Current Trends on Biostatistics & Biometrics

DOI: 10.32474/CTBB.2020.02.000140

ISSN: 2644-1381

Research Article

Non-Negative Variance Component Estimation for Mixed-Effects Models

Jaesung Choi*

Department of Statistics, Keimyung University, Korea

***Corresponding author:** Department of Statistics, Keimyung University, Korea

Received: February 27, 2020 **Published:** March 11, 2020

Abstract

This paper discusses projection methods to find nonnegative estimates for variance components of random effects in mixed models. The proposed methods are based on the concepts of projections, which are called projection method I, II and III. These three methods produce the same nonnegative estimates for the same data. Even though each method uses orthogonal projections in its own way, the results are the same for the variance components regardless of which method is used. It is shown that quadratic forms of an observation vector are constructed differently in each method. All sums of squares in quadratic forms of the observation vector can be expressed as squared distances of corresponding projections. A projection model is defined and used to evaluate expected values of quadratic forms of observations that are associated with variance components. Hartley's synthesis is used as a method for finding the coefficients of variance components.

Keywords: Mixed model; projections; quadratic forms; random effects; synthesis

Introduction

Much literature has been devoted to the estimation of variance components in random effects or mixed effects models. A variance component should always be nonnegative by its definition; however, we sometimes get it as negative [1,2]. illustrated this with the simple hypothetical data of a one-way classification having three observations in two classes and insisted that there was nothing intrinsic in the analysis of variance method to prevent it. When a negative estimate happens, it is not easy to handle this situation in interpretation and action. Hence, many papers have been contributed to strategies to deal with the negative values as estimates of variance components [3]. suggests that negative estimates of variance components can occur in certain designs such as split plot and randomized block designs by random- inaction. Thompson discusses the interpretation of the negative estimate and suggests an alternative method when the analysis of variance method yields negative estimates [4]. also suggest a procedure for eliminating negative estimates of variance components in random effects models. The analysis of the variance method is almost exclusively applied to balanced data for estimating variance components. However, there are multiple methods for unbalanced data. Therefore, it is necessary to identify the types of data before choosing a method. Though balanced data have the same numbers of observations in each cell, unbalanced data have unequal numbers of observations in the subclasses made by the levels of classification factors. Depending on the types of data, many methods can be applied to the estimation of variance components in a vector space. Representing data as vectors, the vector space of an observation vector can be partitioned in many ways, depending on the data structure. For balanced data, the vector space can always be partitioned into orthogonal vector subspaces according to the sources of variation, but it is not true for unbalanced data. This is the main difference between balanced and unbalanced data from the view point of a vector space. 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, while a fixed effect is an unknown constant denoting the effect of a predetermined level of a factor. A linear model with these two types of effects is called a mixed effects model. The primary concern with the model in this paper is naturally in the nonnegative estimation of variance components of random effects. A negative estimate can happen in any method that contributes to the estimation.

Hence, many papers have investigated strategies for interpretation and alternatives. Such strategies are seen in [5-9]. However, it is necessary to have a method that yields nonnegative estimates despite all such efforts [10].suggested a method that uses reductions in sums of squares due to fitting both the full model and different sub-models of it for estimating variance components of random effects in mixed models. This method is called the fitting constants method or Henderson's Method 3. Even though it has been used extensively for the estimation of variance components in mixed models, it still has some defects producing negative estimates in case [11]. synthesis is also used for calculating the coefficients of variance components in the method. Although this method is very useful, we should recognize whether quadratic forms for variance components are in the right form or not. Otherwise, expectations of the quadratic forms can be different from the real ones. This is going to be discussed in detail in projection model building. This paper suggests three methods to produce nonnegative estimates for variance components in mixed models. They are based on the concept of projection defined on a vector space. The definition of a projection and its related concepts are discussed in [12,13]. Quadratic forms in the observations can be obtained as squared distances of projections defined in proper vector subspaces. Each method requires that all vector subspaces for projections should be orthogonal to each other at the stage of fitting sub-models serially. When the orthogonality is satisfied with vector subspaces, it is possible to get nonnegative estimates. Hence, we also discuss how to construct orthogonal vector subspaces from a given mixed model. Quadratic forms as sums of squares due to random effects are then used to evaluate their expected values. Hereafter, equating quadratic forms to their expected values represents available equations for the estimates. For calculating the coefficients of variance components, Hartley's synthesis is applied but in a different manner, which will be discussed.

