Sankhyā : The Indian Journal of Statistics
1999, Volume 61, Series A, Pt. 3, pp. 431-443

ON THE EXISTENCE OF THE MAXIMUM LIKELIHOOD ESTIMATE IN VARIANCE COMPONENTS MODELS

By EUGENE DEMIDENKO
Dartmouth Medical School, New Hampshire
and
HÉLÈNE MASSAM
York University, Toronto and *University of Virginia, Charlottesville*

SUMMARY. We consider the variance components model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{r} \mathbf{Z}_i\mathbf{u}_i + \boldsymbol{\epsilon}$ where the random effects $\mathbf{u}_i$ are normally distributed as $\mathcal{N}(0, \sigma_i^2\mathbf{I}_{k_i}), i = 1, ...r$ and the common random error $\boldsymbol{\epsilon}$ is normally distributed as $\mathcal{N}(\mathbf{0},\sigma_0^2\mathbf{I}_n)$. The parameters $\{\boldsymbol{\beta},\sigma_0^2, ..., \sigma_r^2\}$ are constrained to satisfy the conditions $\sigma_0^2 > 0$, $\sigma_i^2 \geq 0, i = 1, ..., r$. The main results in this paper are Theorem 3.1 and Theorem 3.4 giving, respectively, necessary and sufficient conditions for the existence of the maximum likelihood estimate and the restricted maximum likelihood estimate of the parameters.

1. **Introduction**

The variance components model has a long history and has been considered by a number of authors. The book by Searle, Casella and McCulloch (1992) is one of the recent publications on the subject and covers most of what is currently known. Rao and Kleffe (1988) give an extensive coverage of variance components estimation and examine the problem of the existence of the estimates for a certain form of the variance component model. Criteria for the existence of the maximum likelihood estimate in general nonlinear regression models are given in Demidenko (1996).

One can also find in the literature several numerical algorithms for the maximum likelihood estimation of variance components. Two of the first papers on the subject are Hartley and Rao (1967) and Hemmerle and Hartley (1973). In Harville and Callanan (1990) and Callanan and Harville (1991) one can find a

Paper received. July 1998; revised February 1999.
AMS (1991) *subject classification.* 62H12.
Key words and phrases. Variance components, existence, maximum likelihood.

comparative study of the different algorithms. The reader is also referred to Searle *et al.* (1992), Chapter 8, for a survey of algorithms to compute the estimates of variance components. For the actual implementation of the calculations, a variety of procedures are available, in S-plus or SAS for example (see Littel *et al.* (1996) or Verbeke and Molenberghs (1997)).

In this paper, we are concerned with the existence of the maximum likelihood and restricted maximum likelihood estimates of the variance components in a variance components model. The model is written in matrix form as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \sum \mathbf{Z}_i\mathbf{u}_i + \boldsymbol{\epsilon}$ where $\mathbf{y}$ is an $n \times 1$ random vector, $\mathbf{X}$ is the $n \times m$ design matrix of fixed effects, $\boldsymbol{\beta}$ is an $m \times 1$ parameter vector, $\mathbf{Z}_i$ is the $n \times k_i$ design matrix of the random effects, $\mathbf{u}_i \sim \mathcal{N}(\mathbf{0}, \sigma_i^2\mathbf{I}_{k_i})$ are independent random effects, $i = 1, ..., r$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_0^2\mathbf{I}_n)$ is the random error associated with the common effect. The variance components $(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2)$ are such that $\sigma_0^2 > 0$, $\sigma_i^2 \geq 0$ and the parameter space $\Theta = \{\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2) \in R^m \times (0, \infty) \times [0, +\infty)^r\}$ is not compact. Therefore, the maximum of the likelihood function might not be achieved on the parameter space Θ. Our main purpose, here, is to give a necessary and sufficient condition for the existence of the maximum likelihood estimate (henceforth abbreviated as MLE) of $\boldsymbol{\theta}$ and this is done in Section 3. Indeed, Theorem 3.1 , states that the MLE exists if and only if the data vector $\mathbf{y}$ does not belong to the subspace of $\mathbf{R}^n$ spanned by the columns of $\mathbf{X}$ and $\mathbf{Z}_i, i = 1, ..., r$. If the random effects are removed from the model, i.e. if we consider the simple linear regression problem, this necessary and sufficient condition for the existence of the MLE says that the MLE exists if and only if the data vector does not belong to the mean space, which seems to be intuitively obvious. Thus, it is natural to wonder whether this necessary and sufficient condition is a common condition for the existence of other types of estimates of the variance components. We will see in Section 3.2 that this is not so, by considering the conditions of existence of one of the well known variance component estimate, the MINQUE.

In many practical instances, one prefers the maximum likelihood estimate of variance components based on a linear transformation $\mathbf{K}'\mathbf{y}$ of the data where $\mathbf{K}'$ is an $(n-m) \times n$ matrix of full rank such that $\mathbf{K}'\mathbf{X} = \mathbf{0}$. This estimate is called the restricted maximum likelihood estimate (henceforth abbreviated as RMLE). In Section 3.3, we give necessary and sufficient conditions for the existence of the RMLE. These results are parallel to those given for the MLE, of course. In Section 3.4 we compare our existence results for the MLE and RMLE to the existence results given in Rao and Kleffe (1988). As we shall see, our problem is somewhat different from the problem considered by Rao and Kleffe, in that we do not allow the variance components to be negative and σ_0^2 must be positive. We should also emphasize here that our necessary and sufficient conditions of existence of the MLE and RMLE are very easy to verify. The condition of existence of the MLE and RMLE derived in this paper should be integrated in all commercial software on variance components estimation. Indeed, if the existence conditions are not verified, the user may encounter unexpected computer program overflow and failure to estimate.

2. The Model

In this section the variance components model is introduced and the log-likelihood function is given with two different parametrizations.

