BEST LINEAR UNBIASED ESTIMATION IN MIXED MODELS OF THE ANALYSIS OF VARIANCE

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, New York 14853

BU-864-M December 1984

Abstract

A broad definition is given of balanced data in mixed models. For all such models, it is shown that the BLUE (best linear unbiased estimator) of an estimable function of the fixed effects is the same as the ordinary least squares estimator (OLSE).

1. INTRODUCTION

a. Fixed effects models

Analysis of variance models are traditionally formulated in terms of additive main effects and additive interaction effects. For example, suppose y_{ijk} is the k'th observation on treatment i of variety j in a two-factor experiment concerned with fertilizer treatments and plant varieties. Then a usual analysis of variance model is of the form

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \tag{1}$$

where μ is a general mean, α_i is the effect on the response variable due to the i'th treatment, β_j is the effect due to the j'th variety, γ_{ij} is the interaction effect between treatment i and variety j, and e_{ijk} is the residual error term defined as $e_{ijk} = y_{ijk} - E(y_{ijk})$ for

$$E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

Paper BU-864-M in the Biometrics Unit series.

where E denotes expectation over repeated sampling.

Models such as (1), where estimation of (and testing of hypotheses about) parameters are the features of interest, are known as fixed effects models, and in such models the customary assumptions about variances and covariances are that each observation has the same variance and that every pair of observations has zero covariance. The dispersion matrix $\underset{\sim}{V}$ of the vector of observations $\underset{\sim}{y}$ then has the form

$$\text{var}(\underset{\sim}{y}) = \underset{\sim}{V} = \sigma^2 \underset{\sim}{I} \quad , \tag{2}$$

$\underset{\sim}{I}$ being an identity matrix and σ^2 being the variance of every observation. An assumption about $\underset{\sim}{V}$ more general than (2) is that it is simply a symmetric, positive semi-definite matrix; and in many cases that it be not just positive semi-definite but positive definite, and hence non-singular.

b. Mixed models

Variations of (1) are models where some or all of the α_i, β_j and γ_{ij} terms are assumed not to be parameters to be estimated, but are modeled as being random variables with zero means and some assumed variance-covariance structure. For example, suppose in the no-interaction form of (1), with one observation y_{ij} on treatment i and variety j, namely

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \quad , \tag{3}$$

that the β_j for $j = 1, \cdots, b$, are modeled as random variables with zero mean $E(\beta_j) = 0 \; \forall \; j$. The β_j are then called random effects and, along with the random error terms e_{ij}, usually have the following variance-covariance structure attributed to them:

$$\text{var}(\beta_i) = \sigma^2_\beta \; \forall \; j \quad , \quad \text{cov}(\beta_j, \beta_{j'}) = 0 \; \forall \; j \neq j' \tag{4}$$

$$\text{var}(e_{ij}) = \sigma^2_e \; \forall \; i,j, \quad \text{cov}(e_{ij}, e_{i'j'}) = 0 \text{ except for } i=i' \text{ and } j=j'$$

and

$$\text{cov}(\beta_j, e_{ij'}) = 0 \; \forall \; i,j,j' \quad .$$

Then with μ and the α_i in (3) being fixed effects and the β_j being random effects, (3) is known as a mixed model. And the variances σ^2_β and σ^2_e of (4) are the variance components. The structure of (4) then leads to $\underset{\sim}{V}$ having elements that are either zero, $\sigma^2_\beta + \sigma^2_e$, or σ^2_β; in general to elements that are either zero, or one of the variance components or a sum of them.

Example 1 Consider (3) and (4), where the β factor represents blocks in a randomized complete blocks experiment. Suppose there are 2 treatments and 3 blocks. Then for a zero element of a matrix being shown as a dot,

$$\underset{\sim}{V} = \text{var}\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} \sigma^2_\beta+\sigma^2_e & \cdot & \cdot & \sigma^2_\beta & \cdot & \cdot \\ \cdot & \sigma^2_\beta+\sigma^2_e & \cdot & \cdot & \sigma^2_\beta & \cdot \\ \cdot & \cdot & \sigma^2_\beta+\sigma^2_e & \cdot & \cdot & \sigma^2_\beta \\ \sigma^2_\beta & \cdot & \cdot & \sigma^2_\beta+\sigma^2_e & \cdot & \cdot \\ \cdot & \sigma^2_\beta & \cdot & \cdot & \sigma^2_\beta+\sigma^2_e & \cdot \\ \cdot & \cdot & \sigma^2_\beta & \cdot & \cdot & \sigma^2_\beta+\sigma^2_e \end{bmatrix} .$$