Mixed Models

Mixed models are used to describe data from experimental situations where some factors are fixed, and others are random. When two types of factors are considered in experiments, one is interested in both parts, that is, the fixed-effects part and the random-effects part, in models. Let α be a vector of all the fixed effects except μ in a mixed model and let δ_i denote a set of random effects for random factor i for i = 1, 2, r. Then, δ_i could be interaction effects or nested-factor effects when they are simply regarded as effects from random factors. The matrix notation of the mixed model for an observation vector y is

$$y = j\mu + X_F\alpha_F + X_R\alpha_R + \in$$

$$= j\mu + X_F\alpha_F + \sum_{i=1}^{r} X_i\delta_i + \in \qquad (1)$$

where $j\mu + X_F\alpha_F$ is the fixed part of the model and $X_R\delta_R + \in$ is the random part of the model. δ_i s are assumed to be independent and identically distributed as $N(0,\sigma^2_{\delta_i} I)$, and ∈ is assumed to be distributed as $N(0,\sigma^2_{\in} I)$. The mean and variance of y from (1) is

$$E(y) = j\mu + X_F\alpha_F$$

$$\Sigma = \mathrm{var}(y) = \sum_{i=1}^{\gamma} X_i \,\mathrm{var}(\delta_i) X_i^T + \sigma^2_{\in} I \qquad (2)$$

The expected value of the quadratic form y_T, Q_y is

$$E(y^T Qy) = tr(Q\Sigma) + E(y)^T QE(y) \qquad (3)$$

Substituting the terms of (2) for (3) is

$$E\left(y^T Qy\right) = \sum_{i=1}^{r} \sigma^2_{\delta_i} tr\left(QX_iX_i^T\right) + \sigma^2_{\in} tr(Q) + E\left(j\mu + X_F\alpha_F\right)^T QE\left(j\mu + X_F\alpha_F\right). \qquad (4)$$

The expectation of any quadratic form in the observations of a vector y is represented as a function of variance components and fixed effects. The variance components of the full model can be estimated by the fitting constants method of using reductions in the sums of squares due to fitting the full model and the sub-model of it. This method provides unbiased estimators of the variance components that do not depend on any fixed effects in the model, and it has been widely used for the estimation of variance components for unbalanced data. However, it still has an unsolved problem having negative solutions as estimates. As an alternative, a method which is based on the concepts of projections is suggested [14]. To discuss it, we consider the model (1) as representative. Since there are two parts in the model, we naturally divide the model into a fixed part and a random part. The random part of the model consists of random effects and errors:

$$y = j\mu + X_F\alpha_F + \in_R$$

$$=(j, X_F)(\mu, \alpha_F)^T + \in_R \qquad (5)$$

where $\in_R = \sum_{i=1}^{r} X_i\delta_i + \in$ The general mean μ and fixed effects α_F of (5) can be estimated from normal equations. Regarding y as an observation vector in the n-dimensional vector space, it can be decomposed into two component vectors orthogonal to each other. The decomposition of y is done by projecting y onto the vector subspace generated by (j, X_F).

Projection method

Since a method based on the concept of projection is discussed, it will be called the projection method. For a mixed model such as (5), we can decompose y into two components by means of projections. Denoting (j, X_F) and $(\mu, \alpha_F)^T$ by X_m^y and μ_m, respectively, the projection of y onto the vector subspace spanned by X_m is $X_m\ X^-_m y$, where X_m^- denotes a Moore-Penrose generalized inverse of XM . Then, y can be decomposed into two vectors, that is, $X_m X^-_m y$ and $\left(I - X_M X_M^-\right) y$ which are orthogonal [15,16]. Instead of the fitting constants method, the projection method is attempted to estimate the nonnegative estimates of the variance components in a mixed model. To explain the method simply, suppose there are two factors A and B for a two-way cross-classified unbalanced data where A is fixed with a levels and B is random with b levels. The model for this is

$$y = j\mu + X_F\alpha_F + X_\beta\alpha_\beta + X_{\alpha\beta}\alpha_{\alpha\beta} + \in$$
$$= X_M\alpha_M + \in_M, \quad (6)$$