2.1. *Specification of the model.* The variance components model is written in the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{r} \mathbf{Z}_i \mathbf{u}_i + \boldsymbol{\epsilon} \qquad \ldots (1)$$

where $\mathbf{y}$ is a $n \times 1$ random vector, $\mathbf{X}$ is a $n \times m$ matrix of full rank $m < n$, $\boldsymbol{\beta}$ is the $m \times 1$ parameter vector, $\mathbf{Z}_i$ is a fixed $n \times k_i$ design matrix, $\mathbf{u}_i \sim \mathcal{N}(\mathbf{0}, \sigma_i^2 \mathbf{I}_{k_i})$ are independent random effects and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_0^2 \mathbf{I}_n)$ is an $n \times 1$ vector of random common effect. We assume that all random terms are independent. Denoting $\boldsymbol{\sigma}^2 = (\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2)$ the parameter space is

$$\Theta = \{\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2) = (\boldsymbol{\beta}, \boldsymbol{\sigma}^2) \in R^m \times (0, +\infty) \times [0, \infty)^r\}.$$

Letting $\boldsymbol{\eta} = \sum \mathbf{Z}_i \mathbf{u}_i + \boldsymbol{\epsilon}$, (1) can be rewritten as a general linear regression model,

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\eta} \qquad \ldots (2)$$

where

$$\mathbf{V} = \mathbf{V}(\boldsymbol{\sigma}^2) = \sigma_0^2 \mathbf{I} + \sum_{i=1}^{r} \sigma_i^2 \mathbf{Z}_i \mathbf{Z}_i' \qquad \ldots (3)$$

is the covariance matrix of $\boldsymbol{\eta}$. We denote $\mathbf{Z} = [\mathbf{Z}_1, \ldots, \mathbf{Z}_r]$, the combined $n \times k$ matrix where $k = \sum k_i$.We assume that $k < n$.

2.2. *The log-likelihood function.* Since $\mathbf{y}$ is normally distributed, minus twice the log-likelihood function is given as

$$l(\boldsymbol{\theta}) = \log |\mathbf{V}(\boldsymbol{\sigma}^2)| + (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{V}^{-1}(\boldsymbol{\sigma}^2)(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \qquad \ldots (4)$$

To compute the MLE of the vector of variances $\boldsymbol{\sigma}^2$, we have to minimize this function over Θ. Since Θ is not compact, the MLE need not exist and in Section 3, we give a necessary and sufficient condition for the existence of the MLE.

It is convenient to introduce a slightly different parametrization. We define $\kappa_0 = \sigma_0^2$ and $\kappa_i = \sigma_i^2/\sigma_0^2$, $i = 1, \ldots, r$. Then, for $\boldsymbol{\kappa} = (\kappa_1, ..., \kappa_r)$,

$$\mathbf{V}(\boldsymbol{\sigma}^2) = \kappa_0(\mathbf{I} + \sum \kappa_i \mathbf{Z}_i \mathbf{Z}_i') = \kappa_0 \tilde{\mathbf{V}}(\boldsymbol{\kappa}) \qquad \ldots (5)$$

where $\tilde{\mathbf{V}}(\boldsymbol{\kappa}) = \mathbf{I} + \sum \kappa_i \mathbf{Z}_i \mathbf{Z}_i'$. Denote

$$\tilde{\Theta} = \{\tilde{\boldsymbol{\theta}} = (\boldsymbol{\beta}, \kappa_0, \kappa_1, \ldots, \kappa_r) \in R^m \times (0, +\infty) \times [0, +\infty)^r\}.$$

Then minus twice the log-likelihood function becomes

$$l(\tilde{\boldsymbol{\theta}}) = n \log \kappa_0 + \log |\tilde{\mathbf{V}}(\boldsymbol{\kappa})| + \kappa_0^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \tilde{\mathbf{V}}^{-1}(\boldsymbol{\kappa})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \qquad \ldots (6)$$

Allowing some ambiguity, we use the same notation $l(\boldsymbol{\theta})$ and $l(\tilde{\boldsymbol{\theta}})$ for minus twice the log likelihood function in terms of $\boldsymbol{\theta}$ or $\tilde{\boldsymbol{\theta}}$.

To compute the RMLE of $\boldsymbol{\sigma}^2$ we consider the log-likelihood function based on the linear transformation $\mathbf{y}_* = \mathbf{K}'\mathbf{y}$ of the original data $\mathbf{y}$, where $\mathbf{K}$ is an $n \times (n - m)$ matrix of full rank such that $\mathbf{K}'\mathbf{X} = \mathbf{0}$, e.g. Searle *et al.* (1992). The key-point of this transformation is that now minus twice the log-likelihood function is no longer a function of $\boldsymbol{\beta}$, but a function of $\boldsymbol{\sigma}^2$ only, namely,

$$l_K(\boldsymbol{\sigma}^2) = \log |\mathbf{K}'\mathbf{V}(\boldsymbol{\sigma}^2)\mathbf{K}| + \mathbf{y}'\mathbf{K} \left(\mathbf{K}'\mathbf{V}(\boldsymbol{\sigma}^2)\mathbf{K}\right)^{-1} \mathbf{K}'\mathbf{y}. \qquad \ldots (7)$$

Using the same parametrization $\kappa_i, i = 1, ..., r$ as above, we can express $l_K(\boldsymbol{\sigma}^2)$ as a function of κ_0 and the vector $\boldsymbol{\kappa} = (\kappa_1, ..., \kappa_r)$ in the following way

$$l_K(\kappa_0, \boldsymbol{\kappa}) = (n - m) \log \kappa_0 + \log |\mathbf{K}'\tilde{\mathbf{V}}(\boldsymbol{\kappa})\mathbf{K}| + \kappa_0^{-1} \mathbf{y}'\mathbf{K} \left(\mathbf{K}'\tilde{\mathbf{V}}(\boldsymbol{\kappa})\mathbf{K}\right)^{-1} \mathbf{K}'\mathbf{y}, \qquad \ldots (8)$$

since $\mathbf{K}'\tilde{\mathbf{V}}(\boldsymbol{\kappa})\mathbf{K}$ is an $(n - m) \times (n - m)$ nonsingular matrix.

Again, allowing some ambiguity, we use the same notation $l_K(\boldsymbol{\sigma}^2)$ and $l_K(\kappa_0, \boldsymbol{\kappa})$ for minus twice the log-likelihood function in terms of $\boldsymbol{\sigma}^2$ and $(\kappa_0, \boldsymbol{\kappa})$.

3. Existence of the MLE and RMLE

We are now going to derive necessary and sufficient conditions for the existence of the maximum likelihood and restricted maximum likelihood estimates. As mentioned in the introduction, we will, also, briefly consider the necessary and sufficient conditions of existence of the MINQUE of the variance components to see that the conditions of existence for the MLE are not equivalent to the conditions of existence for MINQUE. Finally, we will compare our conditions to the conditions given by Rao and Kleffe (1988).