c. Estimation with balanced data

Section 3 formulates a set of models that specifies a wide class of balanced data. First, though, we appeal to the general understanding that balanced data have equal numbers of observations in the subclasses. Model equations (1) and (3) are examples, having, for each treatment-variety combination, one observation and (with $k = 1, 2, \cdots, n$) n observations, respectively. In both cases the best linear unbiased estimator (BLUE) of a treatment difference is a well known, simple function of means. Thus when each of (1) and (3) are fixed effects models, the BLUE of $\alpha_i - \alpha_{i'}$ is

$$\text{BLUE}(\alpha_i - \alpha_{i'}) = \bar{y}_{(i)} - \bar{y}_{(i')} \qquad (5)$$

where $\bar{y}_{(i)}$ is the mean of all observations on treatment i. Moreover, the right-hand side of (5) is also the ordinary least squares estimator of $\alpha_i - \alpha_{i'}$. Hence, for these examples

$$\text{BLUE}\ (\alpha_i - \alpha_{i'}) = \text{OLSE}(\alpha_i - \alpha_{i'}) \quad . \tag{6}$$

Of additional importance is the fact that although (5) is true when (1) and (3) are fixed effects models, it is also true when (1) and (3) are mixed models with αs fixed. The generalization of (6) is that for any estimable function of fixed effects in a mixed model with balanced data, BLUE = OLSE. The utility of this result is that although a BLUE is a desirable estimator, its direct derivation generally involves inverting $\underset{\sim}{V}$, which can be tedious; in contrast, with balanced data, the OLSE is often easily derived as a simple function of observed means. Moreover, the equality BLUE = OLSE for balanced data is broad in scope. For example, (5) is true for (1) being not only a fixed effects model, but also a mixed model with βs, or γs, or βs and γs taken as random effects. Furthermore (5), as an example of (6) is also true if (1) is extended by the additional of other random effects: for example, in the model $y_{ijk\ell m} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \theta_k + \tau_\ell + \delta_{jk} + e_{ijk\ell m}$, with μ and αs being fixed effects and all other effects being random, (5) is still true.

We proceed to establish (6) for any mixed model with balanced data. To do so we first describe a general mixed model and then give a broad definition of balanced data.

2. A GENERAL MIXED MODEL

a. Description

The elements of the mixed model (3) are of two kinds: μ and α_i that are fixed effects, and β_i and e_{ij} that are random variables. Recognizing the dichotomy of fixed and random effects in a mixed model, we write the model equation for a vector of observations $\underset{\sim}{y}$ as

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{Z}\underset{\sim}{u} \tag{7}$$

where $\underset{\sim}{\beta}$ is a vector of fixed effects and $\underset{\sim}{u}$ is a vector of random effects, including error terms. The matrices and vectors of (7) are partitioned thus:

$$\begin{aligned}\underset{\sim}{X} &= [\underset{\sim}{X}_1 \;\; \underset{\sim}{X}_2 \;\; \cdots \;\; \underset{\sim}{X}_d \;\; \cdots \;\; \underset{\sim}{X}_f] \\ \underset{\sim}{\beta} &= [\underset{\sim}{\beta}_1' \;\; \underset{\sim}{\beta}_2' \;\; \cdots \;\; \underset{\sim}{\beta}_d' \;\; \cdots \;\; \underset{\sim}{\beta}_f']'\end{aligned} \quad \text{and} \quad \begin{aligned}\underset{\sim}{Z} &= [\underset{\sim}{Z}_1 \;\; \underset{\sim}{Z}_2 \;\; \cdots \;\; \underset{\sim}{Z}_q \;\; \cdots \;\; \underset{\sim}{Z}_r] \\ \underset{\sim}{u} &= [\underset{\sim}{u}_1' \;\; \underset{\sim}{u}_2' \;\; \cdots \;\; \underset{\sim}{u}_q' \;\; \cdots \;\; \underset{\sim}{u}_r']'\end{aligned} \tag{8}$$

Each $\underset{\sim}{\beta}_d$ for $d = 1, 2, \cdots, f$ has as its element the h_d effects corresponding to the h_d levels of the d'th fixed effect (main effect or interaction) factor, and $\underset{\sim}{X}_d$ is the incidence matrix corresponding to $\underset{\sim}{\beta}_d$. Similarly, $\underset{\sim}{u}_q$ (of p_q elements) and $\underset{\sim}{Z}_q$ for $q = 1, 2, \cdots, r-1$ are defined for the random effect (main effect or interaction) factors analogously to $\underset{\sim}{\beta}_d$ and $\underset{\sim}{X}_d$ for fixed effect factors. For $q = r$, we define $\underset{\sim}{u}_r = \underset{\sim}{e}$, the vector of error terms, and accordingly $\underset{\sim}{Z}_r = \underset{\sim}{I}_N$ where N is the total number of observations, and $p_r = N$.