where y is an observation vector in the n dimensional vector space, α_F is a vector of fixed effects of A, δ_β and $\delta_{\alpha\beta}$ represent vectors of random effects of B and AB interaction respectively, and $X_M = (j, X_F)$, $\alpha_M = (\mu, \alpha_F)^T$ and $\in_M = X_\beta\delta_\beta + X_{\alpha\beta}\delta_{\alpha\beta} + \in$. The second $\in$expression of (6) represents the fixed-effects part and the random part. The random part S_M is obtained by the projection of y onto a vector subspace generated by the X_M, which is $\left(I - X_M X_M^-\right)y$. So, y is represented as

$$y = X_M X_M^- y + \left(I - X_M X_M^-\right)y$$
$$= y_M + e_M \quad (7)$$

where $y_M = X_M X_M^- y$ satisfies the two conditions for being the projection of y onto a vector subspace spanned by the columns of X_M. The projection should be obtained by the orthogonal projection to the subspace and denoted as a linear combination of the column vectors of $X_M \cdot X_M X_M^- y$ of (7) satisfies the conditions. Since y_M is orthogonal to e_M, the random part $e_M = \left(I - X_M X_M^-\right)y$ is not affected by the fixed effects and has all the information about the variance components and random error variance. Since there are two random effects and random error terms in the model of (6), we can use e_M for finding the related variance components. The model for the estimation of σ_β^2 is

$$e_M = \left(I - X_M X_M^-\right)y$$
$$= X_B\delta_\beta + \in_\beta, \quad (8)$$

where $X_B = \left(I - X_M X_M^-\right)X_\beta$ and $\in_\beta = \left(I - X_M X_M^-\right)\left(X_{\alpha\beta}\delta_{\alpha\beta} + \in\right)$. The projection of e_M onto the subspace spanned by X_B is $X_B X_B^- e_M$, which is $\left[\left(I - X_M X_M^-\right)X_\beta\right]\left[\left(I - X_M X_M^-\right)X_\beta\right]^- e_M$. Then,

$$e_M = X_B X_B^- e_M + \left(I - X_B X_B^-\right)e_M$$
$$= y_B + e_B, \quad (9)$$

where $y_B = X_B X_B^- e_M$ is the projection of e_M onto the column space of $X_B \cdot y_B$ and e_B are orthogonal each other. Hence, e_B is not affected by the random effects δ_B Therefore, e_B is used for finding the subspace that has information about $\sigma_{\alpha\beta}^2$. The model for this is

$$e_B = \left(I - X_B X_B^-\right)e_M = X_{AB}\delta_{\alpha\beta} + \in_{\alpha\beta}, \quad (10)$$

where $X_{AB} = \left(I - X_M X_M^- - X_B X_B^-\right)X_{\alpha\beta}$ and $\in_{\alpha\beta} = \left(I - X_M X_M^- - X_B X_B^-\right)\in$ Hence, the projection of eB onto the subspace generated by X_{AB} is $y_{AB} = X_{AB}X_{AB}^- e_B$. Then,

$$e_B = X_{AB}X_{AB}^- e_B + (I - X_{AB}X_{AB}^-)e_B$$
$$= y_{AB} + e_B, \quad (11)$$

where e_{AB} is $(I - X_{AB}X_{AB}^-)e_B$. Finally, we can use e_{AB} for finding the coefficient matrix of the random error vector which generates the error space orthogonal to all the other spaces.

$$e_{AB} = (I - X_{AB}X_{AB}^-)e_B$$
$$= \left(I - X_M X_M^- - X_B X_B^- - X_{AB}X_{AB}^-\right)\in \quad (12)$$

Thus, we can know that e_{AB} has all the information about $\sigma_\in^2$ of the random error vector $\in$. Denoting y as the sum of orthogonal projections and error part,

$$y = y_M + y_B + y_{AB} + e_{AB}$$
$$= X_M X_M^- y + X_B X_B^- e_M + X_{AB}X_{AB}^- e_B + \left(I - X_{AB}X_{AB}^-\right)e_B \quad (13)$$