3.1. *The maximum likelihood estimate.* The log-likelihood function $l(\boldsymbol{\theta})$ is a continuous function on the noncompact parameter set Θ. The following theorem gives a necessary and sufficient condition for the existence of the MLE of $\boldsymbol{\theta}$. Before stating the theorem, we introduce some notation. Let $M = \langle \mathbf{X} \rangle$ and $H = \langle \mathbf{Z} \rangle$ denote the subspace generated by the columns of $\mathbf{X}$ and the subspace generated by the columns of $\mathbf{Z} = [\mathbf{Z}_1, \ldots, \mathbf{Z}_r]$ respectively, so that $\langle \mathbf{X}, \mathbf{Z} \rangle =$

$M + H$. We denote by $\mathbf{P}_{H^\perp}$ the orthogonal projection from $\mathbf{R}^n$ to $H^\perp$, the subspace orthogonal to H, and by $\mathbf{P}_{(H+M)^\perp}$ the orthogonal projection from $\mathbf{R}^n$ to $(H+M)^\perp$. We will also need the quantity

$$s_{X,Z} = \mathbf{y}'\mathbf{P}_{(H+M)^\perp}\mathbf{y}.$$

THEOREM 3.1. (MLE EXISTENCE CRITERION). *The MLE in the variance components model exists if and only if*

$$\mathbf{y} \notin \langle \mathbf{X}, \mathbf{Z} \rangle. \qquad \ldots (9)$$

More precisely:

1. *If* $\mathbf{y} \notin \langle \mathbf{X}, \mathbf{Z} \rangle$*, then* $l(\boldsymbol{\theta}) \geq n \log s_{X,Z} - n \log n - n$*, and* $l(\boldsymbol{\theta})$ *reaches its minimum in* Θ.

2. *If* $\mathbf{y} \in \langle \mathbf{X}, \mathbf{Z} \rangle$*, the infimum of* $l(\boldsymbol{\theta})$ *in* Θ *is* $-\infty$.

We note that if (9) holds, then $s_{X,Z} > 0$ and $\log s_{X,Z}$ is well defined.

To prove Theorem 3.1 we need the following proposition, the proof of which is provided in the Appendix.

PROPOSITION 3.2. *Let* $\mathbf{Z}_i,\ i = 1, \ldots, r$ *be* $n \times k_i$ *matrices such that* $\sum k_i < n$. *Let* H *be the subspace generated by the columns of all the* $\mathbf{Z}_i$*'s , let* q *be the dimension of* H *and let* M *be the subspace generated by the columns of an* $n \times m$ *matrix* $\mathbf{X}$ *of rank* m. *Let* $\tilde{\mathbf{V}}(\boldsymbol{\kappa})$ *be defined as in (5). Then:*

(1) *There exists an* $n \times n$ *orthogonal matrix* $\mathbf{U}$, r *positive definite* $q \times q$ *matrices* $\mathbf{A}_1, ..., \mathbf{A}_r$ *such that*

$$\tilde{\mathbf{V}}(\boldsymbol{\kappa}) = \mathbf{U} \begin{bmatrix} \mathbf{I}_q + \sum_{i=1}^{r} \kappa_i \mathbf{A}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-q} \end{bmatrix} \mathbf{U}'. \qquad \ldots (10)$$

Moreover, when $\kappa_1 = ... = \kappa_r = t$ *, there exists an* $n \times n$ *orthogonal matrix* $\mathbf{U}$ *and a* $q \times q$ *diagonal matrix* $\mathbf{D}$ *with positive diagonal elements such that*

$$\tilde{\mathbf{V}}(t, ..., t) = \mathbf{U} \begin{bmatrix} \mathbf{I}_q + t\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-q} \end{bmatrix} \mathbf{U}'. \qquad \ldots (11)$$

(2) *The matrix* $\tilde{\mathbf{V}}^{-1}(\boldsymbol{\kappa}) - \mathbf{P}_{H^\perp}$ *is a positive definite symmetric matrix.*

(3) *For any given* $\mathbf{y}$ *in* $\mathbf{R}^n$ *and* $\mathbf{m}$ *in* M*, we have*

$$\begin{aligned} (\mathbf{y} - \mathbf{m})'\tilde{\mathbf{V}}^{-1}(\boldsymbol{\kappa})(\mathbf{y} - \mathbf{m}) &\geq (\mathbf{y} - \mathbf{m})'\mathbf{P}_{H^\perp}(\mathbf{y} - \mathbf{m}) \\ &\geq (\mathbf{y} - \mathbf{m})'\mathbf{P}_{(H+M)^\perp}(\mathbf{y} - \mathbf{m}), \end{aligned} \qquad \ldots (12)$$

which is equal to $\mathbf{y}'\mathbf{P}_{(H+M)^\perp}\mathbf{y}$ *since* $\mathbf{m} \in M$.

PROOF OF THEOREM 3.1. *Sufficient condition.* Let us assume $\mathbf{y} \notin \langle \mathbf{X}, \mathbf{Z} \rangle = H + M$. We consider expression (6) . By Proposition 3.2,

$$\begin{aligned} l(\widetilde{\boldsymbol{\theta}}) &= n \log(\kappa_0) + \log |\tilde{\mathbf{V}}(\boldsymbol{\kappa})| + \kappa_0^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\tilde{\mathbf{V}}^{-1}(\boldsymbol{\kappa})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &\geq n \log(\kappa_0) + \log |\tilde{\mathbf{V}}(\boldsymbol{\kappa})| + \kappa_0^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{P}_{(H+M)^\perp}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \end{aligned}$$

We note that $(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{P}_{(H+M)^\perp}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) = \mathbf{y}'\mathbf{P}_{(H+M)^\perp}\mathbf{y}$ since $\mathbf{X}\boldsymbol{\beta} \in M$. Moreover $\tilde{\mathbf{V}}(\boldsymbol{\kappa}) > \mathbf{I}$ and therefore $\log|\tilde{\mathbf{V}}(\boldsymbol{\kappa})| > \log 1 = 0$, and we have

$$l(\widetilde{\boldsymbol{\theta}}) \geq n\log\kappa_0 + \kappa_0^{-1}\mathbf{y}'\mathbf{P}_{(H+M)^\perp}\mathbf{y}.$$

Since $\mathbf{y} \notin H{+}M$, we have $\mathbf{y}'\mathbf{P}_{(H+M)^\perp}\mathbf{y} = s_{X,Z} > 0$ and therefore

$$l(\tilde{\boldsymbol{\theta}}) \geq n\log\kappa_0 + \kappa_0^{-1}s_{X,Z}. \qquad \ldots(13)$$

To find the minimum of the right hand side of (13) with respect to κ_0, we take its derivative. The minimum is reached at $\kappa_0 = s_{X,Z}/n$, and it follows that the function l is bounded below as follows

$$l(\tilde{\boldsymbol{\theta}}) \geq n\log s_{X,Z} - n\log n + n. \qquad \ldots(14)$$

