Nonnegative estimates of variance components in a two-way random model

Jaesung Choi[1,a]

[a]Department of Statistics, Keimyung University, Korea

Abstract

This paper discusses a method for obtaining nonnegative estimates for variance components in a random effects model. A variance component should be positive by definition. Nevertheless, estimates of variance components are sometimes given as negative values, which is not desirable. The proposed method is based on two basic ideas. One is the identification of the orthogonal vector subspaces according to factors and the other is to ascertain the projection in each orthogonal vector subspace. Hence, an observation vector can be denoted by the sum of projections. The method suggested here always produces nonnegative estimates using projections. Hartley's synthesis is used for the calculation of expected values of quadratic forms. It also discusses how to set up a residual model for each projection.

Keywords: projections, random effects, vector space, quadratic forms, synthesis

1. Introduction

When a linear model contains more than one random component the model is said to be a random effects model or a mixed effects model depending on whether the fixed effects are or not. A random effect is a random variable representing the effect of a randomly chosen level from a population of levels that a random factor can assume. Many papers have tried to solve negative values of variance components of random effects included in a linear model, which is an underlying model for the analysis of data from a certain experiment. Henderson (1953) suggested three methods to estimate variance components that came to be known as Henderson's method 1, 2, and 3. Searle (1971) dealt with them in detail. There are many papers devoted to the estimation of variance components; however, it is unfortunate that any method discussed in papers can have negative estimates of variances. Milliken and Johnson (1984) suggested a proposal for eliminating negative estimates. EI-Leithy *et al.* (2016) and Searle *et al.* (2009) discussed the nonnegative estimation of variance components in mixed linear models. This paper proposes a proper method that always yields nonnegative estimates for variance components. It is based on the concept of projections defined on a vector space. The basic idea of projection and related concepts are well-defined by Johnson and Wichern (1988). Milliken and Johnson (1984) used the following data as an example for estimating variance components. It is again used here to compare estimates between the Henderson's method 1 and the method proposed in this paper that produces nonnegative estimates for variance components.

Table 1: Data for two-way design

RC11	RC12	RC13	RC21	RC22	RC23
10	13	21	16	13	11
12	15	19	18	19	13
11				14	

2. Data and model

The data in Table 1 are from Milliken and Johnson (1984). R and C denote the random factors of a two-way treatment structure in a completely randomized design structure.

First, consider a more general two-way random components model for analyzing such a data. The model of general form is given by

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}, \quad i = 1, 2, \ldots, a; \ j = 1, 2, \ldots, b; \ k = 1, 2, \ldots, n_{ij},$$

where y_{ijk} is the response from the kth experimental unit in the treatment combination of ith level of R and j^{th} level of C. μ is the grand mean, α_i is the effect of the ith level of the random factor R and is assumed to be distributed i.i.d. $N(0, \sigma_\alpha^2)$, and β_j is the effect of the jth level of the random factor C, which is assumed to be distributed i.i.d. $N(0, \sigma_\beta^2)$. It also assumes $(\alpha\beta)_{ij}$ the effect of the interaction $R \times C$ and is assumed to be distributed i.i.d. $N(0, \sigma_{\alpha\beta}^2)$, and ϵ_{ijk} is the random error associated with y_{ijk} where it is assumed ϵ_{ijk} are distributed i.i.d. $N(0, \sigma_\epsilon^2)$. These assumptions allow the variances and covariances of the observations to be evaluated. The variance of an observation is

$$\text{Var}(y_{ijk}) = \text{Var}\left(\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}\right)$$
$$= \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2 + \sigma_\epsilon^2.$$

