2$
Reductions in sums of squares for sub-models of (22)				
$R(\underline{u}_1,\underline{u}_2\mid\mu,\underline{\beta}_1)$:	$[\underline{1}\ \ \underline{X}_1]$	$[\underline{Z}_1\ \ \underline{Z}_2]$	$\begin{bmatrix}\mu\\ \underline{\beta}_1\end{bmatrix}$	$\begin{bmatrix}\underline{u}_1\\ \underline{u}_2\end{bmatrix}$
$R(\underline{u}_1\mid\mu,\underline{\beta}_1,\underline{u}_2)$:	$[\underline{1}\ \ \underline{X}_1\ \ \underline{Z}_2]$	$\underline{Z}_1$	$\begin{bmatrix}\mu\\ \underline{\beta}_1\\ \underline{u}_2\end{bmatrix}$	$\underline{u}_1$
$R(\underline{u}_2\mid\mu,\underline{\beta}_1,\underline{u}_1)$:	$[\underline{1}\ \ \underline{X}_1\ \ \underline{Z}_1]$	$\underline{Z}_2$	$\begin{bmatrix}\mu\\ \underline{\beta}_1\\ \underline{u}_1\end{bmatrix}$	$\underline{u}_2$

Expectation of each of the 3 reductions listed in Table 1 stems from (17). Thus for the first one, with

$$E\left[\begin{bmatrix}\underline{u}_1\\ \underline{u}_2\end{bmatrix}[\underline{u}_1'\ \ \underline{u}_2]\right] = \begin{bmatrix}\sigma_1^2\underline{I}_a & \underline{0}\\ \underline{0} & \sigma_2^2\underline{I}_b\end{bmatrix},$$

we have

$$E\left[R(\underline{u}_1,\underline{u}_2\mid\mu,\underline{\beta}_1)\right]$$

$$= \mathrm{tr}\left\{\begin{bmatrix}\underline{Z}_1'\\ \underline{Z}_2'\end{bmatrix}\left[\underline{I} - [\underline{1}\ \ \underline{X}_1]\begin{bmatrix}N & \underline{1}'\underline{X}_1\\ \underline{X}_1'\underline{1} & \underline{X}_1'\underline{X}_1\end{bmatrix}^{-1}\begin{bmatrix}\underline{1}'\\ \underline{X}_1'\end{bmatrix}\right][\underline{Z}_1\ \ \underline{Z}_2]\begin{bmatrix}\sigma_1^2\underline{I}_a & \underline{0}\\ \underline{0} & \sigma_2^2\underline{I}_b\end{bmatrix}\right\}$$

$$+\ \sigma_e^2\left\{r[\underline{1}\ \ \underline{X}_1\ \ \underline{Z}_1\ \ \underline{Z}_2] - r[\underline{1}\ \ \underline{X}_1]\right\} \qquad \text{(A12)}$$

$$= \sigma_1^2 \mathrm{tr}(\underline{Z}_1'\underline{Z}_1) + \sigma_2^2 \mathrm{tr}(\underline{Z}_2'\underline{Z}_2) + \sigma_e^2\left[c + a + b - 1 - (c + 1)\right]$$

$$- N^{-1}\mathrm{tr}\left\{\begin{bmatrix} \underline{Z}_1' \\ \underline{Z}_2' \end{bmatrix} \underline{1}\underline{1}'[\underline{Z}_1 \quad \underline{Z}_2]\begin{bmatrix} \sigma_1^2\underline{I}_a & \underline{0} \\ \underline{0} & \sigma_2^2\underline{I}_b \end{bmatrix}\right\}$$

$$- \mathrm{tr}\left\{\begin{bmatrix} \underline{Z}_1' \\ \underline{Z}_2' \end{bmatrix}[\underline{X}_1 - \underline{1}\bar{\underline{x}}']\underline{T}^{-1}[\underline{X}_1' - \bar{\underline{x}}\underline{1}'][\underline{Z}_1 \quad \underline{Z}_2]\begin{bmatrix} \sigma_1^2\underline{I}_a & \underline{0} \\ \underline{0} & \sigma_2^2\underline{I}_b \end{bmatrix}\right\}$$

$$= ab\sigma_1^2 + ab\sigma_2^2 + \sigma_e^2(a + b - 2) - \sigma_1^2 ab^2/ab - \sigma_2^2 ba^2/ab$$