Each term of (13) can be used to calculate the sums of squares that are quadratic forms in the observations. Since y is partitioned as four terms, there are four available sums of squares. We denote them SS_M, SS_B, SS_{AB} and SS_E where subscripts are corresponding factors. They are defined as

$$SS_M = y^T X_M X_M^- y$$
$$SS_B = y^T (I - X_M X_M^-) X_B X_B^- (I - X_M X_M^-) y$$
$$SS_{AB} = y^T (I - X_M X_M^- - X_B X_B^-) X_{AB} X_{AB}^- (I - X_M X_M^- - X_B X_B^-) y$$
$$SS_E = y^T X_{Ey} \quad (14)$$

where each SS term is given as the squared length of the projection of y onto its own vector subspace, and $X_E = \left(I - X_M X_M^- - X_B X_B^- - X_{AB}X_{AB}^-\right)$. All the sums of squares are evaluated by using the eigenvalues and eigenvectors of the projection matrices associated with the quadratic forms in y. Since projections are defined on subspaces that are orthogonal to each other, we can identify the coefficient matrices spanning them.

Projection model

Since y is made up by the sum of mutual orthogonal projections such as (13), y can be represented by the orthogonal coefficient's matrices of the effects of the assumed model (6). Temporarily, we denote y as y_p for differentiating the model based on projections to the classical model (6). Then, the model for y_p is

$$y = X_M\alpha_M + X_B\alpha_\beta + X_{AB}\delta_{\alpha\beta} + X_E \in \quad (15)$$

where $y_p = y$. Since each coefficient matrix of the effects is derived from the corresponding orthogonal projection, the equation of (15) defines a projection model that is different from a classical two-way linear mixed model (6). It is useful for evaluating the coefficients of the variance components in the expectations of the quadratic form of an observation vector y_p. In the model, all the coefficient matrices are orthogonal to each other. δ_β, $\delta_{\alpha\beta}$ and s are assumed to be $N(0,\sigma_\beta^2 I_b)$, $N(0,\sigma_{\alpha\beta}^2 I_{ab})$ and $N(0,\sigma_\in^2 I_n)$ respectively. The expectation and the covariance matrix of y_p of the projection model (15) is

$$E(y_p) = X_M\alpha_M$$

$$\Sigma = \sigma_\beta^2 X_B X_B^T + \sigma_{\alpha\beta}^2 X_{AB} X_{AB}^T + \sigma_\varepsilon^2 X_E X_E^T \quad (16)$$

Expectations of the SS terms except SS_M of (14) are

$$E(SS_B) = \sigma_\beta^2 tr\left(X_B^T X_B\right), E(SS_{AB}) = \sigma_{\alpha\beta}^2 tr\left(X_{AB}^T X_{AB}\right)$$

$$E(SS_E) = \sigma_\in^2 tr(X_E) \quad (17)$$

Equating the three sums of squares, SS_B, SS_{AB}, and SS_E of (14) to their correspond- ing expectations leads to linear equations in the variance components, the solutions to which are taken as the estimators of those components. Now, the equations are

$$SS_{AB} = \hat{\sigma}_\beta^2 t\gamma\left(X_B^T X_B\right), SS_{AB} = \hat{\sigma}_{\alpha\beta}^2 t\gamma\left(X_{AB}^T X_{AB}\right),$$
$$SS_E = \hat{\sigma}_\beta^2 t\gamma(X_E). \quad (18)$$

Solutions from the linear equations (18) are nonnegative estimates of the variance components. Since there are three different ways of getting sums of squares by means of projections, we will differentiate them with projection method I, II, and III. The procedure using the system of linear equations like (18) is called projection method I. The projection method II uses residual vectors after projecting y onto orthogonal subspaces. That is, e_M , e_B, and e_{AB} are used such as. Then,

$$e_M = \left(I - X_M X_M^-\right) y_p$$
$$= \left(I - X_M X_M^-\right)\left(X_M\alpha_M + X_B\delta_\beta + X_{AB}\delta_{\alpha\beta} + X_E \in\right) \quad (19)$$

Since e_M has three random components, $e_M^T e_M$ in the quadratic form of y_p in which the coefficients matrices of the projection model are orthogonal is available for esti- mating their variance components. Denoting $e_M^T e_M$ as RSS_M ,

$$RSS_M = e_M^T e_M, \quad (20)$$

where RSS_M measures the variation due to the three random effects, and thus, the quantity is used for the estimation of three variance components σ_β^2, $\sigma_{\alpha\beta}^2$, and $\sigma_\in^2$. Representing the residual random vector e_B as y_p, has two random components as follows.