There are four components in the variance of y_{ijk}. The above model can be described in matrix notation as

$$\boldsymbol{y} = \boldsymbol{j}\mu + \boldsymbol{X}_1\boldsymbol{\alpha} + \boldsymbol{X}_2\boldsymbol{\beta} + \boldsymbol{X}_3(\boldsymbol{\alpha\beta}) + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\boldsymbol{j}$ is an $n \times 1$ vector of ones ($n = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}$), $\boldsymbol{X}_1$ is an $n \times a$ coefficient matrix of $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}$ is the $a \times 1$ random vector assumed to have $N(\boldsymbol{0}, \sigma_\alpha^2 \boldsymbol{I}_a)$, $\boldsymbol{X}_2$ is an $n \times b$ coefficient matrix of $\boldsymbol{\beta}$, $\boldsymbol{\beta}$ is the $b \times 1$ random vector assumed to have $N(\boldsymbol{0}, \sigma_\beta^2 \boldsymbol{I}_b)$, $\boldsymbol{X}_3$ is an $n \times ab$ coefficient matrix of $(\boldsymbol{\alpha\beta})$, $(\boldsymbol{\alpha\beta})$ is the $ab \times 1$ random vector assumed to have $N(\boldsymbol{0}, \sigma_{\alpha\beta}^2 \boldsymbol{I}_{ab})$, and $\boldsymbol{\epsilon}$ is the $n \times 1$ random vector assumed to be distributed $N(\boldsymbol{0}, \sigma_\epsilon^2 \boldsymbol{I}_n)$. The covariance matrix of the observation vector $\boldsymbol{y}$ denoted by Σ is

$$\Sigma = \text{Var}(\boldsymbol{y})$$
$$= \text{Var}\left(\boldsymbol{j}\mu + \boldsymbol{X}_1\boldsymbol{\alpha} + \boldsymbol{X}_2\boldsymbol{\beta} + \boldsymbol{X}_3(\boldsymbol{\alpha\beta}) + \boldsymbol{\epsilon}\right)$$
$$= \sigma_\alpha^2 \boldsymbol{X}_1\boldsymbol{X}_1' + \sigma_\beta^2 \boldsymbol{X}_2\boldsymbol{X}_2' + \sigma_{\alpha\beta}^2 \boldsymbol{X}_3\boldsymbol{X}_3' + \sigma_\epsilon^2 \boldsymbol{I}_n.$$

The analysis of the variance method called Henderson's Method 1 can be used for the estimation of variance components. But the problem is that it sometimes yields negative values with which we do not agree as an estimate of variance. This unpleasant thing can be avoided by way of projections. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be vectors in an n dimensional vector space. The projection of $\boldsymbol{y}$ on the vector $\boldsymbol{x}$ denoted by $\boldsymbol{p}_x$ is

$$\boldsymbol{p}_x = \boldsymbol{x}(\boldsymbol{x}'\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{y}.$$

Let X be a model matrix from (2.1) where X is defined as $X = (j, X_1, X_2, X_3)$. Projection of y onto the vector subspace generated by X is

$$
\begin{aligned}
p_X &= X \left(X'X \right)^- X'y \\
&= XX^- y,
\end{aligned}
\tag{2.2}
$$

where X^- is Moore-Penrose generalized inverse and denotes $(X'X)^- X'$. The definition of it is well defined on Graybill (1983) as terminology g-inverse. If X is of full rank, $(X'X)^{-1}$ is used instead of $(X'X)^-$.

3. Projection method

A linear model is called a random components model when it has random components only except μ. Random components may happen from the treatment structure or design structure (or both) in an experiment. Henderson (1953) suggested Method 1 for the estimation of variance components in a random model. An alternative method that always produces nonnegative values is discussed because it allows impermissible estimates as variances. An observation vector y in a vector space can be decomposed into projections that are defined on orthogonal vector subspaces generated by model matrices from fitting sub-models of model (2.1). This implies that the projection defined on each orthogonal vector subspace has its own nonnegative variation not affected by other projections. Henceforth, there is the possibility of finding nonnegative variance estimates. To find nonnegative estimates of variance components in model (2.1) first consider fitting a full model. In fitting (2.1), we get (2.2). After fitting the whole model we get the residual vector r,

$$
\begin{aligned}
r &= (I - XX^-)(X\delta + \epsilon) \\
&= (I - XX^-)\epsilon,
\end{aligned}
\tag{3.1}
$$