Example 2 Using (3) and (4) as the model for a randomized complete blocks experiment for a treatments in b blocks, μ and $[\alpha_1 \cdots \alpha_a]'$ would be $\underset{\sim}{\beta}_1$ and $\underset{\sim}{\beta}_2$ of (8), respectively, and $[\beta_1 \cdots \beta_b]$ and the e_{ij}-terms of (3) would be $\underset{\sim}{u}_1$ and $\underset{\sim}{u}_2$ of (8), respectively.

The variance and covariance properties of (4) generalized to $\underset{\sim}{u}$ are

$$\text{var}(\underset{\sim}{u}_q) = \sigma_q^2 I_{p_q} \qquad \text{for} \qquad q = 1, 2, \cdots, r$$

and (9)

$$\text{cov}(\underset{\sim}{u}_q, \underset{\sim}{u}'_{q'}) = \underset{\sim}{0}_{p_q \times p_{q'}} \qquad \text{for} \quad q \neq q' = 1, 2, \cdots, r \quad .$$

Hence from (7) the variance-covariance matrix of $\underset{\sim}{y}$ is

$$\underset{\sim}{V} = \text{var}(\underset{\sim}{y}) = \text{var}(\underset{\sim}{Z}\underset{\sim}{u}) = \sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{Z}_q \underset{\sim}{Z}'_q \quad . \tag{10}$$

Thus (7) through (10) constitute a description of a general mixed model.

b. Estimation

The OLSE estimator of an estimable function $\underset{\sim}{\lambda}'\underset{\sim}{X}\underset{\sim}{\beta}$ of the parameters in $\underset{\sim}{\beta}$ in the model (7) will be denoted by $\text{OLSE}(\underset{\sim}{\lambda}'\underset{\sim}{X}\underset{\sim}{\beta})$ and is, as is well-known,

$$\text{OLSE}(\underset{\sim}{\lambda}'\underset{\sim}{X}\underset{\sim}{\beta}) = \underset{\sim}{\lambda}'\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{y} \tag{11}$$

where $(\underset{\sim}{X}'\underset{\sim}{X})^{-}$ is a generalized inverse of $\underset{\sim}{X}'\underset{\sim}{X}$, i.e., $(\underset{\sim}{X}'\underset{\sim}{X})^{-}$ is any matrix satisfying

$$\underset{\sim}{X}'\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{X}'\underset{\sim}{X} \quad .$$

Similarly the BLUE of that same estimable $\underset{\sim}{\lambda}'\underset{\sim}{X}\underset{\sim}{\beta}$ is

$$\text{BLUE}(\underset{\sim}{\lambda}'\underset{\sim}{X}\underset{\sim}{\beta}) = \underset{\sim}{\lambda}'\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} \quad , \tag{12}$$

where $\underset{\sim}{V}$ is assumed to be positive definite.

In fixed effects models, $\underset{\sim}{V} = \sigma^2\underset{\sim}{I}$, as in (2), whereupon (12) very simply reduces to (11), as is well known. An extension to $\underset{\sim}{V} = [(1-\rho)\underset{\sim}{I} + \rho\underset{\sim}{J}]\sigma^2$ is given by McElroy (1967) and, in complete generality, Zyskind (1967) has shown that these two estimators are equal, if and only if

$$\underset{\sim}{V}\underset{\sim}{X} = \underset{\sim}{X}\underset{\sim}{Q} \qquad \text{for some } \underset{\sim}{Q} \quad . \tag{13}$$

Graybill (1976, p. 209) also has this result, restricted to $\underset{\sim}{X}$ of full column rank. We use (13) to show for a broad definition of balanced data that for mixed models of the form (7) through (10) the BLUE of an estimable function of the fixed effects parameters is the same as the OLSE.

3. BALANCED DATA

We deal with data categorized by a number of factors, each of which is either a main effects factor (including the possibility of nested main effects factors), or an interaction factor representing the interaction of two or more main effects factors. Suppose there are m main effects factors, with the t'th one having N_t levels, for t = 1, 2, ⋯, m. Then the k'th observation in the "cell" defined by the i_t'th level (for $i_t = 1, \cdots, N_t$) of the t'th main effect for t = 1, ⋯, m, where there are $n_{i_1 i_2 \cdots i_t \cdots i_m}$ such observations, is $y_{i_1 i_2 \cdots i_t \cdots i_m k}$ for k = 1, 2, ⋯, $n_{i_1 i_2 \cdots i_t \cdots i_m}$. On defining $\underset{\sim}{i} = [i_1 \; i_2 \; \cdots \; i_m]$, a typical observation can then be denoted as $y_{\underset{\sim}{i}k}$ for k = 1, 2, ⋯, $n_{\underset{\sim}{i}}$. Furthermore, the total number of observations is

$$N = p_r = \sum_{\underset{\sim}{i}=\underset{\sim}{1}'_m}^{\underset{\sim}{i}=\underset{\sim}{N}'} n_{\underset{\sim}{i}} \qquad \text{for } \underset{\sim}{N}' = [N_1 \;\; N_2 \; \cdots \; N_t \; \cdots \; N_m] \quad .$$

($\underset{\sim}{1}'_m$ is a row vector of m unities.)

