Preliminary Report

On the Pseudo-Expectation Methods Using Animal Models for Variance Component Estimation

Kohichi Yamagishi*¹, Hiroaki Iwaisaki*² and Ichiro Ashida*²

*¹Course of Environmental Management Science,
Graduate School of Science and Technology, Niigata University,
Niigata 950-2181, Japan

*²Department of Animal Science, Faculty of Agriculture,
Niigata University,
Niigata 950-2181, Japan

e-mail: iwska@agr.niigata-u.ac.jp

The pseudo-expectation (PE) methods for estimating variance components in the mixed linear models are equivalent to the restricted maximum likelihood (REML) procedure in the balanced case, but are a variety of the approximations in the unbalanced case. In animal breeding applications, the PE methods are used, for instance, to quickly obtain an initial value in the REML estimation. In this article, we reveal that the PE approach can never be defined under the so-called individual animal model, when each animal has only one record. We then exploit the use of another sort of animal models, or the so-called reduced animal model, in the PE approach and present sets of the possible quadratic forms. It is shown that the current PE procedure actually performs numerically, as expected, but the sampling variance of the estimator is quite likely to be relatively large, and the similarity of the current estimate to the REML estimate would considerably be affected by the given data-structure, definitely indicating the necessity of further theoretical investigation for the improved procedure.

Key words: Key words: Animal model; Mixed linear model; Pseudo-expectation method; REML; Unbalanced data; Variance component.

1. Introduction

In animal breeding applications, data used to estimate variance components and thereby genetic parameters are usually unbalanced, and operational statistical models employed are often a mixed-effects model to take account of fixed macro-environmental effects and random genetic and residual effects affecting economically important quantitative traits. Of the numerous methods for variance component estimation developed until now (for a review, e.g., Khuri and Sahai, 1985; Searle et al., 1992), approximately before the 1970s, Method III of Henderson (1953) had been the method of choice. From the early 1970s, the restricted maximum likelihood (REML;

Patterson and Thompson, 1971) procedure has widely been used in many livestock species. In the recent decade, Method R of Reverter et al. (1994) and Bayesian methods using Gibbs sampling (e.g., Wang et al., 1993; Van Tassell et al., 1995) are also of interest to animal breeders.

The pseudo-expectation (PE) methods (Thompson, 1969; Schaeffer, 1986; Van Raden and Jung, 1988; Iwaisaki, 1992; Shi et al., 1993) are sometimes used and classified into one sort of the REML-like methods (Harville, 1977). These are actually approximations to REML that mimic the computational steps in REML using the expectation-maximization algorithm (EM-REML) in different degrees, and consequently their computational requirements are relatively small. (e.g., Ouweltjes et al., 1988; Iwaisaki, 1992; Yamagishi and Iwaisaki, 1998). In the applications, it has been considered that the PE methods are useful for estimating variance components in the populations with no or weak selection, or for quickly obtaining an initial value for the REML estimation. Thus, a traditional sire model or a sire-maternal grandsire model was used in the PE estimation (e.g., Bertrand and Kriese, 1990; Shi et al., 1993).

We recently applied the PE method of Thompson (1969) and Schaeffer (1986), using the individual animal model (IAM), to several data sets on beef carcass traits in order to get pre-assigned values for the EM-REML estimation. However, our attempts were all unsuccessful. Following these results, the purposes of this preliminary report are to theoretically examine the property of the PE method using IAM, and to investigate the possibility of the PE approach using the animal model of another type.

2. Property of the PE method using IAM

In this section, assuming IAM, we first derive a set of PE quadratic forms from the viewpoint of Thompson's (1969) and Schaeffer's (1986) methodology and then discuss their validity.

Consider the following mixed linear model, or IAM:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e}, \quad (1)$$

where $\mathbf{y}$ is an $N \times 1$ vector of observations, $\mathbf{b}$ is a $r \times 1$ vector of fixed effects, $\mathbf{a}$ is a $q \times 1$ vector of animals' breeding values, or additive genotypic values, $\mathbf{e}$ is an $N \times 1$ vector of residuals, and $\mathbf{X}$ and $\mathbf{Z}$ are $N \times r$ and $N \times q$ known incidence matrices, respectively. We assume that the expectation and the dispersion parameter for the random vectors are

$$E \begin{pmatrix} \mathbf{y} \\ \mathbf{a} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{X}\mathbf{b} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \quad \text{and} \quad \text{Var} \begin{pmatrix} \mathbf{a} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{A}\sigma_a^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_e^2 \end{pmatrix},$$

respectively, where σ_a^2 and σ_e^2 are the components of additive genetic variance and residual variance, respectively, and $\mathbf{A}$ is the additive relationship matrix for animals. Under the model (1), with the prior information $\tilde{\alpha} = \tilde{\sigma}_e^2 / \tilde{\sigma}_a^2$ for $\alpha = \sigma_e^2 / \sigma_a^2$, the mixed model equations are written

as

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\tilde{\alpha} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}. \quad (2)$$