where δ denotes $(\mu, \alpha, \beta, (\alpha\beta))'$. From (3.1) we see that the error vector subspace is generated by sub-model matrix $(I - XX^-)$. Let the residual model matrix be X_f. The residual sum of squares can be obtained by the squared distance of the projection of r onto the error space. The residual random vector r has only one random component ϵ which has the information about σ_ϵ^2. The projection of r denoted by p_e is given by

$$
\begin{aligned}
p_e &= [(I - XX^-)][(I - XX^-)]^- r \\
&= X_f X_f^- r,
\end{aligned}
$$

where $X_f = I - XX^-$. So, the residual sum of squares denoted by SS_e is

$$
\begin{aligned}
p_e'p_e &= r'X_f X_f^- r \\
&= y'(I - XX^-)y.
\end{aligned}
\tag{3.2}
$$

XX^- is the projection of y into the vector subspace generated by the model matrix X. Second, think of the residual random model for the estimation of $\sigma_{\alpha\beta}^2$ and σ_ϵ^2. The model is

$$
\begin{aligned}
r_b &= \left(I - X_b X_b^- \right)(X_b\delta_b + X_3(\alpha\beta) + \epsilon) \\
&= \left(I - X_b X_b^- \right) X_3(\alpha\beta) + \epsilon_b,
\end{aligned}
\tag{3.3}
$$

where $X_b = (j, X_1, X_2)$, $\delta_b = (\mu, \alpha, \beta)'$, and $\epsilon_b = (I - X_b X_b^-)\epsilon$. Letting SS_b be the sum of squares adjusted for δ_b, $SS_b = r_b' r_b$. $E(SS_b)$ will reveal information about the two variance components $\sigma_{\alpha\beta}^2$ and σ_ϵ^2. The projection of r_b into the vector space generated by sub-model matrix is revealed as

$$p_{ab} = \left[\left(I - X_b X_b^-\right) X_3\right]\left[\left(I - X_b X_b^-\right) X_3\right]^- r_b$$
$$= X_{\alpha\beta} X_{\alpha\beta}^- r_b,$$

where $X_{\alpha\beta} = (I - X_b X_b^-)X_3$. The sum of squares due to $(\alpha\beta)$ is the squared length of the projection which is expressed by $p_{ab}' p_{ab}$. Then we can see that $SS_b = p_{ab}' p_{ab} + p_e' p_e$. From (3.3), it can be seen the residual random vector r_b has two variance components $\sigma_{\alpha\beta}^2$ and σ_ϵ^2. These are estimated separately from each corresponding vector subspace that is orthogonal to each other. Third, for the estimation of variance component σ_β^2 a model to be fitted is given by

$$r_a = (I - X_a X_a^-)(X_a \delta_a + X_2 \beta + X_3(\alpha\beta) + \epsilon)$$
$$= (I - X_a X_a^-)X_2 \beta + \epsilon_a, \tag{3.4}$$

where $X_a = (j, X_1)$, $\delta_a = (\mu, \alpha)'$, and $\epsilon_a = (I - X_a X_a^-)(X_3(\alpha\beta) + \epsilon)$. It should be noted that r_a has three variance components from (3.4), one of which is estimated from the estimation space and the others are from the error space. Hence, it has the information about σ_β^2, $\sigma_{\alpha\beta}^2$, and σ_ϵ^2. Let SS_a be $r_a' r_a$. The projection denoted by p_b is given as

$$p_b = [(I - X_a X_a^-) X_2][(I - X_a X_a^-) X_2^-] r_a$$
$$= X_\beta X_\beta^- r_a,$$

where $X_\beta = (I - X_a X_a^-)X_2$. The sum of squares due to β is given by the squared distance of the projection $p_b' p_b$. So, $SS_a = p_b' p_b + p_{ab}' p_{ab} + p_e' p_e$. Lastly, the fitting model for σ_α^2 is

$$r_j = \left(I - X_j X_j^-\right)\left(X_j \delta_j + X_1 \alpha + X_2 \beta + X_3(\alpha\beta) + \epsilon\right)$$
$$= \left(I - X_j X_j^-\right)X_1 \alpha + \epsilon_j, \tag{3.5}$$