To show that $l(\tilde{\boldsymbol{\theta}})$ achieves its minimum on $\tilde{\Theta}$, we need only show that it can become arbitrarily large outside a compact set of $\tilde{\Theta}$, that is, we have to show that given $A > n\log s_{X,Z} - n\log n + n$, there exists a compact set $C_A \subset \tilde{\Theta}$ such that $l(\tilde{\boldsymbol{\theta}}) \geq A, \forall\tilde{\boldsymbol{\theta}} \notin C_A$. For this purpose:

- choose ϵ in the interval $(0,1)$ such that

$$\text{for } \kappa_0 \in (0,\epsilon), \quad n\log\kappa_0 + \kappa_0^{-1}s_{X,Z} \geq A; \qquad \ldots(15)$$

$$\text{for } \kappa_0 > \epsilon^{-1}, \quad n\log\kappa_0 \geq A; \qquad \ldots(16)$$

$$\text{for } \epsilon \leq \kappa_0 \leq \epsilon^{-1}, \quad s_{X,Z} \geq (A - n\log\epsilon)\epsilon^{-1}; \qquad \ldots(17)$$

- choose any bounded set $B \subset M$;
- choose $b > 0$ such that

$$\text{for } \max_{1\leq i\leq r}\kappa_i \geq b, \quad \log|\tilde{\mathbf{V}}(\boldsymbol{\kappa})| \geq A - n\log\epsilon. \qquad \ldots(18)$$

Now suppose that $(\boldsymbol{\beta},\kappa_0,\kappa_1,\ldots,\kappa_r) \notin C_A = \{\boldsymbol{\beta} : \mathbf{X}\boldsymbol{\beta} \in B\} \times \left(\epsilon,\epsilon^{-1}\right) \times [0,b)^r$. Then we have several mutually exclusive possibilities:

1. $\kappa_0 \leq \epsilon$ or $\kappa_0 \geq \epsilon^{-1}$: if $\kappa_0 \leq \epsilon$, since $l(\tilde{\boldsymbol{\theta}}) \geq n\log\kappa_0 + \kappa_0^{-1}s_{X,Z}$, from (15) we have $l(\tilde{\boldsymbol{\theta}}) \geq A$; if $\kappa_0 \geq \epsilon^{-1}$, since $l(\tilde{\boldsymbol{\theta}}) \geq n\log\kappa_0$, from (16) we have $l(\tilde{\boldsymbol{\theta}}) \geq A$.

2. $\epsilon \leq \kappa_0 \leq \epsilon^{-1}$: by (13) and (17), $l(\tilde{\boldsymbol{\theta}}) \geq A$.

3. $\epsilon \leq \kappa_0 \leq \epsilon^{-1}$ and $\max_{1\leq i\leq r}\kappa_i \geq b$: by (18) $l(\tilde{\boldsymbol{\theta}}) \geq A$.

We have therefore proved that if $\tilde{\boldsymbol{\theta}} \notin C_A$ then $l \geq A$. Thus, $l(\tilde{\boldsymbol{\theta}})$ is a continuous function bounded from below and becomes arbitrarily large outside a compact set. Therefore, it attains a minimum on $\tilde{\Theta}$.

Necessary condition. Let us now assume that $\mathbf{y} \in \langle \mathbf{X}, \mathbf{Z} \rangle = M + H$. Then $\mathbf{y} = \mathbf{m}_0 + \mathbf{h}_0$ for some $\mathbf{m}_0 = \mathbf{X}\boldsymbol{\beta}_0$ in M and $\mathbf{h}_0$ in H and therefore

$$\inf_{\widetilde{\boldsymbol{\theta}}} l(\widetilde{\boldsymbol{\theta}}) \leq l(\boldsymbol{\beta}_0, \boldsymbol{\kappa}) = n \log \kappa_0 + \log |\tilde{V}(\boldsymbol{\kappa})| + \kappa_0^{-1} \mathbf{h}_0' \tilde{\mathbf{V}}^{-1}(\boldsymbol{\kappa}) \mathbf{h}_0. \qquad \ldots (19)$$

If $\mathbf{h}_0' \tilde{\mathbf{V}}^{-1}(\boldsymbol{\kappa}) \mathbf{h}_0 = 0$, i.e. if $\mathbf{h}_0 = \mathbf{0}$, for fixed $\kappa_1, \ldots, \kappa_r$, let us take the minimum of the right hand side of (19): as $\kappa_0 \to 0, n \log \kappa_0 \to -\infty$ and $\inf l(\widetilde{\boldsymbol{\theta}}) = -\infty$. If $\mathbf{h}_0 \neq \mathbf{0}$, taking $\kappa_0 = n^{-1} \mathbf{h}_0' \tilde{\mathbf{V}}^{-1}(\boldsymbol{\kappa}) \mathbf{h}_0$, (19) becomes

$$\inf_{\widetilde{\boldsymbol{\theta}}} l(\widetilde{\boldsymbol{\theta}}) \leq \log |\tilde{\mathbf{V}}(\boldsymbol{\kappa})| + n \log(\mathbf{h}_0' \mathbf{V}^{-1}(\boldsymbol{\kappa}) \mathbf{h}_0) - n \log n + n. \qquad \ldots (20)$$

Since $\mathbf{h}_0 \in H$, for $\mathbf{U}$ as defined in Proposition 3.2, $\mathbf{h}_0' \mathbf{U} = (h_{01}, \ldots, h_{0q}, 0, \ldots, 0)$. Let us be more specific and choose $\mathbf{U}$ such that in the basis $(\mathbf{U}_1, \ldots, \mathbf{U}_q)$, $\sum_{i=1}^r \mathbf{Z}_i \mathbf{Z}_i'$ is equal to a diagonal matrix $\mathbf{D}$. By Proposition 3.2, we know that the matrix $\mathbf{D}$ is positive definite with eigenvalues $\lambda_j > 0$, $j = 1, \ldots, q$. Let us take $\kappa_1 = \kappa_2 = \cdots = \kappa_r = t > 0$. Then, in the chosen basis $(\mathbf{U}_1, \ldots, \mathbf{U}_q, \mathbf{U}_{q+1}, \ldots, \mathbf{U}_n)$, (20) becomes

$$\inf_{\widetilde{\boldsymbol{\theta}}} l(\widetilde{\boldsymbol{\theta}}) \leq \sum_{j=1}^{q} \log(1 + t\lambda_j) + n \log \sum_{j=1}^{q} \frac{h_{0j}^2}{1 + t\lambda_j} - n \log n + n. \qquad \ldots (21)$$

Doing a series expansion of the logarithms, we see that, if $q < n$, the right hand side of (21) tends to $-\infty$ when $t \to +\infty$, and therefore $\inf_{\widetilde{\boldsymbol{\theta}}} l(\widetilde{\boldsymbol{\theta}}) = -\infty$. We note that q, the dimension of H, is less than k and we assumed that $k < n$. Therefore, the condition $q < n$ is satisfied and the theorem is proved. □

REMARKS. (1) If condition (9) takes place, then $\langle \mathbf{X}, \mathbf{Z} \rangle$ does not span the entire space R^n. Therefore (9) implies that the MLE exists with probability 1 if $\sigma_0^2 > 0$.