The solutions to these equations are then expressed, with $\mathbf{Z} = \mathbf{I}$, as

$$\hat{\mathbf{a}} = (\mathbf{M} + \mathbf{A}^{-1}\tilde{\alpha})^{-1}\mathbf{M}\mathbf{y}, \quad (3)$$

where $\mathbf{M}$ is a projection matrix of the form:

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

We now consider the pseudo-expectations (denoted as $\mathcal{E}$) of Thompson's (1969) and Schaffer's (1986) quadratic forms with $\mathbf{Z} = \mathbf{I}$:

$$\hat{\mathbf{a}}'\mathbf{A}^{-1}\mathbf{M}\mathbf{y} \quad \text{and} \quad \mathbf{y}'\mathbf{M}\mathbf{y} - \hat{\mathbf{a}}'\mathbf{M}\mathbf{y}.$$

The 'pseudo-expectation' is the term given to deriving expectations of quadratic forms as if the priori values are equal to the true values. Then, we have

$$\mathcal{E}(\hat{\mathbf{a}}'\mathbf{A}^{-1}\mathbf{M}\mathbf{y}) = \{N - r(\mathbf{X})\}\sigma_a^2 \quad (4)$$

and

$$\mathcal{E}(\mathbf{y}'\mathbf{M}\mathbf{y} - \hat{\mathbf{a}}'\mathbf{M}\mathbf{y}) = \{N - r(\mathbf{X})\}\sigma_e^2 \quad (5)$$

with $\text{tr}(\mathbf{M}) = N - r(\mathbf{X})$, where tr is the trace operator and $r(\mathbf{X})$ is the rank of $\mathbf{X}$. The quadratic form $\hat{\mathbf{a}}'\mathbf{A}^{-1}\mathbf{M}\mathbf{y}$ can be written as

$$\mathbf{y}'\mathbf{M}(\mathbf{M} + \mathbf{A}^{-1}\tilde{\alpha})^{-1}\mathbf{A}^{-1}\tilde{\alpha}\mathbf{M}\mathbf{y}\tilde{\alpha}^{-1},$$

which is further expressed as

$$\mathbf{y}'\mathbf{M}\{\mathbf{I} - (\mathbf{M} + \mathbf{A}^{-1}\tilde{\alpha})^{-1}\mathbf{M}\}\mathbf{M}\mathbf{y}\tilde{\alpha}^{-1}.$$

Consequently, the estimates of the two variance components in one round of iteration are given as

$$\hat{\sigma}_a^2 = \{\mathbf{y}'\mathbf{M}\mathbf{y} - \mathbf{y}'\mathbf{M}(\mathbf{M} + \mathbf{A}^{-1}\tilde{\alpha})^{-1}\mathbf{M}\mathbf{y}\} / \{N - r(\mathbf{X})\}\tilde{\alpha} \quad (6)$$

and

$$\hat{\sigma}_e^2 = \{\mathbf{y}'\mathbf{M}\mathbf{y} - \mathbf{y}'\mathbf{M}(\mathbf{M} + \mathbf{A}^{-1}\tilde{\alpha})^{-1}\mathbf{M}\mathbf{y}\} / \{N - r(\mathbf{X})\}. \quad (7)$$

This finding straightforwardly leads to

$$\hat{\sigma}_e^2 / \hat{\sigma}_a^2 = \tilde{\alpha} \quad (8)$$

Hence, it is seen that the PE estimation using IAM does not perform, if each animal does not have repeated records.