where $X_j = j$, $\delta_j = \mu$, and $\epsilon_j = (I - X_j X_j^-)(X_2 \beta + X_3(\alpha\beta) + \epsilon)$. Let SS_j be $r_j' r_j$. Since the residual random vector has all four random components, it has the information about σ_α^2, σ_β^2, $\sigma_{\alpha\beta}^2$, and σ_ϵ^2. From (3.5), σ_α^2 can be estimated in the estimation space independently of the others. The projection of r_j into the space generated by $(I - X_j X_j^-)X_1$ is

$$p_a = \left[\left(I - X_j X_j^-\right) X_1\right]\left[\left(I - X_j X_j^-\right) X_1\right]^- r_j$$
$$= X_\alpha X_\alpha^- r_j,$$

where $X_\alpha = (I - X_j X_j^-)X_1$. The sum of squares adjusted for α is $p_a' p_a$. Hence, $SS_j = p_a' p_a + p_b' p_b + p_{ab}' p_{ab} + p_e' p_e$. It should be noted that random residual models are derived in a reverse order using the residual vector. After establishing the model in each step, identify the model matrix associated with the random effects and then derive the corresponding projection at that step using the matrix. Then calculate the sum of squares due to random effects in terms of projection.

4. Expectations of quadratic forms

Variance components are estimated by equating the sums of squares to expected values. The sums of squares are quadratic forms in the observations. Hartleys synthesis (Hartley, 1967) is used for the calculation of the expected values of quadratic forms. The method applies to the calculations of the coefficients of variance components of the expected value of quadratic forms in observations as a variation source. Let Q be a quadratic form in the observation vector y and $q = y'Qy$ where Q is assumed to be a symmetric matrix of constants. Then the expectation of Q is normally given by

$$E\left(y'Qy\right) = \text{Tr}\left[QE\left(yy'\right)\right]$$
$$= \text{Tr}(Q\Sigma) + \mu^2 j'Qj.$$

The matrix of a quadratic form in y for the analysis of variance can be constructed such that $\mu^2 j'Qj = 0$. Hence, the expectation of the quadratic form that does not depend on μ is

$$E(y'Qy) = \text{Tr}(Q\Sigma)$$
$$= \text{Tr}\left[Q\left(\sigma_\alpha^2 X_1 X_1' + \sigma_\beta^2 X_2 X_2' + \sigma_{\alpha\beta}{}^2 X_3 X_3' + \sigma_\epsilon^2 I_n\right)\right]$$
$$= \sigma_\alpha^2 \text{Tr}\left(X_1' Q X_1\right) + \sigma_\beta^2 \text{Tr}\left(X_2' Q X_2\right) + \sigma_{\alpha\beta}^2 \text{Tr}\left(X_3' Q X_3\right) + \sigma_\epsilon^2 \text{Tr}(Q),$$

where $\text{Tr}(Q)$ denotes the trace of the square matrix Q. It should be noted that the $\text{Tr}(X_i' Q X_i)$ can vary at times because of the dependence of sub-model matrices X_i's which are not orthogonal to each other. For the calculation of the coefficients of variance components orthogonal sub-model matrices are used instead of X_i's. First, think of the coefficients of σ_ϵ^2. From the equation of (3.2) the expectation of SS_e is

$$E\left(p_e' p_e\right) = E\left(r'(I - XX^-)(I - XX^-)^- r\right)$$
$$= E\left[y'(I - XX^-)y\right]$$
$$= c_{e1}\sigma_\epsilon^2 + c_{e2}\sigma_{\alpha\beta}^2 + c_{e3}\sigma_\beta^2 + c_{e4}\sigma_\alpha^2, \tag{4.1}$$

where y is supposed of consisting of orthogonal components vectors. First, the coefficient of σ_ϵ^2, c_{e1} is obtained by using X_f, which is the residual model matrix of (3.1). That is,

$$\text{Tr}\left(X_f'\left(I - XX^-\right) X_f\right) = \text{Tr}\left((I - XX^-)'(I - XX^-)(I - XX^-)\right)$$
$$= \text{Tr}(I - XX^-).$$