A tight, rigorous, formal and complete definition of balanced data is elusive. Development of such a definition would, as Cornfield and Tukey (1956) write, involve "⋯ systematic algebra [which] can take us deep into the forest of notation. But the detailed manipulation will, sooner or later, blot out any understanding we may have started with." Nevertheless, one formulation of a model that yields a wide class of balanced data situations is as follows. It is similar to that used by Smith and Hocking

(1978), Searle and Henderson (1979), Seifert (1979), Khuri (1981) and Anderson *et al.* (1984).

The balanced data models we consider are those that have $n_{\underset{\sim}{i}} = n \; \forall \; \underset{\sim}{i}$. They also have each $\underset{\sim}{X}_d$ and each $\underset{\sim}{Z}_q$ of (8) being a Kronecker product (KP, for brevity) of m + 1 matrices, each of which is either an $\underset{\sim}{I}$-matrix or a $\underset{\sim}{1}$-vector; i.e.,

$$\text{each } \underset{\sim}{X}_d \text{ and each } \underset{\sim}{Z}_q \text{ is a KP of m+1 matrices that are each } \underset{\sim}{I} \text{ or } \underset{\sim}{1} \;. \tag{14}$$

The occurrence of the $\underset{\sim}{I}$-matrices and $\underset{\sim}{1}$-vectors in these KPs is as follows. First, corresponding to the scalar parameter μ in the model is $\underset{\sim}{X}_1$ which is $\underset{\sim}{1}_N$, and so every matrix in its KP is a $\underset{\sim}{1}$:

$$\underset{\sim}{X}_1 = \underset{\sim}{1}_N = \underset{\sim}{1}_{N_1} * \underset{\sim}{1}_{N_2} * \cdots * \underset{\sim}{1}_{N_t} * \cdots * \underset{\sim}{1}_{N_m} * \underset{\sim}{1}_n \quad ,$$

where * represents the operation of Kronecker multiplication. Second, corresponding to $\underset{\sim}{u}_r = \underset{\sim}{e}$ is $\underset{\sim}{Z}_N$, and so each of the m + 1 matrices in the KP that is $\underset{\sim}{Z}_r = \underset{\sim}{I}_N$ is an $\underset{\sim}{I}$-matrix:

$$\underset{\sim}{Z}_r = \underset{\sim}{I}_N = \underset{\sim}{I}_{N_1} * \underset{\sim}{I}_{N_2} * \cdots * \underset{\sim}{I}_{N_t} * \cdots * \underset{\sim}{I}_{N_m} * \underset{\sim}{I}_n \quad .$$

Finally, in the KP for each $\underset{\sim}{X}_d$ and $\underset{\sim}{Z}_q$ (other than $\underset{\sim}{X}_1$ and $\underset{\sim}{Z}_r$), the t'th matrix corresponds to the t'th main effects factor and is $\underset{\sim}{I}_{N_t}$ when that factor is part of the definition of the factor corresponding to $\underset{\sim}{X}_d$ or $\underset{\sim}{Z}_q$; otherwise it is $\underset{\sim}{1}_{N_t}$. This is for t = 1, $\cdots$, m. And for all $\underset{\sim}{X}_d$ and $\underset{\sim}{Z}_q$, other than $\underset{\sim}{Z}_r$, the (m+1)'th matrix in the KP is $\underset{\sim}{1}_n$.

The phrase "part of the definition" demands explanation. It is exemplified in the 2-factor model (1), wherein the two main effects factors are each part of the definition of the interaction factor. Similarly, if nested within an α-factor there is a β-factor then the α-factor is part of the definition of that β-factor. (See also, comments B and C which follow the examples.)

Each h_d and p_q (number of levels in the d'th fixed factor and the q'th random factor, respectively) in the balanced data we have defined is the product of the numbers of columns in the $\underset{\sim}{I}$ and $\underset{\sim}{1}$ terms in the KP (14) that is $\underset{\sim}{X}_d$ and $\underset{\sim}{Z}_q$. Hence h_d is the product of the N_t values for the main effects factors that are part of the definition of the d'th fixed effect factor; p_q is a similar product for the q'th random effects factor.