$$e_B = \left(I - X_B X_B^-\right) e_M$$
$$= \left(I - X_M X_M^- - X_B X_B^-\right)\left(X_{AB}\delta_{\alpha\beta} + X_E \in\right) \quad (21)$$

Hence, $e_B^T e_B$ is used as an variation quantity for two random effects vectors. Denoting $e_B^T e_B$ as RSS_B,

$$RSS_B = e_B^T e_B$$
$$= y_p^T\left(I - X_M X_M^- - X_B X_B^-\right) y_p \quad (22)$$

where RSS_B is used for estimating the two variance components $\sigma_{\alpha\beta}^2$ and $\sigma_\in^2$ since e_B has just two random effects. Finally, expressing e_{AB} as y_p,

$$e_{AB} = \left(I - X_{AB} X_{AB}^-\right) e_B = X_E \in \quad (23)$$

which has just one random component s. Therefore, $e_{AB}^T e_{AB}$ shows the variation due to the random error vector only, and this quantity is used for estimating the variance component $\sigma_\in^2$. Denoting $e_{AB}^T e_{AB}$ as RSS_{AB},

$$RSS_{AB} = e_{AB}^T e_{AB} \quad (24)$$

Hence, RSS_M, RSS_B, and RSS_{AB} are another set of sums of squares for estimating variance components instead of using sums of squares derived from the projections as an alternative method. RSS_M, RSS_B, and RSS_{AB} are also evaluated by using the eigenvalues and eigenvectors of the projection matrices associated with the quadratic forms in y. Now, the expected values of the RSS's are

$$E(RSS_M) = tr\left(\left(I - X_M X_M^-\right)\Sigma\right)$$
$$= \sigma_\beta^2\gamma M_B + \sigma_{\alpha\beta}^2\gamma M_{AB} + \sigma_\in^2\gamma M_E$$
$$E(RSS_B) = tr\left(\left(I - X_B X_B^-\right)\left(I - X_M X_M^-\right)\Sigma\right)$$
$$= \sigma_{\alpha\beta}^2\gamma B_{AB} + \sigma_\in^2\gamma B_E$$
$$E(RSS_{AB}) = tr\left(\left(I - X_{AB} X_{AB}^-\right)\left(I - X_B X_B^-\right)\left(I - X_M X_M^-\right)\Sigma\right)$$
$$= \sigma_\in^2\gamma AB_E \quad (25)$$

Then, the linear equations of variance components are obtained by equating the RSS's to their expected values, the solutions for which always produce nonlinear estimates.

That is,

$$RSS_M = \hat{\sigma}_\beta^2\gamma M_B + \hat{\sigma}_{\alpha\beta}^2\gamma M_{AB} + \hat{\sigma}_\in^2\gamma M_E$$
$$RSS_B = \hat{\sigma}_{\alpha\beta}^2\gamma B_{AB} + \hat{\sigma}_\beta^2\gamma B_E$$
$$RSS_{AB} = \hat{\sigma}_\in^2\gamma AB_E \quad (26)$$

Even though two systems of linear equations are not the same, either system will produce the same estimates of the variance components that are nonnegative. As another method, projection method III is also available for the estimation of variance components. This method is done as follows. For the model of (6), $y = X\theta + \in$, where $X = \left(j, X_F, X_\beta, X_{\alpha\beta}\right)$ and $\theta = \left(\mu, \alpha_F, \delta_\beta, \delta_{\alpha\beta}\right)^T$. This method splits the vector space of an observation vector into two subspaces, one for the projection part and the other for the error part at each step. Then, the projection of y onto the subspace spanned by XX^- is given by XX^-y, and the error vector in the error vector space is $\left(I-XX^-\right)y$. Therefore, the coefficient matrix of s is derived as $\left(I-XX^-\right)$ from it. The quadratic form $y'\left(I-XX^-\right)y$ denoted by BSS_0 is the sum of squares due to random error only, which has all the information about $\sigma_\in^2$. For information about both $\sigma_{\alpha\beta}^2$ and $\sigma_\in^2$, the vector space of the observation vector can be decomposed into two parts one for the projection part and the other for the error part. For this, the model to be fitted is $y = X_1\theta_1 + \in_1$, where $X_1 = \left(j, X_F, X_\beta\right)$, $\theta_1 = \left(\mu, \alpha_F, \delta_\beta\right)$ and $\in_1 = X_{\alpha\beta}\delta_{\alpha\beta} + \in$. Then, the projection of y onto the subspace spanned by $X_1X_1^-$ is given by $X_1X_1^-y$, and the error vector in the errror vector space is $\left(I-X_1X_1^-\right)y$. The quadratic form $y^T\left(I-X_1X_1^-\right)y$ denoted by BSS_1 has information about $\sigma_{\alpha\beta}^2$ and $\sigma_\in^2$. Now, the error vector is represented by