Second, the coefficient of $\sigma_{\alpha\beta}^2$, c_{e2} is given as

$$\text{Tr}\left(X_{\alpha\beta}'(I - XX^-)X_{\alpha\beta}\right) = \text{Tr}\left(X_3'\left(I - X_b X_b^-\right)(I - XX^-)\left(I - X_b X_b^-\right) X_3\right)$$
$$= 0,$$

which shows $X_{\alpha\beta}$ and $I - XX^-$ are orthogonal matrices. Third, the coefficient of σ_β^2, c_{e3} is given as

$$\text{Tr}\left(X_\beta'(I - XX^-)X_\beta\right) = \text{Tr}\left(X_2'(I - X_a X_a^-)(I - XX^-)(I - X_a X_a^-)X_2\right),$$
$$= 0,$$

where X_β denotes the model matrix of (3.4). It shows X_β and $I - XX^-$ are orthogonal matrices. Fourth, the coefficient of σ_α^2, c_{e4}, is given as

$$\text{Tr}\left(X_\alpha'(I - XX^-)X_\alpha\right) = \text{Tr}\left(X_1'(I - X_aX_a^-)(I - XX^-)(I - X_aX_a^-)X_1\right)$$
$$= 0,$$

where X_α is the model matrix of (3.5). This expression also shows X_α and $I - XX^-$ are orthogonal matrices. All matrices associated with the variance components $\sigma_{\alpha\beta}^2$, σ_β^2, and σ_α^2 are orthogonal to the coefficient matrix of the random error vector ϵ. Hence, (4.1) is reduced to

$$E\left(p_e'p_e\right) = \text{Tr}(I - XX^-)\sigma_\epsilon^2$$
$$= c_{e1}\sigma_\epsilon^2.$$

Next, we compute $E(SS_b)$ having two variance components because the random residual vector r_b has random effects related to $\sigma_{\alpha\beta}^2$ and σ_ϵ^2. $E(SS_b)$ is

$$E\left(r_b'r_b\right) = E\left(y'\left(I - X_bX_b^-\right)y\right)$$
$$= c_{b1}\sigma_\epsilon^2 + c_{b2}\sigma_{\alpha\beta}^2 + c_{b3}\sigma_\beta^2 + c_{b4}\sigma_\alpha^2.$$

First, the coefficient of σ_ϵ^2, c_{b1} is obtained using the model matrix $I - XX^-$ of (3.1) denoted by X_f. That is,

$$\text{Tr}\left(X_f'\left(I - X_bX_b^-\right)X_f\right) = \text{Tr}\left((I - XX^-)\left(I - X_bX_b^-\right)(I - XX^-)\right)$$
$$= \text{Tr}(I - XX^-).$$

Second, the coefficient of $\sigma_{\alpha\beta}^2$, c_{b2} is obtained using the model matrix $(I - X_bX_b^-)X_3$ of (3.3) denoted by $X_{\alpha\beta}$. This turns out to be

$$\text{Tr}\left(X_{\alpha\beta}'\left(I - X_bX_b^-\right)X_{\alpha\beta}\right) = \text{Tr}\left(X_3'\left(I - X_bX_b^-\right)\left(I - X_bX_b^-\right)\left(I - X_bX_b^-\right)X_3\right)$$
$$= \text{Tr}\left(X_3'\left(I - X_bX_b^-\right)X_3\right).$$

Third, the coefficient of σ_β^2, c_{b3} is obtained using the model matrix $(I - X_bX_b^-)X_2$ of (3.4) denoted by X_β. That is,

$$\text{Tr}\left(X_\beta'\left(I - X_bX_b^-\right)X_\beta\right) = \text{Tr}\left(X_2'(I - X_aX_a^-)\left(I - X_bX_b^-\right)(I - X_aX_a^-)X_2\right)$$
$$= 0,$$

where two matrices X_β and $I - X_bX_b^-$ are orthogonal. Lastly, the coefficient of σ_α^2, c_{b4} is obtained using the model matrix $(I - X_jX_j^-)X_1$ of (3.5) denoted by X_α. That is,

$$\text{Tr}\left(X_\alpha'\left(I - X_bX_b^-\right)X_\alpha\right) = \text{Tr}\left(X_1'\left(I - X_aX_a^-\right)\left(I - X_bX_b^-\right)(I - X_aX_a^-)X_1\right)$$
$$= 0.$$