$$- \mathrm{tr}\left\{\underline{T}^{-1}[\underline{X}_1' - \bar{\underline{x}}\underline{1}'][\underline{Z}_1 \quad \underline{Z}_2]\begin{bmatrix} \sigma_1^2\underline{I}_a & \underline{0} \\ \underline{0} & \sigma_2^2\underline{I}_b \end{bmatrix}\begin{bmatrix} \underline{Z}_1' \\ \underline{Z}_2' \end{bmatrix}[\underline{X}_1 - \underline{1}\bar{\underline{x}}']\right\}$$

where $\bar{\underline{x}}$ and $\underline{T}$ are defined in (20), and $N = ab$. This may be further simplified to give

$$E\left[R(\underline{u}_1,\underline{u}_2|\mu,\underline{\beta}_1)\right] = b(a - 1)\sigma_1^2 + a(b - 1)\sigma_2^2 + (a + b - 2)\sigma_e^2$$

$$- \mathrm{tr}\left\{\underline{T}^{-1}[b\underline{G}_1' \quad a\underline{G}_2']\begin{bmatrix} \sigma_1^2\underline{I}_a & \underline{0} \\ \underline{0} & \sigma_2^2\underline{I}_b \end{bmatrix}\begin{bmatrix} b\underline{G}_1 \\ a\underline{G}_2 \end{bmatrix}\right\}$$

$$= b\left[a - 1 - b\ \mathrm{tr}(\underline{T}^{-1}\underline{B}_1)\right]\sigma_1^2 + a\left[b - 1 - a\ \mathrm{tr}(\underline{T}^{-1}\underline{B}_2)\right]\sigma_2^2$$

$$+ (a + b - 2)\sigma_e^2 \qquad \text{(A13)}$$

as shown in (24). In these expressions

$$\underline{G}_1 = \left\{\bar{x}_{ti\cdot} - \bar{x}_{t\cdot\cdot}\right\} \quad \text{for } i = 1, 2, \cdots, a \text{ and } t = 1, 2, \cdots, c$$

is an a $\times$ c matrix of the covariates, of the deviations of their means for each level of the u_1-factor from their overall means; and

$$\underset{\sim}{G}_2 = \left\{\bar{x}_{t\cdot j} - \bar{x}_{t\cdot\cdot}\right\} \quad \text{for } j = 1, 2, \cdots, b \text{ and } t = 1, 2, \cdots, c$$

is a b × c matrix of the covariates, of the deviations of their means for each level of the u_2-factor from their overall means. Hence

$$\underset{\sim}{B}_1 = \underset{\sim}{G}_1'\underset{\sim}{G}_1 = \left\{\sum_{i=1}^{a}(\bar{x}_{ti\cdot} - \bar{x}_{t\cdot\cdot})(\bar{x}_{t'i\cdot} - \bar{x}_{t'\cdot\cdot})\right\} \quad \text{for } t,t' = 1, 2, \cdots, c \quad \text{(A14)}$$

is a c × c matrix of the between levels (of the u_1-factor) sums of squares and products of the covariates, and

$$\underset{\sim}{B}_2 = \underset{\sim}{G}_2'\underset{\sim}{G}_2 = \left\{\sum_{j=1}^{b}(\bar{x}_{t\cdot j} - \bar{x}_{t\cdot\cdot})(\bar{x}_{t'\cdot j} - \bar{x}_{t'\cdot\cdot})\right\} \quad \text{for } t,t' = 1, 2, \cdots, c \quad \text{(A15)}$$

is a c × c matrix of the between levels (of the u_2-factor) sums of squares and products of the covariates.

The expected value of the second term in Table 1 proceeds from (17) in similar fashion as follows:

$$E\left[R(\underset{\sim}{u}_1|\mu,\underset{\sim}{\beta}_1,\underset{\sim}{u}_2)\right]$$