$$\left(I-X_1X_1^-\right)y = \left(I-X_1X_1^-\right)\left(X_1\theta_1 + \in_1\right)$$
$$= \left(I-X_1X_1^-\right)X_{\alpha\beta}\delta_{\alpha\beta} + \left(I-X_1X_1^-\right)\in \quad (27)$$

Hence, the coefficient matrix of $\delta_{\alpha\beta}$ is given by $\left(I-X_1X_1^-\right)X_{\alpha\beta}$. For information about three variance components $\sigma_\in^2$, $\sigma_{\alpha\beta}^2$ and σ_β^2, the vector space can be divided into two subspaces considering the model matrix of the equation, $y = X_2\theta_2 + \in_2$, where $X_2 = \left(j, X_F\right)$, $\theta_2 = \left(\mu, \alpha_F\right)^T$ and $\in_2 = X_\beta\delta_\beta + X_{\alpha\beta}\delta_{\alpha\beta} + \in$. Then, the projection of y onto the subspace spanned by $X_2X_2^-$ is given by $X_2X_2^-y$, and the error vector in the error vector space is $\left(I-X_2X_2^-\right)y$. The quadratic form $y'\left(I-X_2X_2^-\right)y$ denoted by BSS_2 has information about $\sigma_\in^2$, $\sigma_{\alpha\beta}^2$ and σ_β^2. Now, the error vector is represented by

$$\left(I-X_2X_2^-\right)y = \left(I-X_2X_2^-\right)\left(X_2\theta_2 + \in_2\right)$$
$$= \left(I-X_2X_2^-\right)\left(X_\beta\delta_\beta + X_{\alpha\beta}\delta_{\alpha\beta} + \in\right) \quad (28)$$

Hence, the coefficient matrix of δ_β is given by $\left(I-X_2X_2^-\right)X_\beta$. It is necessary to evaluate the expected values of the quadratic forms for constructing the equations for the variance components. They are

$$E\left(BSS_2\right) = \sigma_\beta^2 c_{2\beta} + \sigma_{\alpha\beta}^2 c_{2\alpha\beta} + \sigma_\in^2 c_{2\in}$$
$$E\left(BSS_1\right) = \sigma_{\alpha\beta}^2 c_{2\alpha\beta} + \sigma_\in^2 c_{2\in}$$
$$E\left(BSS_0\right) = \sigma_\in^2 c_{2\in} \quad (29)$$

The nonnegative estimates of variance components are given as solutions of linear equations of $\hat{\sigma}_\beta^2$, $\hat{\sigma}_{\alpha\beta}^2$ and $\hat{\sigma}_\in^2$. The above equations are summarized as follows:

$$BSS_2 = \hat{\sigma}_\beta^2 c_{2\beta} + \hat{\sigma}_{\alpha\beta}^2 c_{2\alpha\beta} + \hat{\sigma}_\in^2 c_{2\in};$$
$$BSS_2 = \hat{\sigma}_{\alpha\beta}^2 c_{2\alpha\beta} + \hat{\sigma}_\in^2 c_{2\in};$$
$$BSS_0 = \hat{\sigma}_\in^2 c_{2\in}, \quad (30)$$

where c_{ij}'s are coefficients of variance components of expected values of quadratic forms of (29).