Hence, the expectation of SS_b is represented by

$$E\left(r_b'r_b\right) = \text{Tr}\left(I - XX^-\right)\sigma_\epsilon^2 + \text{Tr}\left(X_3'\left(I - X_bX_b^-\right)X_3\right)\sigma_{\alpha\beta}^2$$
$$= c_{b1}\sigma_\epsilon^2 + c_{b2}\sigma_{\alpha\beta}^2.$$

Since r_a defined on (3.4) has three random effects, $E(SS_a)$ has the information about the coefficients of σ^2_β, $\sigma^2_{\alpha\beta}$, and σ^2_ϵ. $E(SS_a)$ is

$$E\left(r'_a r_a\right) = E\left(y'\left(I - X_a X_a^-\right) y\right)$$
$$= c_{a1}\sigma^2_\epsilon + c_{a2}\sigma^2_{\alpha\beta} + c_{a3}\sigma^2_\beta + c_{a4}\sigma^2_\alpha. \tag{4.2}$$

First, the coefficient of σ^2_ϵ, c_{a1} is obtained using the model matrix of (3.1) denoted by X_f. That is,

$$\text{Tr}\left(X'_f\left(I - X_a X_a^-\right) X_f\right) = \text{Tr}\left((I - XX^-)(I - X_a X_a^-)(I - XX^-)\right)$$
$$= \text{Tr}(I - XX^-).$$

Second, the coefficient of $\sigma^2_{\alpha\beta}$, c_{a2} is obtained using the model matrix of (3.3) denoted by $X_{\alpha\beta}$. That is,

$$\text{Tr}\left(X'_{\alpha\beta}\left(I - X_a X_a^-\right) X_{\alpha\beta}\right) = \text{Tr}\left(X'_3\left(I - X_b X_b^-\right)(I - X_a X_a^-)\left(I - X_b X_b^-\right) X_3\right)$$
$$= \text{Tr}\left(X'_3\left(I - X_b X_b^-\right) X_3\right).$$

Third, the coefficient of σ^2_β, c_{a3} is found using the model matrix of (3.4) denoted by X_β. That is,

$$\text{Tr}\left(X'_\beta\left(I - X_a X_a^-\right) X_\beta\right) = \text{Tr}\left(X'_2\left(I - X_a X_a^-\right)(I - X_a X_a^-)(I - X_a X_a^-) X_2\right)$$
$$= \text{Tr}\left(X'_2\left(I - X_a X_a^-\right) X_2\right),$$

where two matrices X_β and $I - X_b X_b^-$ are orthogonal. Lastly, the coefficient of σ^2_α, c_{a4} is obtained using the model matrix of (3.5) denoted by X_α. That is,

$$\text{Tr}\left(X'_\alpha\left(I - X_a X_a^-\right) X_\alpha\right) = \text{Tr}\left(X'_1\left(I - X_a X_a^-\right)(I - X_a X_a^-)(I - X_a X_a^-) X_1\right)$$
$$= \text{Tr}\left(X'_1\left(I - X_a X_a^-\right) X_1\right)$$
$$= 0.$$

Hence, $E(SS_a)$ is represented by

$$E\left(r'_a r_a\right) = \text{Tr}\left(I - XX^-\right)\sigma^2_\epsilon + \text{Tr}\left(X'_3\left(I - X_b X_b^-\right) X_3\right)\sigma^2_{\alpha\beta} + \text{Tr}\left(X'_2\left(I - X_a X_a^-\right) X_2\right)\sigma^2_\beta$$
$$= c_{a1}\sigma^2_\epsilon + c_{a2}\sigma^2_{\alpha\beta} + c_{a3}\sigma^2_\beta.$$

Finally, $E(SS_j)$ is

$$E\left(r'_j r_j\right) = E\left(y'\left(I - X'_j X_j^-\right) y\right)$$
$$= c_{j1}\sigma^2_\epsilon + c_{j2}\sigma^2_{\alpha\beta} + c_{j3}\sigma^2_\beta + c_{j4}\sigma^2_\alpha.$$