Examples We give four examples that are each in terms of those of the following vectors that are appropriate: $\underset{\sim}{\alpha} = [\alpha_1, \cdots, \alpha_a]'$, $\underset{\sim}{\beta} = [\beta_1, \cdots, \beta_b]'$ or $\underset{\sim}{\beta}_+ = [\beta_{11} \cdots \beta_{1b} \; \beta_{21} \cdots \beta_{2b} \cdots \beta_{a1} \cdots \beta_{ab}]'$, $\underset{\sim}{\gamma} = [\gamma_{11} \cdots \gamma_{1b} \; \gamma_{21} \cdots \gamma_{2b} \cdots \gamma_{a1} \cdots \gamma_{ab}]'$, and $\underset{\sim}{e}$, the vector of error terms, the same order as $\underset{\sim}{y}$. Determination of which KPs are $\underset{\sim}{X}$-matrices and which are $\underset{\sim}{Z}$-matrices is governed by which factors are defined as fixed effects and which are random. This is illustrated for only example (iii).

(i) <u>One-way classification</u>: $y_{ij} = \mu + \alpha_i + e_{ij}$ with $i=1, \cdots, a$ and $j=1, \cdots, n$.

$$\underset{\sim}{y} = (\underset{\sim}{1}_a * \underset{\sim}{1}_n)\mu + (\underset{\sim}{I}_a * \underset{\sim}{1}_n)\alpha + (\underset{\sim}{I}_a * \underset{\sim}{I}_n)e \quad . \tag{15}$$

(ii) <u>Two-way crossed classification, no interaction, and one observation per cell</u>: $y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$ for $i=1, \cdots, a$ and $j=1, \cdots, b$.

$$\underset{\sim}{y} = (\underset{\sim}{1}_a * \underset{\sim}{1}_b)\mu + (\underset{\sim}{I}_a * \underset{\sim}{1}_b)\underset{\sim}{\alpha} + (\underset{\sim}{1}_a * \underset{\sim}{I}_b)\underset{\sim}{\beta} + (\underset{\sim}{I}_a * \underset{\sim}{I}_b)\underset{\sim}{e} \quad . \tag{16}$$

(iii) <u>Two-way crossed classification, with interaction and n observations per cell</u>: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$ with $i=1, \cdots, a$, $j=1, \cdots, b$ and $k=1, \cdots, n$.

$$\underset{\sim}{y} = (\underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)\mu + (\underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)\underset{\sim}{\alpha} + (\underset{\sim}{1}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n)\underset{\sim}{\beta} + (\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n)\underset{\sim}{\gamma} + (\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_n)\underset{\sim}{e} \quad . \tag{17}$$

Suppose in (17) that elements of $\underset{\sim}{\beta}$ and $\underset{\sim}{\chi}$ were taken to be random effects. Then the terms of (8) for the general mixed model would have the following values:

$$
\begin{array}{llllllll}
m=3, & f=2 \text{ with } & h_1 = N_1 & = 1 & \text{and} & \underset{\sim}{X}_1 = \underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n & \text{for } \underset{\sim}{\beta}_1 = \mu & , \\
 & \text{and} & h_2 = N_2 & = a & \text{and} & \underset{\sim}{X}_2 = \underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n & \text{for } \underset{\sim}{\beta}_2 = \underset{\sim}{\alpha} & ;
\end{array}
$$

$$
\begin{array}{llllllll}
r=3 & \text{with} & p_1 = N_3 & = b & \text{and} & \underset{\sim}{Z}_1 = \underset{\sim}{1}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n & \text{for } \underset{\sim}{u}_1 = \underset{\sim}{\beta} & , \\
 & & p_2 = N_2N_3 & = ab & \text{and} & \underset{\sim}{Z}_2 = \underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n & \text{for } \underset{\sim}{u}_2 = \underset{\sim}{\chi} & , \\
 & \text{and} & p_3 = N_2N_3n & = abn & \text{and} & \underset{\sim}{Z}_3 = \underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_n & \text{for } \underset{\sim}{u}_3 = \underset{\sim}{e} & .
\end{array}
$$

(iv) <u>Two-way nested classification</u>: $y_{ij} = \mu + \alpha_i + \beta_{ij} + e_{ijk}$ for $i=1,\cdots,a$, $j=1,\cdots,b$ and $k=1,\cdots,n$.

$$
\begin{aligned}
\underset{\sim}{y} = (\underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)\mu + (\underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)\underset{\sim}{\alpha} + (\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n)\underset{\sim}{\beta}_+ \\
+ (\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_n)\underset{\sim}{e} \quad .
\end{aligned}
\tag{18}
$$