$$= \mathrm{tr}\left\{\underset{\sim}{Z}_1'\left[\underset{\sim}{I} - [\underset{\sim}{1} \;\; \underset{\sim}{X}_1 \;\; \underset{\sim}{Z}_2]\begin{bmatrix} N & \underset{\sim}{1}'\underset{\sim}{X}_1 & \underset{\sim}{1}'\underset{\sim}{Z}_2 \\ \underset{\sim}{X}_1'\underset{\sim}{1} & \underset{\sim}{X}_1'\underset{\sim}{X}_1 & \underset{\sim}{X}_1'\underset{\sim}{Z}_2 \\ \underset{\sim}{Z}_2'\underset{\sim}{1} & \underset{\sim}{Z}_2'\underset{\sim}{X}_1 & \underset{\sim}{Z}_2'\underset{\sim}{Z}_2 \end{bmatrix}^{-}\begin{bmatrix} \underset{\sim}{1}' \\ \underset{\sim}{X}_1' \\ \underset{\sim}{Z}_2' \end{bmatrix}\right]\underset{\sim}{Z}_1\right\}\sigma_1^2$$

$$+ \sigma_e^2\left\{r[\underset{\sim}{1} \;\; \underset{\sim}{X}_1 \;\; \underset{\sim}{Z}_1 \;\; \underset{\sim}{Z}_2] - r[\underset{\sim}{1} \;\; \underset{\sim}{X}_1 \;\; \underset{\sim}{Z}_2]\right\} .$$

After considerable algebraic manipulation involving a double use of the (generalized) inverse of a partitioned matrix and the use of properties of $\underset{\sim}{Z}_1$ and $\underset{\sim}{Z}_2$ such as

$$\underset{\sim}{Z}_1'\underset{\sim}{Z}_1 = b\underset{\sim}{I}_a , \quad \underset{\sim}{Z}_1\underset{\sim}{Z}_1' = \sum_{i=1}^{a}{}^{+} \underset{\sim}{1}_b\underset{\sim}{1}_b' ,$$

$$\underset{\sim}{1}_{ab}'\underset{\sim}{Z}_2 = a\underset{\sim}{1}_b', \quad \underset{\sim}{1}_b'\underset{\sim}{Z}_2' = \underset{\sim}{1}_{ab}' , \quad \underset{\sim}{Z}_2'\underset{\sim}{Z}_2 = a\underset{\sim}{I}_b ,$$

$$\underset{\sim}{Z}_2(\underset{\sim}{Z}_2'\underset{\sim}{Z}_2)^{-1}\underset{\sim}{Z}_2' = a^{-1}\underset{\sim}{Z}_2\underset{\sim}{Z}_2' = a^{-1}\begin{bmatrix} \underset{\sim}{I}_b & \cdots & \underset{\sim}{I}_b \\ \vdots & & \\ \underset{\sim}{I}_b & \cdots & \underset{\sim}{I}_b \end{bmatrix}$$

and
$$\underset{\sim}{Z}_1'\underset{\sim}{Z}_2 = \underset{\sim}{1}_a\underset{\sim}{1}_b', \quad \underset{\sim}{Z}_1\underset{\sim}{Z}_1'\underset{\sim}{Z}_2 = \underset{\sim}{1}_{ab}\underset{\sim}{1}_b' ,$$

we get

$$\begin{aligned} &E\left[R(\underset{\sim}{u}_1|\mu,\underset{\sim}{\beta}_1,\underset{\sim}{u}_2)\right] \\ &= \sigma_1^2 N - \sigma_1^2 \operatorname{tr}\left\{ \underset{\sim}{Z}_1'[\underset{\sim}{1} \quad \underset{\sim}{Z}_2] \begin{bmatrix} \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & (\underset{\sim}{Z}_2'\underset{\sim}{Z}_2)^{-1} \end{bmatrix} \begin{bmatrix} \underset{\sim}{1}' \\ \underset{\sim}{Z}_2' \end{bmatrix} \underset{\sim}{Z}_1 \right. \\ &\qquad \left. + \underset{\sim}{Z}_1'\left[\underset{\sim}{X}_1 - \underset{\sim}{Z}_2(\underset{\sim}{Z}_2'\underset{\sim}{Z}_2)^{-1}\underset{\sim}{Z}_2'\underset{\sim}{X}_1\right]\underset{\sim}{W}_2^{-1}\left[\underset{\sim}{X}_1' - \underset{\sim}{X}_1'\underset{\sim}{Z}_2(\underset{\sim}{Z}_2'\underset{\sim}{Z}_2)^{-1}\underset{\sim}{Z}_2'\right]\underset{\sim}{Z}_1 \right\} \\ &\qquad + \sigma_e^2\left[c + a + b - 1 - (c + b)\right] \\ &= b\left[(a - 1) - b \operatorname{tr}(\underset{\sim}{W}_2^{-1}\underset{\sim}{B}_2)\right]\sigma_1^2 + (b - 1)\sigma_e^2 \end{aligned} \tag{A16}$$

where $\underset{\sim}{W}_2$ is a $c \times c$ matrix of the within levels (of the u_2-factor) sums of squares and products of the covariates, i.e.,

$$\underset{\sim}{W}_2 = \left\{ \sum_{j=1}^{b} \sum_{i=1}^{a} (\bar{x}_{tij} - \bar{x}_{t\cdot j})(\bar{x}_{t'ij} - \bar{x}_{t'\cdot j}) \right\} \quad \text{for } t,t' = 1, 2, \cdots, c. \tag{A17}$$