3. A PE approach using another sort of animal models

Now, we notice the characteristics of the data-structure of beef carcass records. That is, carcass data fulfill the conditions: the number of records per animal is only one, none of the animals having a record becomes a parent, and the animals producing progeny have no own record. Then, with these conditions satisfied, the breeding value of animal i having one record under IAM can be written as

$$a_i = 0.5(a_s + a_d) + m_i, \quad (9)$$

where a_i is the breeding value of animal i , a_s and a_d are the breeding values of the parents (namely, the sire and the dam), and m_i is the Mendelian sampling effect. In matrix notation, we have

$$\mathbf{a}_o = \mathbf{F}\mathbf{a}_p + \mathbf{m} \quad (10)$$

with

$$\begin{aligned} \mathbf{a} &= \begin{pmatrix} \mathbf{a}_p \\ \mathbf{a}_o \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{I} \\ \mathbf{F} \end{pmatrix} \mathbf{a}_p + \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \end{pmatrix} \mathbf{m}, \end{aligned}$$

where $\mathbf{a}_p$ is the vector of breeding values of animals having one or more progeny, $\mathbf{a}_o$ is the vector of breeding values of animals having no progeny, and $\mathbf{m}$ is the vector of the Mendelian sampling term. Then, in this case, the model (1) can equivalently be rewritten as another type of animal models, or the reduced animal model (RAM):

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_p\mathbf{a}_p + \boldsymbol{\varepsilon} \quad (11)$$

with

$$E \begin{pmatrix} \mathbf{y} \\ \mathbf{a}_p \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} \mathbf{X}\mathbf{b} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \text{Var} \begin{pmatrix} \mathbf{a}_p \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_p\sigma_a^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_p \end{pmatrix},$$

and

$$\text{Var}(\mathbf{y}) = \mathbf{Z}_p\mathbf{A}_p\mathbf{Z}_p'\sigma_a^2 + \mathbf{R}_p; \quad \mathbf{R}_p = \mathbf{I}\sigma_e^2 + \mathbf{D}\sigma_a^2,$$

where $\mathbf{a}_p$ is a $q_p \times 1$ vector of breeding values of animals having progeny, $\boldsymbol{\varepsilon}$ is an $N \times 1$ vector of residual effects, and $\mathbf{Z}_p$ is an $N \times q_p$ known incidence matrix, $\mathbf{A}_p$ is the additive relationship matrix for animals having progeny, and $\mathbf{D}$ is a diagonal matrix whose diagonal elements d_i are given by

$$\begin{aligned} d_i &= 1 - 0.25(a_{ss} + a_{dd}) \\ &= 0.5 - 0.25(f_s + f_d), \end{aligned} \quad (12)$$

where f_s and f_d are inbreeding coefficients for the sire and the dam, respectively, and $a_{ss} = 1 + f_s$ and $a_{dd} = 1 + f_d$.

The empirical mixed model equations in this case are given as

$$\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}_p^{-1}\mathbf{X} & \mathbf{X}'\tilde{\mathbf{R}}_p^{-1}\mathbf{Z}_p \\ \mathbf{Z}_p'\tilde{\mathbf{R}}_p^{-1}\mathbf{X} & \mathbf{Z}_p'\tilde{\mathbf{R}}_p^{-1}\mathbf{Z}_p + \mathbf{A}_p^{-1}\tilde{\sigma}_a^{-2} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}}_p \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}_p^{-1}\mathbf{y} \\ \mathbf{Z}_p'\tilde{\mathbf{R}}_p^{-1}\mathbf{y} \end{pmatrix}, \quad (13)$$

where $\tilde{\sigma}_a^{-2}$ and $\tilde{\mathbf{R}}_p^{-1}$ are the prior information for σ_a^{-2} and $\mathbf{R}_p^{-1}$, respectively, and these equations can be rewritten as

$$\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{S}}^{-1}\mathbf{X} & \mathbf{X}'\tilde{\mathbf{S}}^{-1}\mathbf{Z}_p \\ \mathbf{Z}_p'\tilde{\mathbf{S}}^{-1}\mathbf{X} & \mathbf{Z}_p'\tilde{\mathbf{S}}^{-1}\mathbf{Z}_p + \mathbf{A}_p^{-1}\tilde{\alpha} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}}_p \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\tilde{\mathbf{S}}^{-1}\mathbf{y} \\ \mathbf{Z}_p'\tilde{\mathbf{S}}^{-1}\mathbf{y} \end{pmatrix} \quad (14)$$

with $\tilde{\mathbf{S}} = \mathbf{I} + \mathbf{D}\tilde{\alpha}^{-1}$.