Comments on the examples. Several comments are in order. (A) In every case $\underset{\sim}{X}_1$ for μ is $\underset{\sim}{1}$, a KP of 1-vectors; and $\underset{\sim}{Z}_r$ for $\underset{\sim}{e}$ is $\underset{\sim}{I}$, a KP of $\underset{\sim}{I}$-matrices. (B) In every case the KP that is the coefficient of $\underset{\sim}{\alpha}$ has only one $\underset{\sim}{I}$-matrix in it, namely $\underset{\sim}{I}_a$. This is so because, obviously, the definition of $\underset{\sim}{\alpha}$ involves only $\underset{\sim}{\alpha}$. The same is true of the coefficient of $\underset{\sim}{\beta}$ in (16) and (17). (C) In contrast, the KP that is the coefficient of $\underset{\sim}{\beta}_+$ in (18) has two $\underset{\sim}{I}$-matrices, $\underset{\sim}{I}_a$ and $\underset{\sim}{I}_b$. This is because $\underset{\sim}{\beta}_+$ has elements that represent the nesting of the β-factor within the α-factor. Thus the α-factor is involved in the definition of $\underset{\sim}{\beta}_+$ and so the coefficient of $\underset{\sim}{\beta}_+$ contains $\underset{\sim}{I}_a$ and $\underset{\sim}{I}_b$. Thus the coefficient of $\underset{\sim}{\beta}_+$ in (18) is the same as that of $\underset{\sim}{\chi}$, the interaction term, in (17). Judged solely by their coefficients, $\underset{\sim}{\beta}_+$ and $\underset{\sim}{\chi}$ would therefore appear to be the same. What makes $\underset{\sim}{\chi}$ an interaction term is that both main effect factors that go into defining it are also present on their own in (17), but with β_+, only one factor that goes into defining it is present on its own in (18), and so β_+ represents nesting. In other words, a factor that looks like an interaction factor is such when all of its associated main effects factors are present in the model; otherwise it is a nested factor. (D) Equation (16) is a special case of (17) with χ omitted and n=1 and hence, for example, $1_a * 1_b * 1_n = 1_a * 1_b * 1 = 1_a * 1_b$.

A final observation concerns $V = \sum_{q=1}^{r} \sigma_q^2 Z_q Z_q'$ of (10), based on the general result that $(A * B)(P * Q) = AP * BQ$, given the necessary conformability requirements. Thus, for $1_n 1_n' = J_n$ being a square matrix of order n with every element unity, we have from (14) that every $Z_q Z_q'$ is a KP of I and J matrices. Hence we rewrite (10) as

$$V = \sum_{q=1}^{r} \sigma_q^2 \text{ (the KP of } I \text{ and } J \text{ matrices that is } Z_q Z_q') \quad . \tag{19}$$

4. ESTIMATION FROM BALANCED DATA

We now show for mixed models as specified in (7) – (10), with balanced data as defined in the preceding section, that the BLUE of (12) equals the OLSE of (11). We do this by showing that (13) is satisfied for V of (19) and $X = \{X_d\}$, $d = 1,\cdots,f$ of (14) with X_d being a KP of I-matrices and 1-vectors.

Writing W_q for $Z_q Z_q'$ of (19) we have

$$W_q = Z_q Z_q' = (W_{q1} * W_{q2} * \cdots * W_{qt} * \cdots * W_{q,m+1}) = \mathop{*}_{t=1}^{m+1} W_{qt} \quad , \tag{20}$$

where, from (19) each W_{qt} is either an I or a J matrix. Similarly, from (8),

$$X = [X_1 \ \ X_2 \ \cdots \ X_d \ \cdots \ X_f] \ \text{ with } \ X_d = \mathop{*}_{t=1}^{m+1} X_{dt} \quad , \tag{21}$$

where each X_{dt} is either I_{N_t} or 1_{N_t}. Then from (19)

$$\underset{\sim\sim}{VX} = \left\{ \sum_{q=1}^{r} \sigma^2_q \underset{\sim}{Z}_q \underset{\sim}{Z}'_q \underset{\sim}{X}_d \right\}_{d=1}^{d=f}$$

where, by the curly braces notation, we mean that $\underset{\sim\sim}{VX}$ is partitioned into a row of f sub-matrices. Thus

$$\underset{\sim\sim}{VX} = \left\{ \sum_{q=1}^{r} \sigma^2_q \underset{\sim}{W}_q \underset{\sim}{X}_d \right\}_{d=1}^{d=f} \tag{22}$$

$$= \left\{ \sum_{q=1}^{r} \sigma^2_q \overset{m+1}{\underset{t=1}{*}} \underset{\sim}{W}_{qt} \underset{\sim}{X}_{dt} \right\}_{d=1}^{d=f} . \tag{23}$$