To derive the expected value of the third term in Table 1, observe that its notation is identical to that for the second except for interchanging $\underset{\sim}{u}_1$ with $\underset{\sim}{u}_2$ and $\underset{\sim}{Z}_1$ with $\underset{\sim}{Z}_2$. Hence, making these replacements in (A16) gives

$$E\left[R(\underset{\sim}{u}_2|\mu,\underset{\sim}{\beta}_1,\underset{\sim}{u}_1)\right] = a\left[(b - 1) - a\ \mathrm{tr}(\underset{\sim}{W}_1^{-1}\underset{\sim}{B}_1)\right]\sigma_2^2 + (a - 1)\sigma_e^2 \qquad \text{(A18)}$$

with

$$\underset{\sim}{W}_1 = \left\{\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}_{tij} - \bar{x}_{ti.})(\bar{x}_{t'ij} - \bar{x}_{t'i.})\right\} \quad \text{for } t,t' = 1, 2, \cdots, c. \qquad \text{(A19)}$$

Results (A16) and (A18) are those shown in (25) and (26), respectively.

To obtain alternative expressions for the coefficients of σ_1^2 and σ_2^2 in (A13), notice that (A12) simplifies to

$$E\left[R(\underset{\sim}{u}_1,\underset{\sim}{u}_2|\mu,\underset{\sim}{\beta}_1)\right] = \sigma_1^2\ \mathrm{tr}\left\{\underset{\sim}{Z}_1'\underset{\sim}{M}\underset{\sim}{Z}_1\right\} + \sigma_2^2\ \mathrm{tr}\left\{\underset{\sim}{Z}_2'\underset{\sim}{M}\underset{\sim}{Z}_2\right\} + \sigma_e^2(a + b - 2)$$

where

$$\underset{\sim}{M} = \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}' \quad \text{and} \quad \underset{\sim}{X} = [\underset{\sim}{1}\ \ \underset{\sim}{X}_1].$$

As $\underset{\sim}{M}$ is symmetric and idempotent, the coefficient of σ_1^2 is equal to

$$\mathrm{tr}\left\{(\underset{\sim}{M}\underset{\sim}{Z}_1)'(\underset{\sim}{M}\underset{\sim}{Z}_1)\right\} = \mathrm{tr}\left\{(\underset{\sim}{Z}_1 - \hat{\underset{\sim}{Z}}_1)'(\underset{\sim}{Z}_1 - \hat{\underset{\sim}{Z}}_1)\right\} \qquad \text{(A20)}$$

where $\hat{\underset{\sim}{Z}}_1 = \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Z}_1$ is the ordinary least squares predicted value of $\underset{\sim}{Z}_1$ derived by regressing $\underset{\sim}{Z}_1$ on $\underset{\sim}{X}$. Equivalently, the coefficient of σ_2^2 is equal to

$$\mathrm{tr}\left\{(\underset{\sim}{M}\underset{\sim}{Z}_2)'(\underset{\sim}{M}\underset{\sim}{Z}_2)\right\} = \mathrm{tr}\left\{(\underset{\sim}{Z}_2 - \hat{\underset{\sim}{Z}}_2)'(\underset{\sim}{Z}_2 - \hat{\underset{\sim}{Z}}_2)\right\} \qquad \text{(A21)}$$

where $\hat{\underset{\sim}{Z}}_2 = \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Z}_2$ is the ordinary least squares predicted value of $\underset{\sim}{Z}_2$ derived by regressing $\underset{\sim}{Z}_2$ on $\underset{\sim}{X}$. Thus (A20) and (A21) are the alternative computing formulae for the coefficients of σ_1^2 and σ_2^2, respectively, in (A13) and, equivalently, in (24); they are the formulae shown in (29) and (30).