The coefficients of variance components of (4.2) can be obtained similarly as before with the model matrices X_f, $X_{\alpha\beta}$, X_β, and X_α. With the X_f, the coefficient of σ^2_ϵ, c_{j1} is given as

$$\text{Tr}\left(X'_f\left(I - X_j X_j^-\right) X_f\right) = \text{Tr}\left((I - XX^-)\left(I - X_j X_j^-\right)(I - XX^-)\right)$$
$$= \text{Tr}\left(I - XX^-\right).$$

With the $X_{\alpha\beta}$, the coefficient of $\sigma_{\alpha\beta}^2$, c_{j2} is given as

$$\mathrm{Tr}\left(X'_{\alpha\beta}\left(I - X_j X_j^-\right)X_{\alpha\beta}\right) = \mathrm{Tr}\left(X'_3\left(I - X_b X_b^-\right)\left(I - X_j X_j^-\right)\left(I - X_b X_b^-\right)X_3\right)$$
$$= \mathrm{Tr}\left(X'_3\left(I - X_b X_b^-\right)X_3\right).$$

With the X_β, the coefficient of σ_β^2, c_{a3} is given as

$$\mathrm{Tr}\left(X'_\beta\left(I - X_j X_j^-\right)X_\beta\right) = \mathrm{Tr}\left(X'_2\left(I - X_a X_a^-\right)\left(I - X_j X_j^-\right)\left(I - X_a X_a^-\right)X_2\right)$$
$$= \mathrm{Tr}\left(X'_2\left(I - X_a X_a^-\right)X_2\right).$$

With the X_α, the coefficient of σ_α^2, c_{j4} is given as

$$\mathrm{Tr}\left(X'_\alpha\left(I - X_j X_j^-\right)X_\alpha\right) = \mathrm{Tr}\left(X'_1\left(I - X_a X_a^-\right)\left(I - X_j X_j^-\right)\left(I - X_a X_a^-\right)X_1\right)$$
$$= \mathrm{Tr}\left(X'_1\left(I - X_a X_a^-\right)X_1\right).$$

Hence, the expectation of SS_a is represented by

$$E\left(r'_j r_j\right) = \mathrm{Tr}\left(I - X X^-\right)\sigma_\epsilon^2 + \mathrm{Tr}\left(X'_3\left(I - X_b X_b^-\right)X_3\right)\sigma_{\alpha\beta}^2$$
$$+ \mathrm{Tr}\left(X'_2\left(I - X_a X_a^-\right)X_2\right)\sigma_\beta^2 + \mathrm{Tr}\left(X'_1\left(I - X_a X_a^-\right)X_1\right)\sigma_\alpha^2$$
$$= c_{j1}\sigma_\epsilon^2 + c_{j2}\sigma_{\alpha\beta}^2 + c_{j3}\sigma_\beta^2 + c_{j4}\sigma_\alpha^2. \tag{4.3}$$

The equations solving for variance components are

$$SS_e = c_{e1}\hat\sigma_\epsilon^2 + c_{e2}\hat\sigma_{\alpha\beta}^2 + c_{e3}\hat\sigma_\beta^2 + c_{e3}\hat\sigma_\alpha^2,$$
$$SS_b = c_{b1}\hat\sigma_\epsilon^2 + c_{b2}\hat\sigma_{\alpha\beta}^2 + c_{b3}\hat\sigma_\beta^2 + c_{b4}\hat\sigma_\alpha^2,$$
$$SS_a = c_{a1}\hat\sigma_\epsilon^2 + c_{a2}\hat\sigma_{\alpha\beta}^2 + c_{a3}\hat\sigma_\beta^2 + c_{a4}\hat\sigma_\alpha^2,$$
$$SS_j = c_{j1}\hat\sigma_\epsilon^2 + c_{j2}\hat\sigma_{\alpha\beta}^2 + c_{j3}\hat\sigma_\beta^2 + c_{j4}\hat\sigma_\alpha^2.$$

The solutions from a system of linear equations in (σ_i^2)'s (4.3) are given as nonnegative estimates of variance components for the assumed model (2.1).