Now in (20), $\underset{\sim}{W}_{qt}$ is either $\underset{\sim}{I}$ or $\underset{\sim}{J}$, and in (21) each $\underset{\sim}{X}_{dt}$ is either $\underset{\sim}{I}$ or $\underset{\sim}{1}$, all of order $\underset{\sim}{N}_t$. Therefore the four possible values of the product $\underset{\sim}{W}_{qt} \underset{\sim}{X}_{dt}$, together with the definition of a matrix $\underset{\sim}{M}_{qdt}$ such that $\underset{\sim}{W}_{qt} \underset{\sim}{X}_{dt} = \underset{\sim}{X}_{dt} \underset{\sim}{M}_{qdt}$ in each case, are as follows:

$\underset{\sim}{W}_{qt}$	$\underset{\sim}{X}_{dt}$	$\underset{\sim}{W}_{qt}\underset{\sim}{X}_{dt} = \underset{\sim}{X}_{dt}\underset{\sim}{M}_{qdt}$	$\underset{\sim}{M}_{qdt}$
$\underset{\sim}{I}$	$\underset{\sim}{I}$	$\underset{\sim}{I} = \underset{\sim\sim}{II}$	$\underset{\sim}{I}$
$\underset{\sim}{I}$	$\underset{\sim}{1}$	$\underset{\sim}{1} = \underset{\sim}{1}1$	1
$\underset{\sim}{J}$	$\underset{\sim}{I}$	$\underset{\sim}{J} = \underset{\sim\sim}{IJ}$	$\underset{\sim}{J}$
$\underset{\sim}{J}$	$\underset{\sim}{1}$	$N_t\underset{\sim}{1} = \underset{\sim}{1}N_t$	N_t

Therefore from (23)

$$\underset{\sim\sim}{VX} = \left\{ \sum_{q=1}^{r} \sigma^2_q \overset{m+1}{\underset{t=1}{*}} \underset{\sim}{X}_{dt} \underset{\sim}{M}_{qdt} \right\}_{d=1}^{d=f} , \tag{24}$$

$$= \left\{ \sum_{q=1}^{r} \sigma^2_q \underset{\sim}{X}_d \underset{\sim}{M}_{qd} \right\}_{d=1}^{d=f} , \tag{25}$$

for

$$\underset{\sim}{M}_{qd} = \underset{\sim}{M}_{qd1} * \underset{\sim}{M}_{qd2} * \cdots * \underset{\sim}{M}_{qdt} * \cdots * \underset{\sim}{M}_{q,d,m+1} . \tag{26}$$

Derivation both of (23) from (22) and of (25) from (24) is based both on $\underset{\sim}{X}_d$ and $\underset{\sim}{M}_q$ each being a KP, and on the product rule for KP quoted earlier.

The conformability requirements of the regular products in (24) might seem to be lacking because, from the preceding table, two forms of $\underset{\sim}{M}_{qdt}$ are scalars. However, both regular and Kronecker products of matrices do exist when one or more of the matrices is a scalar; e.g., for scalar θ, both $\underset{\sim}{A}\theta$ and $(\underset{\sim}{A} * \underset{\sim}{B})(\theta * \underset{\sim}{L}) = \underset{\sim}{A}\theta * \underset{\sim}{B}\underset{\sim}{L}$ exist. Therefore (25) exists. Hence, on writing

$$\underset{\sim}{Q} = \text{diag}\left\{\sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{M}_{qd}\right\}_{d=1}^{d=f} ,$$

the block diagonal matrix of matrices $\sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{M}_{qd}$, we get from (25)

$$\underset{\sim}{V}\underset{\sim}{X} = [\underset{\sim}{X}_1 \cdots \underset{\sim}{X}_d \cdots \underset{\sim}{X}_f]\begin{bmatrix} \sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{M}_{q1} & & & & \underset{\sim}{0} \\ & \ddots & & & \\ & & \sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{M}_{qd} & & \\ & & & \ddots & \\ \underset{\sim}{0} & & & & \sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{M}_{qf} \end{bmatrix} \qquad (27)$$

$$= \underset{\sim}{X}\underset{\sim}{Q} \quad .$$

Thus Zyskind's condition of (13) is satisfied. Hence, with balanced data as here defined, the BLUE of an estimable function of the fixed effects in any mixed model is the same as the OLSE.

A final note: each sum $\sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{M}_{qd}$ in (27) does exist because, as a result of (26), the order of $\underset{\sim}{M}_{qd}$ is the product of the orders of $\underset{\sim}{M}_{qdt}$ for $t = 1, \cdots, m+1$; and (from the Table) each $\underset{\sim}{M}_{qdt}$ is square of order either N_t or 1. Furthermore, that order is N_t only when $\underset{\sim}{X}_{dt} = \underset{\sim}{I}$; and this is so only when the t'th main effects factor is involved in defining the d'th fixed effects factor. Hence the order of $\underset{\sim}{M}_{qd}$ is the product of such N_t values, and this is h_d; thus $\underset{\sim}{M}_{qd}$ has order h_d for all q and so $\sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{M}_{qd}$ exists.