(2) We have proved that under condition (9), $l(\boldsymbol{\theta}) \to +\infty$ outside a compact set. Therefore the level sets $S_0 = \{\boldsymbol{\theta} \in \Theta : l(\boldsymbol{\theta}) \leq l(\boldsymbol{\theta}_0)\}$ for a given $\boldsymbol{\theta}_0$ are compact and the sequence of parameters generated by any minimization procedure will have at least one limit point.

(3) The existence criterion (9) can easily be verified by regressing $\mathbf{y}$ on $\langle \mathbf{X}, \mathbf{Z} \rangle$. Clearly $\mathbf{y} \in \langle \mathbf{X}, \mathbf{Z} \rangle$ if and only if the residuals of the regression are zero.

(4) When (9) is satisfied, the estimate of σ_0^2 has to be positive but the estimates of the individual variance components $\sigma_i^2, i = 1, \ldots, r$ might be zero. In the next theorem, we consider the special case $r = 1$ and give a sufficient condition for the estimate of σ_1^2 to be positive. We pay special attention to this case in the example given in Section 6.

(5) Condition (9) provides a condition for the existence of the MLE but does not address the problems of uniqueness and identifiability. Rao and Kleffe (1988)

give the following conditions for identifiability of $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$ respectively: the rank of $\mathbf{X}$ is equal to m and the matrices $\mathbf{I}, \mathbf{Z}_i\mathbf{Z}_i', \ldots, \mathbf{Z}_r\mathbf{Z}_r'$ are linearly independent, i.e. $\mathbf{V}(\boldsymbol{\sigma}_1^2) = \mathbf{V}(\boldsymbol{\sigma}_2^2)$ implies $\boldsymbol{\sigma}_1^2 = \boldsymbol{\sigma}_2^2$.

THEOREM 3.3. *Assume the existence condition (9) holds and $r = 1$. We denote $\hat{\boldsymbol{\beta}}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ and $\sigma_X^2 = \mathbf{y}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{y}/n = \|\mathbf{P}_{M^\perp}\mathbf{y}\|^2/n$. Then if*

$$\sigma_X^2 < \frac{\|\mathbf{Z}_1'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS})\|^2}{tr\mathbf{Z}_1'\mathbf{Z}_1} \qquad \ldots(22)$$

the MLE of σ_1^2 is positive.

PROOF. When $\sigma_1^2 = 0$, the variance components model (1) is the ordinary regression model with variance $\mathbf{V}(\boldsymbol{\sigma}^2) = \sigma_0^2\mathbf{I}$ and the MLE of $\boldsymbol{\beta}$ and σ_0^2 are respectively $\widehat{\boldsymbol{\beta}}_{OLS}$ and σ_X^2. If

$$\left.\frac{\partial l(\widehat{\boldsymbol{\beta}}_{OLS}, \sigma_X^2, \sigma_1^2)}{\partial \sigma_1^2}\right|_{\sigma_1^2=0} < 0,$$

then there exists a value of $\sigma_1^2 > 0$ which makes the value of $l(\boldsymbol{\theta})$, as defined in (4), smaller than $l(\hat{\boldsymbol{\beta}}_{OLS}, \sigma_X^2, 0)$. Let us differentiate the log-likelihood function (4) with respect to σ_1^2. Then we have

$$\frac{\partial l}{\partial \sigma_1^2} = tr(\mathbf{V}^{-1}\mathbf{Z}_1\mathbf{Z}_1') - (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{V}^{-1}\mathbf{Z}_1\mathbf{Z}_1'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

At $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}_{OLS}$ and $\sigma_1^2 = 0$, when $r = 1$, we have $\mathbf{V}(\boldsymbol{\sigma}^2) = \sigma_0^2\mathbf{I}$ and

$$\frac{\partial l}{\partial \sigma_1^2} = \frac{1}{\sigma_0^2}tr(\mathbf{Z}_1\mathbf{Z}_1') - \frac{1}{\sigma_0^4}\|\mathbf{Z}_1'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS})\|^2.$$

Therefore, $\partial l/\partial\sigma_1^2 < 0$ is equivalent to (22) and the MLE of σ_1^2 must be positive. □

We illustrate the last theorem with the following particular one-way unbalanced variance component model:

$$y_{ij} = \boldsymbol{\beta}'\mathbf{x}_i + u_i + \epsilon_{ij}, \quad i = 1, \ldots, p, \quad j = 1, \ldots, n_i \qquad \ldots(23)$$

where $\boldsymbol{\beta}$ is the $m \times 1$ vector of unknown parameters, $\mathbf{x}_i$ is the $m \times 1$ vector of explanatory variables at level i, $u_i \sim \mathcal{N}(0, \sigma_1^2)$ and $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_0^2)$ assuming the ϵ_{ij} mutually independent and independent of u_i, we denote $e_{ij} = y_{ij} - \hat{\boldsymbol{\beta}}_{OLS}'\mathbf{x}_i$ and condition (22) becomes

$$\sum_i\sum_j e_{ij}^2 < \sum_i\left(\sum_j e_{ij}\right)^2 \qquad \ldots(24)$$

Indeed $\mathbf{Z}_1$ is the $n \times k_1$ matrix whose ith column vectors is made up of zeros except for the n_i elements corresponding to that level, $\mathbf{Z}_1'\mathbf{Z}_1 = diag(n_1, \ldots, n_r)$ and $tr\mathbf{Z}_1\mathbf{Z}_1' = \sum_{i=1}^{p} n_i$. Then $\sigma_X^2 = \sum_i \sum_j e_{ij}^2/\Sigma n_i$, $\|\mathbf{Z}_1'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS})\|^2 = \sum_{i=1}^{r} \left(\sum_{j=1}^{n_i} e_{ij}\right)^2$ and clearly (22) is equivalent to (24).