Examples

As a first example of nonnegative estimates of random effects for a two-way mixed model, Montgomery (2013)'s data are illustrated. The data are from an experiment for a gauge capability study where parts are randomly selected, and three operators are fixed. An instrument or gauge is used to measure a critical dimension on a part. Twenty parts have been selected from the production process, and only three operators are assumed to use the gauge. The assumed model for the data in Table1 is $y_{ijk} = \mu + \alpha_i + \gamma_j + \left(\alpha\gamma\right)_{ij} + \in_{ijk}$, where they $\alpha_i\left(i = 1,2,3\right)$ are fixed effects such that $\Sigma_{i=1}^3 \alpha_i = 0$ and $\gamma_j\left(j = 1,2,\ldots,20\right)$, $\left(\alpha\gamma\right)_{ij}$, and $\in_{ijk}$ are uncorrelated random variables having zero means and variances $V_{ar}\left(\gamma j\right) = \sigma_\gamma^2$, $V_{ar}\left(\left(\alpha\gamma\right)_{ij}\right) = \sigma_{\alpha\gamma}^2$ and $V_{ar}\left(\in_{ijk}\right) = \sigma^2$. Under the assumed unrestricted model, estimated variance components are $\hat{\sigma}_\gamma^2 = 10.2798$, $\hat{\sigma}_{\alpha\gamma}^2 = -0.1399$, and $\hat{\sigma}_\in^2 = 0.9917$. Applying the projection method, I to the data, the linear equations of variance components are given as follows:

$$SS_{part} = 1185.425 = 114\hat{\sigma}_\gamma^2$$
$$SS_{part\times operator} = 27.05 = 76\hat{\sigma}_{\alpha\gamma}^2$$
$$SS_{error} = 59.5 = 60\hat{\sigma}_\in^2 \quad (31)$$

The solutions of the equations are $\hat{\sigma}_\gamma^2 = 10.3985$, $\hat{\sigma}_{\alpha\gamma}^2 = 0.3559$ and $\hat{\sigma}_\in^2 = 0.9917$. All the variance components are estimated nonnegatively. When we apply projection method II to the same data, we get

$$RSS_{fixed} = 114\hat{\sigma}_{\gamma}^{2} + 76\hat{\sigma}_{\alpha\gamma}^{2} + 60\hat{\sigma}_{\in}^{2}$$
$$RSS_{part} = 76\hat{\sigma}_{\alpha\gamma}^{2} + 60\hat{\sigma}_{\in}^{2}$$
$$RSS_{part \times operator} = 60\hat{\sigma}_{\in}^{2} \quad (32)$$

Table 1: Data for a measurement systems capability study from Montgomery.

Part Number	1	2	3
1	21, 20	20, 20	19, 21
2	24, 23	24, 24	23, 24
3	20, 21	19, 21	20, 22
4	27, 27	28, 26	27, 28
5	19, 18	19, 18	18, 21
6	23, 21	24, 21	23, 22
7	22, 21	22, 24	22, 20
8	19, 17	18, 20	19, 28
9	24, 23	25, 23	24, 24
10	25, 23	26, 25	24, 25
11	21, 20	20, 20	21, 20
12	18, 19	17, 19	18, 19
13	23, 25	25, 25	25, 25
14	24, 24	23, 25	24, 25
15	29, 30	30, 28	31, 30
16	26, 26	25, 26	25, 27
17	20, 20	19, 20	20, 20
18	19, 21	19, 19	21, 23
19	25, 26	25, 24	25, 25
20	19, 19	18, 17	19, 17

where $RSS_{fixed} = 1271.975$, $RSS_{part} = 86.55$, and $RSS_{part \times operator} = 59.5$. The solutions for the equations are $\hat{\sigma}_{\gamma}^{2} = 10.3985$, $\hat{\sigma}_{\alpha\gamma}^{2} = 0.3559$ and $\hat{\sigma}_{\in}^{2} = 0.9917$ which are the same as the previous solutions. Hence, either one of the projection methods can be used for the nonnegative estimation of variance components of random effects in a mixed model. Projection method III also gives the same result as projection methods I and II for the data. As a second example, Searle [2]'s hypothetical data are illustrated. Searle explains why a negative estimate can occur in the estimation of variance component of random effects in a random model. The data are shown in Table 1. Since class in Table 2 is a random factor, the one-way random effects model is assumed. The assumed model is $y_{ij} = \mu + \alpha_i + \in_{ij}$, where the $\alpha_i\,(i = 1, 2)$ are random effects and $\in_{ij}$ are uncorrelated random errors having zero means and variances $V_{ar}(\alpha_i) = \sigma_{\alpha}^{2}$ and $V_{ar}(\in_{ij}) = \sigma^{2}$. As a result of the analysis of variance, the estimates of variance components are given as $\hat{\sigma}_{\alpha}^{2} = -15.33$ and $\hat{\sigma}_{\in}^{2} = 52$. Searle demonstrated how negative estimates could come from the analysis of variance and insisted that there would be nothing intrinsic in the method to prevent it. However, the projection methods yield the same nonnegative estimates as $\hat{\sigma}_{\alpha}^{2} = 2$ and $\hat{\sigma}_{\in}^{2} = 52$ in any method.