Example Suppose in (1) and (17) that the βs and γs are random effects. Then

$$\tilde{X} = [\tilde{1}_a * \tilde{1}_b * \tilde{1}_n \quad \tilde{I}_a * \tilde{1}_b * \tilde{1}_n]$$

and

$$\tilde{V} = \sigma^2_\beta(\tilde{J}_a * \tilde{I}_b * \tilde{J}_n) + \sigma^2_\gamma(\tilde{I}_a * \tilde{I}_b * \tilde{J}_n) + \sigma^2_e(\tilde{I}_a * \tilde{I}_b * \tilde{I}_n) \quad .$$

Hence in $\tilde{V}\tilde{X}$ the first sub-matrix is

$$\tilde{V}(\tilde{1}_a * \tilde{1}_b * \tilde{1}_n) = \sigma^2_\beta(a\tilde{1}_a * \tilde{1}_b * n\tilde{1}_n) + \sigma^2_\gamma(\tilde{1}_a * \tilde{1}_b * n\tilde{1}_n) + \sigma^2_e(\tilde{1}_a * \tilde{1}_b * \tilde{1}_n)$$

$$= (\tilde{1}_a * \tilde{1}_b * \tilde{1}_n)[\sigma^2_\beta(a * 1 * n) + \sigma^2_\gamma(1 * 1 * n) + \sigma^2_e(1 * 1 * 1)] \quad . \qquad (28)$$

Similarly, the second sub-matrix of $\tilde{V}\tilde{X}$ is

$$\tilde{V}(\tilde{I}_a * \tilde{1}_b * \tilde{1}_n) = \sigma^2_\beta(\tilde{J}_a * \tilde{1}_b * n\tilde{1}_n) + \sigma^2_\gamma(\tilde{I}_a * \tilde{1}_b * n\tilde{1}_n) + \sigma^2_e(\tilde{I}_a * \tilde{1}_b * \tilde{1}_n)$$

$$= (\tilde{I}_a * \tilde{1}_b * \tilde{1}_n)[\sigma^2_\beta(\tilde{J}_a * 1 * n) + \sigma^2_\gamma(\tilde{I}_a * 1 * n) + \sigma^2_e(\tilde{I}_a * 1 * 1)] \ . \qquad (29)$$

Hence

$$\tilde{V}\tilde{X} = [\tilde{1}_a * \tilde{1}_b * \tilde{1}_n \quad \tilde{I}_a * \tilde{1}_b * \tilde{1}_n]\begin{bmatrix}\tilde{M}_1 & \tilde{0} \\ \tilde{0} & \tilde{M}_2\end{bmatrix} = \tilde{X}\begin{bmatrix}\tilde{M}_1 & \tilde{0} \\ \tilde{0} & \tilde{M}_2\end{bmatrix}$$

for $\tilde{M}_1$ and $\tilde{M}_2$ being the matrices in square braces in (28) and (29), respectively, namely

$$\tilde{M}_1 = an\sigma^2_\beta + n\sigma^2_\gamma + \sigma^2_e \quad \text{and} \quad \tilde{M}_2 = n\sigma^2_\beta\tilde{J}_a + n\sigma^2_\gamma\tilde{I}_a + \sigma^2_e\tilde{I}_a \quad .$$

References

Anderson, R. D., Henderson, H. V., Pukelsheim, F. and Searle, S. R. (1984). Best estimation of variance components from balanced data with arbitrary kurtosis. *Mathematische Operationsforschung und Statistik. Series Statistik* 15, 163-176.

Cornfield, J. and Tukey, J. W. (1956). Average values of mean squares in factorials. *Ann. Math. Statist.* 27, 907-949.

Graybill, F. A. (1976). *Theory and Application of the Linear Model*, Duxbury, North Scituate, Massachusetts.

Khuri, A. I. (1981). Anova sums of squares for balanced mixed-effects models revisited. Technical Report No. 162, University of Florida.

McElroy, F. W. (1967). A necessary and sufficient condition that ordinary least squares estimators be best linear unbiased. *J. American Statistical Association* 62, 1302-1304.

Searle, S. R. and Henderson, H. V. (1979). Dispersion matrices in variance components models. *J. American Statistical Association* 74, 465-470.

Seifert, B. (1979). Optimal testing for fixed effects in general balanced mixed classification models. *Mathematische Operationsforschung und Statistik. Series Statistics* 10, 237-255.

Smith, D. W. and Hocking, R. R. (1978). Maximum likelihood analysis of the mixed model: the balanced case. *Communications in Statistics — Theory and Methods* A7, 1253-1266.

Zyskind, G. (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *Ann. Math. Statistics* 38, 1092-1109.