5. Example

In this section the data presented in Section 2 is used as an example of getting nonnegative estimates for four variance components. From the model fitting procedure discussed previously we get $SS_e = 30.66667$, $SS_b = 139.81176$, $SS_a = 154.57143$, and $SS_j = 155.2143$. The expectations of these are given as

$$E(SS_e) = 8\sigma_\epsilon^2$$
$$E(SS_b) = 8\sigma_\epsilon^2 + 4.52\sigma_{\alpha\beta}^2$$
$$E(SS_a) = 12\sigma_\epsilon^2 + 4.52\sigma_{\alpha\beta}^2 + 9.14\sigma_\beta^2$$
$$E(SS_j) = 8\sigma_\epsilon^2 + 4.52\sigma_{\alpha\beta}^2 + 9.14\sigma_\beta^2 + 7\sigma_\alpha^2.$$

By equating the sums of squares to their respective expected sums of squares, we get

$$\hat{\sigma}_\epsilon^2 = 3.8333, \qquad \hat{\sigma}_{\alpha\beta}^2 = 24.1597, \qquad \hat{\sigma}_\beta^2 = 1.6143, \qquad \hat{\sigma}_\alpha^2 = 0.0918.$$

Nonnegative estimates are obtained by the projection method. These values are compared with the Type I estimates from Milliken and Johnson (1984) which are

$$\hat{\sigma}_\epsilon^2 = 3.8333, \qquad \hat{\sigma}_{\alpha\beta}^2 = 22.4620, \qquad \hat{\sigma}_\beta^2 = -10.5877, \qquad \hat{\sigma}_\alpha^2 = -8.0329.$$

6. Conclusion

This paper suggests a method to estimate variance components in a two-way random model. The method is based on the concept of an orthogonal projection of an observation vector in n-dimensional space onto the vector subspace of itself. Orthogonal projections are used for the estimation of variance components associated with the random effects of the assumed model. There are many methods that are available in the literature; however, most still have the problem of negative estimates of variance components when it happens. To avoid this, it is necessary to build proper models to obtain right projections onto vector subspaces that are orthogonal to each other. Therefore, it is important to identify the residual random vector in each step and then construct a residual random model for it.

The corresponding orthogonal projection is obtained by projecting the residual random vector onto the vector subspace spanned by the residual model matrix of the residual random model. The observation vector can be denoted as the sum of projections since projections are defined on their own vector subspaces that are orthogonal.

Hartley's synthesis was used for the calculations of the expectations of sums of squares expressed as sums of squared distances of projections that are orthogonal. Each projection is defined on its own vector subspace orthogonal to others; therefore, the expectation on the sums of squares for the residual vector defined on each stage has nonzero coefficients for variance components only if it has related random effects.

Variance as a measure of variation of data can never be a negative value from its definition; Nevertheless, we accredit the negative estimate as evidence that the true value of the component is zero as is often. However, the projection method shows that the estimates of variance components are always nonnegative.

Acknowledgements

References

EI-Leithy AH, Abdel-Wahed AZ, and Abdallah SM (2016). On non-negative estimation of variance components in mixed linear models, *Journal of Advanced Research*, **7**, 59–68.

Graybill FA (1983). *Matrices with Applications in Statistics*, Wadsworth, California.

Hartley HO (1967). Expectations, variances and covariances of ANOVA mean squares by "synthesis", *Biometrics*, **23**, 105–114.

Henderson CR (1953). Estimation of variance and covariance components, *Biometrics*, **9**, 226–252.

Johnson RA and Wichern DW (1988). *Applied Multivariate Statistical Analysis*, Prentice-Hall, New Jersey.

Milliken GA and Johnson DE (1984). *Analysis of Messy Data*, Van Nostrand Reinhold, New York.

Searle SR (1971). *Linear Models*, John Wiley & Sons, New York.

Searle SR, Casella G, and McCulloch CE (2009). *Variance Components*, John Wiley & Sons, New York.