3.2. *Existence conditions for MINQUE.* The necessary and sufficient condition of existence given in Theorem 3.1, for the MLE of the variance components states that the estimate exists when the data does not belong to the subspace spanned by the design matrix for the mean, $\mathbf{X}$, and by the design matrix for the variance components, $\mathbf{Z}$. When applied to the simple linear regression problems, this condition says that for the MLE of the variance to exist, the data must not belong to the mean space. This seems to be intuitively obvious and it is tempting to think that condition (9) would be a common condition for the existence of other variance components estimators. To show that this is not so, we consider the most popular variance components estimator, the MINQUE. The quadratic form $\mathbf{y}'\mathbf{A}\mathbf{y}$ is said to be a quadratic unbiased estimator (abbreviated MINQUE) of a given linear combination $\mathbf{f}'\boldsymbol{\sigma}$ of the variance components if it is unbiased, if $\mathbf{A}$ is a nonnegative symmetric matrix and if the norm of $\mathbf{A}$ is minimum. The existence conditions and the expressions of the estimates have been studied by many authors, see Pukelsheim (1977), Matthew (1984), Peddada (1984), Massam (1985), and Rao and Kleffe (1988), for example.

Let us denote $\mathbf{V}_i = \mathbf{Z}_i\mathbf{Z}_i'$. Finding the MINQUE reduces to solving a convex optimization problem over the cone of $n \times n$ nonnegative definite matrices. An estimate will exist if and only if the feasible set of this optimization problem is nonempty. The conditions of existence involve the linear combination vector $\mathbf{f}$ and the r nonnegative matrices $\mathbf{N}\mathbf{V}_i\mathbf{N}$ where $\mathbf{N}$ is the orthogonal projection operator on the subspace orthogonal to the subspace spanned by the column vectors of $\mathbf{X}$ and $\mathbf{Z}$. For the MINQUE, the existence conditions are independent of the data $\mathbf{y}$. Since the existence conditions are independent of $\mathbf{y}$ one might wonder if a MINQUE estimate equal to 0 would be equivalent to some condition on the data.

When the MINQUE estimate $\widehat{\sigma}_i^2$ exists, its expression is of the form $a_i^{-1}\mathbf{y}'\mathbf{A}_i\mathbf{y}$ where a_i is the rank of $\mathbf{A}_i$ and $\mathbf{A}_i$ is a nonnegative definite matrix equal to $((\mathbf{I} - \mathbf{G}_{(i)})\mathbf{V}_i(\mathbf{I} - \mathbf{G}_{(i)}))^+$ where $\mathbf{G}_{(i)}$ denotes the orthogonal projection of $\mathbf{R}^n$ onto the subspace generated by $\mathbf{X}, \mathbf{V}_1, ..., \mathbf{V}_{(i-1)}, \mathbf{V}_{(i+1)}, ..., \mathbf{V}_r$, and for any matrix $\mathbf{Q}$, $\mathbf{Q}^+$ denotes the generalized inverse of $\mathbf{Q}$ (see Theorem 5.6.6 of Rao (1988)). Let us see whether $\widehat{\sigma}_i^2 = 0$ corresponds to a condition on $\mathbf{y}$. Since $\mathbf{G}_{(i)}$ and $\mathbf{V}_i$ are symmetric matrices, this is equivalent to $\mathbf{y}'((\mathbf{I} - \mathbf{G}_{(i)})\mathbf{V}_i(\mathbf{I} - \mathbf{G}_{(i)}))\mathbf{y}$ $= 0$, and when the MINQUE estimate exists, this implies that the projection of $\mathbf{y}$ on the orthogonal complement of the subspace spanned by the columns of $\mathbf{X}, \mathbf{V}_1, ..., \mathbf{V}_{(i-1)}, \mathbf{V}_{(i+1)}, ..., \mathbf{V}_r$ is orthogonal to the subspace generated by the columns of $\mathbf{Z}_i$, which is the same, when $\mathbf{Z}_i$ is of full column rank, as the subspace spanned by the columns of $\mathbf{V}_i$. We cannot, however, conclude that $\mathbf{y}$

belongs to the subspace spanned by the columns of $\mathbf{X}, \mathbf{Z}_1, ..., \mathbf{Z}_r$, i.e., we cannot conclude that condition (9) does not hold. Therefore, the existence condition for the MINQUE is quite different from the condition given by (9).

3.3. *The restricted maximum likelihood estimate.* It is known that the MLE of variances σ_i^2 can be heavily biased especially for small n, because it does not take into account the degrees of freedom involved in estimating the fixed effects. Hence, it might be preferable to use an estimate of the variance components based on the linear transformation $\mathbf{K}'\mathbf{y}$ of the data where $\mathbf{K}'$ is an $(n-m) \times n$ matrix of rank $n-m$ such that $\mathbf{K}'\mathbf{X} = \mathbf{0}$, (see, for example, Searle *et al.* (1992)). Then $\mathbf{K}'\mathbf{y}$ follows a $\mathcal{N}(\mathbf{0}, \mathbf{K}'\mathbf{V}(\boldsymbol{\sigma}^2)\mathbf{K})$ distribution and does not carry any information on the fixed effects. Minus twice the log-likelihood function is $l_K(\boldsymbol{\sigma}^2)$ as given in Section 2.2. The parameter space Θ is now $\Theta = \{\boldsymbol{\sigma}^2 = (\sigma_0^2, \sigma_1^2, ..., \sigma_r^2), \sigma_0^2 > 0, \sigma_i^2 \geq 0\}$. The function $l_K(\boldsymbol{\sigma}^2)$ is a continuous function on a non-compact set Θ and therefore, as in the unrestricted case, its minimum needs not exist. The following theorem gives a necessary and sufficient condition for the existence of the RMLE of $\boldsymbol{\sigma}^2$ based on the given linear transformation $\mathbf{K}'\mathbf{y}$ of the original data. $\mathbf{K}'H$ denotes the image of the linear transformation $\mathbf{K}'$ of subspace H.