Table 2: Hypothetical data of a one-way classification from Searle and Gruber [2].

Class	Observations
1	19, 17, 15
2	25, 5, 15

Conclusion

Variance should be a nonnegative quantity as a measure of variation in data by its definition. In this work, it shows that orthogonal projections are very useful for defining a projection model for nonnegative variance estimation. Although there have been many attempts in literature to fix the problem of negative estimates for variance components over decades, those were not successful. However, the proposed methods in this paper always produce nonnegative estimates of variance components of the random effects in a mixed model. The two most important findings are checked and discussed for the estimation of nonnegative variance component. One is that a projection model should be derived from an assumed mixed-effects model. The other is that expectations of quadratic forms associated with the random effects should be evaluated from the projection model. This paper introduces terms such as projection method I, II, and III related to the methods, and the projection model for emphasizing projection rather than model fitting. Though they are based on the same assumed model, three methods are ap- plied differently in the application. Each method uses in its own way but summing up all orthogonal projections come to the observation vector. Depending on the types of projections, each method produces three different sets of equations for the evaluation of quadratic forms. Nonetheless, all of them show the same nonnegative estimates for variance components. It also shows that projection methods can be used for estimating variance components of the random effects in either random model or mixed model through examples. It should be noted that all the matrices associated with the quadratic forms come from the projection model not from the assumed model. In such a case, Hartley's synthesis can yield correct coefficients of variance components.

Funding

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant No.2018R1D1A1B07043021.

References

1. Searle SR (1971) Linear models. John Wiley and Sons New York, USA p. 1.

2. Searle SR, Gruber MHJ (2016) Linear Models. John Wiley and Sons New York, USA.
3. Nelder J (1954) A The interpretation of negative components of variance. Biometric 41: 544-548.
4. Milliken GA, Johnson DE (1984) Analysis of messy data volume 1: designed experiments. Van Nostrand Reinhold New York, USA.
5. Searle SR, Fawcett RF (1970) Expected mean squares in variance components models having finite populations. Biometrics 26(2): 243-254.
6. Hill BM (1965) Inference about variance components in the one-way mode. J Am Stat Assoc 60(311): 806-825.
7. Hill BM (1967) Correlated errors in the random model. J Am Stat Assoc 62(320): 1387-1400.
8. Searle SR, Casella G, McCulloch CE (2009) Variance components. John Wiley and Sons New York, USA.
9. Harville DA (1969) Variance component estimation for the unbalanced one-way random classification a critique. Aerospace Research Laboratories ARL p. 69.
10. Henderson CR (1953) Estimation of variance and covariance components. Biometrics 9(2): 226-252.
11. Hartley HO (1967) Expectations variances and covariances of ANOVA means squares by synthesis. Biometrics 23: 105-114.
12. GA (1983) Franklin Matrices with Applications in Statistics. Wadsworth Inc New York, USA.
13. Johnson RA, Wichern DW (2014) Applied multivariate statistical analysis. Prentice hall Upper Saddle River New Jersey, USA.
14. Montgomery DC (2013) Design and analysis of experiments. John Wiley and Sons New York, USA.
15. Thompson WA (1967) Negative estimates of variance components: an introduction Bulletin International Institute of Statistics 34: 1-4.
16. Thompson WA, Moore JR (1967) Non negative estimates of variance components. Tech no metrics 5(4): 441-449.

DOI: 10.32474/CTBB.2020.02.000140