Theorem 3.4. (RMLE existence criterion). *Let* $\mathbf{y}$ *be a random variable defined as in variance component model (1). Let* $\mathbf{Z}$ *and* $H = \langle \mathbf{Z} \rangle$ *be as in Section 3.1. Let* $\mathbf{K}'$ *be an* $(n-m) \times n$ *matrix of rank* $n-m$ *such that* $\mathbf{K}'\mathbf{X} = \mathbf{0}$, *and* $\mathbf{N} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ *denote the projection on the subspace orthogonal to the range of* $\mathbf{X}$. *Then the RMLE of* $\boldsymbol{\sigma}^2$ *based on* $\mathbf{K}'\mathbf{y}$ *exists if and only if* $\mathbf{N}\mathbf{y} \notin H$.
More precisely,

1. *if* $\mathbf{N}\mathbf{y} \notin H$ *then*

$$\begin{aligned} l_K(\boldsymbol{\kappa}) &= (n-m)\log(\kappa_0) + \log \mid \mathbf{K}'\widetilde{\mathbf{V}}(\boldsymbol{\kappa})\mathbf{K} \mid + \kappa_0^{-1}\mathbf{y}'\mathbf{K}(\mathbf{K}'\widetilde{\mathbf{V}}(\boldsymbol{\kappa})\mathbf{K})^{-1}\mathbf{K}'\mathbf{y} \\ &\geq (n-m) \parallel \mathbf{P}_{(\mathbf{K}'H)^\perp}(\mathbf{N}\mathbf{y}) \parallel^2 - (n-m)\log(n-m) - (n-m) \end{aligned}$$

and $l_K(\cdot)$ *reaches its minimum in* Θ.

2. *if* $\mathbf{N}\mathbf{y} \in H$ *then the infimum of* $l_K(\boldsymbol{\kappa})$ *on* Θ *is* $-\infty$.

The proof of the theorem follows along the same lines as the proof of Theorem 3.1. The only difference is that now the data is in the image of the transformation $\mathbf{K}'$ rather than in all of $\mathbf{R}^n$ as in Section 3.1. The proof of Theorem 3.1 uses the following Proposition 3.5, which is the parallel of Proposition 3.2. The proofs of Theorem 3.4 and Proposition 3.5 are omitted.

Proposition 3.5. *Let* $\mathbf{Z}_i$, H, M *and* $\widetilde{\mathbf{V}}(\boldsymbol{\kappa})$ *be as in Proposition 3.2. Then* $(\mathbf{K}'\widetilde{\mathbf{V}}(\boldsymbol{\kappa})\mathbf{K})^{-1} - \mathbf{P}_{(\mathbf{K}'H)^\perp}$ *is a symmetric positive definite matrix where* $(\mathbf{K}'H)^\perp$ *denotes the orthogonal complement of* $\mathbf{K}'H$.

It is interesting to note that the existence condition for the MLE is stronger than the existence condition for the RMLE. Indeed, if $\mathbf{N}\mathbf{y} \in H = \langle \mathbf{Z} \rangle$, then there exists a vector γ such that

$$\mathbf{Ny} = \mathbf{Z}\gamma,$$

which implies that $\mathbf{y} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} + \mathbf{Z}\gamma = \mathbf{X}\tau + \mathbf{Z}\gamma$. So, $\mathbf{y} \in \langle \mathbf{X}, \mathbf{Z} \rangle$. We conclude from this that, if the RMLE does not exist, the MLE does not exist either. However, we are going to show that there are cases where the MLE might not exist while the RMLE exists. Consider the model with one fixed effect and one variance component ($r = 1, k_1 = 1$), $\mathbf{y} = \beta\mathbf{x} + \mathbf{z}u + \boldsymbol{\epsilon}$ where $\mathbf{x}$ and $\mathbf{z}$ are non-zero vectors in $\mathbf{R}^n$. Theorem 3.1 tells us that if $\mathbf{y}$ belongs to the subspace spanned by $\mathbf{x}$ and $\mathbf{z}$, then the MLE does not exist. If this is the case, the RMLE might still exist unless $\mathbf{Ny} \in \langle \mathbf{Z} \rangle$. Since $\mathbf{Ny}$ is the projection of $\mathbf{y}$ on the subspace orthogonal to $\mathbf{x}$, this latter condition can only be then satisfied in two cases: either $\mathbf{x}$ and $\mathbf{z}$ are orthogonal, or $\mathbf{x}$ and $\mathbf{z}$ are not orthogonal and $\mathbf{y}$ is collinear to $\mathbf{x}$. So, if $\mathbf{x}$ and $\mathbf{z}$ are not orthogonal, and $\mathbf{y}$ is not collinear to $\mathbf{x}$, then the RMLE exists while the MLE does not.

3.4. *Some remarks on previous work.* The question of existence of the MLE and RMLE is of course very important but has rarely been considered in a systematic way. The book by Rao and Kleffe (1988) is probably the only work where the issues of existence of the MLE and the RMLE are studied. However, the sufficient conditions for existence given there, are for a problem slightly different from the one considered in this paper. Indeed, in Chapter 9 of their book, Rao and Kleffe consider the linear regression model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta}$ where $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_{\boldsymbol{\theta}})$ with linear covariance structure, namely $\mathbf{V}_{\boldsymbol{\theta}} = \theta_1\mathbf{V}_1 + ... + \theta_r\mathbf{V}_r$ where $\mathbf{V}_1, ..., \mathbf{V}_r$ are linearly independent matrices and $\boldsymbol{\theta} = (\theta_1, ..., \theta_r) \in \mathcal{J} = \{\boldsymbol{\theta} : V_{\boldsymbol{\theta}}$ is positive definite$\}$. This model is different from ours, where $\mathbf{V}(\boldsymbol{\theta}) = \sigma_0^2\mathbf{I} + \sum \sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'$, in two respects. First, in Rao and Kleffe's formulation, the θ_i 's are allowed to be negative. Secondly , the matrices $\mathbf{V}_i$ in our model have the special structure $\mathbf{V}_i = \mathbf{Z}_i\mathbf{Z}_i'$. We argue that a negative value for variance component is not desirable from a statistical point of view. The MLE and RMLE that we consider in the present paper are such that $\widehat{\sigma}_0^2$ is positive and $\widehat{\sigma}_i^2$ are nonnegative. For the problem they consider, Rao and Kleffe give necessary and sufficient conditions in the special cases listed below:

1. If $\mathbf{y} \in < \mathbf{X} > = M$, they prove that when the parameters are $(\boldsymbol{\beta}, \alpha\boldsymbol{\theta}_0)$ for arbitrary $\boldsymbol{\beta}$ and arbitrary $\boldsymbol{\theta}_0 \in \mathcal{J}$, then $l(\boldsymbol{\theta}) \to -\infty$ as $\alpha \to 0$.

2. For $n = 2$ and $\mathbf{y} = (y_1, y_2)', \mathbf{Z}_1 = (1, 0)', \mathbf{Z}_2 = (0, 1)'$ and $\sigma_0^2 = 0$, they prove that $l(\boldsymbol{\theta})$ is unbounded.

3. On page 231, they assume that $\mathbf{V}_{\boldsymbol{\theta}} = \boldsymbol{\theta}\mathbf{I} + \mathbf{U}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{U}'$ for some known matrix $\mathbf{U}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ depending on an unknown parameter $\boldsymbol{\theta}$. This form of $\mathbf{V}_{\boldsymbol{\theta}}$ is similar to our $\mathbf{V}(\boldsymbol{\theta})$. In that case, they say that a sufficient condition for the existence of the MLE is that there exists a $\boldsymbol{\beta}$ such that $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \mathbf{U}\mathbf{U}^+)(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) > \mathbf{0}$. For our problem, this condition becomes $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \mathbf{Z}\mathbf{Z}^+)(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) > \mathbf{0}$ for some $\boldsymbol{\beta}$. It is not difficult to show that this is equivalent to our condition $\mathbf{y} \notin \langle \mathbf{X}, \mathbf{Z} \rangle$.

Clearly, the results in cases 1 and 2 are special cases of our result: "If $\mathbf{y} \notin \langle \mathbf{X}, \mathbf{Z} \rangle$ then $l(\boldsymbol{\theta}) \to -\infty$ in some direction". And the sufficient condition in case 3 is the same as condition (9). Finally, on pages 233 and 236, Rao and Kleffe give conditions under which the MLE and RMLE do not admit a solution in the admissible region $\mathcal{J}$, but those conditions are difficult to verify. On the contrary, the necessary and sufficient condition for the existence of the MLE and the existence of the RMLE, given in this paper, are easy to verify.

4. Appendix

4.1. *Proof of proposition 3.2.* Let $(\mathbf{U}_1, \ldots, \mathbf{U}_n)$ be an orthonormal basis of $\mathbf{R}^n$ such that $(\mathbf{U}_1, \ldots, \mathbf{U}_q)$ is a basis of H and $(\mathbf{U}_{q+1}, \ldots, \mathbf{U}_n)$ is a basis of $H^\perp$. Let $\mathbf{U}$ be the orthogonal matrix with column vectors $\mathbf{U}_1, \ldots, \mathbf{U}_n$. Then, since the range of $\mathbf{Z}_i\mathbf{Z}_i'$ is included in H, by the classical change of basis formula, for each $i = 1, ..., r$, there exists a positive $q \times q$ matrix $\mathbf{A}_i$ and an $(n-q) \times q$ matrix $\mathbf{B}_i$ such that

$$\mathbf{Z}_i\mathbf{Z}_i' = \mathbf{U} \begin{bmatrix} \mathbf{A}_i & \mathbf{0} \\ \mathbf{B}_i & \mathbf{0} \end{bmatrix} \mathbf{U}'$$

with obvious notation for the 0-block matrices. Moreover, the matrix $\mathbf{Z}_i\mathbf{Z}_i'$ is symmetric and therefore

$$\mathbf{Z}_i\mathbf{Z}_i' = \mathbf{U} \begin{bmatrix} \mathbf{A}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}'$$

and (10) follows. Since in the basis $(\mathbf{U}_1, \ldots, \mathbf{U}_n)$ we can write

$$\mathbf{P}_{H^\perp} = \mathbf{U} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-q} \end{bmatrix} \mathbf{U}', \quad \tilde{\mathbf{V}}(\boldsymbol{\kappa}) - \mathbf{P}_{H^\perp} = \mathbf{U} \begin{bmatrix} \mathbf{I}_q + \sum_{i=1}^r \kappa_i \mathbf{A}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}'$$

where $\mathbf{A}_i$ are positive definite and $\kappa_i \geq 0$. This proves that matrix $\tilde{\mathbf{V}}(\boldsymbol{\kappa}) - \mathbf{P}_{H^\perp}$ is a nonnegative symmetric matrix and it implies immediately the first inequality in (12). Since $\mathbf{P}_{H^\perp} - \mathbf{P}_{(H+M)^\perp}$ is positive definite, the second inequality also follows immediately. When $\kappa_1 = ... = \kappa_r = t$, since $\sum \mathbf{A}_i$ is positive definite, we can find a $\mathbf{U}$ such that in the basis $(\mathbf{U}_1, ..., \mathbf{U}_q)$, $\mathbf{D}$ is diagonal with positive diagonal elements and this proves (11).

References

Callanan T.P., Harville D.A. (1991). Some new algorithms for computing restricted maximum likelihood estimates of variance components, *Journal of Statistical Computation*, **38**, 239-259.

Demidenko, E. (1996). On the existence of the least squares estimate in nonlinear growth curve models of exponential type, *Communications in Statistics, Theory and Methods*, **25**, 159-182.

HARVILLE D.A., CALLANAN T.P. (1990). Computational aspects of likelihood-based inference for variance components. In: Gianola D. and Hammonds K. (eds.) *Advances in Statistical Methods for Genetic Improvement of Livestock*, 136-176, Springer-Verlag, Heidelberg.

HARTLEY H.O., RAO J.N.K. (1967). Maximum-likelihood estimation for the mixed analysis of variance model, *Biometrika*, **54**, 93-108.

HEMMERLE W.J., HARTLEY H.O. (1973). Computing maximum likelihood estimates for the mixed A.O.V. model using the $W-$transformation, *Technometrics*, **15**, 819-831.

LITTEL, R.C., MILLIKEN, G.A., STROUP, W.W. AND WOLFINGER, R.D. (1996), *SAS System for Mixed Models*, SAS Institute, Inc.,Cary, North Carolina.

MASSAM, H., (1985), The nonnegative MINQUE estimate, *J. of Multivariate Anal.*, **16**, 253-259.

MATTHEW, T. (1984). On nonnegative quadratic unbiased estimability of variance components, *Ann. Statist.,* **12,** 1566-1569.

PEDDADA, S.D., (1984). Some results on nonnegative estimation of variance components, *Tech. Report.* No. 84-12, Center for Multivariate Analysis, University of Pittsburgh.

PUKELSHEIM, F. (1977). On Hsu's model in regression analysis, *Math. Operationsforsch. u. Statist. Ser. Statistics* ,**8**, 323-331.

RAO, C.R., KLEFFE, J. (1988). *Estimation of Variance Components and Applications*, North Holland. Amsterdam.

SEARLE S.R., CASELLA G., MCCULLOCH C.M. (1992). *Variance Components*, New York, J. Wiley.

VERBEKE, G. AND MOLENBERGHS, G. (1997), *Linear Mixed Models in Practice: A SAS Oriented Approach*, Springer Verlag, New York.

EUGENE DEMIDENKO
DARTMOUTH MEDICAL SCHOOL
NEW HAMPSHIRE
USA
e-mail: eugene.demidenko@dartmouth.edu

HÉLÈNE MASSAM
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE
USA
e-mail : hmm2n@virginia.edu