Following Henderson's (1986) work on REML under RAM, with the above equations (14), considering the quadratic forms of Thompson's (1969) and Schaeffer's (1986) type leads to

$$\hat{\mathbf{a}}_p' \mathbf{Q} \mathbf{Z}_p' \mathbf{T} \mathbf{y}, \quad \hat{\boldsymbol{\varepsilon}}' \mathbf{P}_a \mathbf{y} \quad \text{and} \quad \hat{\boldsymbol{\varepsilon}}' \mathbf{P}_e \mathbf{y}$$

with

$$\mathbf{T} = \tilde{\mathbf{S}}^{-1} - \tilde{\mathbf{S}}^{-1} \mathbf{X} (\mathbf{X}' \tilde{\mathbf{S}}^{-1} \mathbf{X})^{-1} \mathbf{X}' \tilde{\mathbf{S}}^{-1}; \quad \mathbf{Q} = \mathbf{A}_p^{-1} \tilde{\sigma}_a^{-4}; \quad \mathbf{P}_a = \tilde{\mathbf{S}}^{-1} \mathbf{D} \tilde{\mathbf{S}}^{-1} \tilde{\sigma}_e^{-4}; \quad \mathbf{P}_e = \tilde{\mathbf{S}}^{-1} \mathbf{S}^{-1} \tilde{\sigma}_e^{-4}.$$

Then the pseudo-expectations of these quadratic forms are expressed as

$$\mathcal{E}(\hat{\mathbf{a}}_p' \mathbf{Q} \mathbf{Z}_p' \mathbf{T} \mathbf{y}) = \text{tr}(\mathbf{Z}_p' \mathbf{T} \mathbf{Z}_p) \sigma_a^{-2}, \quad (15)$$

$$\mathcal{E}(\hat{\boldsymbol{\varepsilon}}' \mathbf{P}_a \mathbf{y}) = [\text{tr}(\mathbf{T} \mathbf{D}) - \text{tr}\{\mathbf{Z}_p \mathbf{C} \mathbf{Z}_p' \mathbf{T} \mathbf{D} (\mathbf{T} - \tilde{\mathbf{S}}^{-1})\}] \sigma_e^{-2} \quad (16)$$

and

$$\mathcal{E}(\hat{\boldsymbol{\varepsilon}}' \mathbf{P}_e \mathbf{y}) = [\text{tr}(\mathbf{T}) - \text{tr}\{\mathbf{Z}_p \mathbf{C} \mathbf{Z}_p' \mathbf{T} (\mathbf{T} - \tilde{\mathbf{S}}^{-1})\}] \sigma_e^{-2}, \quad (17)$$

where

$$\mathbf{C} = (\mathbf{Z}_p' \mathbf{T} \mathbf{Z}_p + \mathbf{A}_p^{-1} \tilde{\alpha})^{-1}.$$

Thus, the PE estimation equations are written as

$$\begin{pmatrix} \hat{\sigma}_a^2 \\ \hat{\sigma}_e^2 \end{pmatrix} = \begin{pmatrix} \text{tr}(\mathbf{Z}_p' \mathbf{T} \mathbf{Z}_p) & \text{tr}(\mathbf{T} \mathbf{D}) - \text{tr}\{\mathbf{Z}_p \mathbf{C} \mathbf{Z}_p' \mathbf{T} \mathbf{D} (\mathbf{T} - \tilde{\mathbf{S}}^{-1})\} \\ \mathbf{0} & \text{tr}(\mathbf{T}) - \text{tr}\{\mathbf{Z}_p \mathbf{C} \mathbf{Z}_p' \mathbf{T} (\mathbf{T} - \tilde{\mathbf{S}}^{-1})\} \end{pmatrix}^{-1} \begin{pmatrix} \hat{\mathbf{a}}_p' \mathbf{A}_p^{-1} \mathbf{Z}_p' \mathbf{T} \mathbf{y} + \hat{\boldsymbol{\varepsilon}}' \tilde{\mathbf{S}}^{-1} \mathbf{D} \tilde{\mathbf{S}}^{-1} \mathbf{y} \\ \hat{\boldsymbol{\varepsilon}}' \tilde{\mathbf{S}}^{-1} \tilde{\mathbf{S}}^{-1} \mathbf{y} \end{pmatrix}. \quad (18)$$

Further, an approximation to equation (18) is given as

$$\begin{pmatrix} \hat{\sigma}_a^2 \\ \hat{\sigma}_e^2 \end{pmatrix} = \begin{pmatrix} \text{tr}(\mathbf{Z}_p' \mathbf{T} \mathbf{Z}_p) & \text{tr}(\mathbf{T} \mathbf{D}) \\ \mathbf{0} & \text{tr}(\mathbf{T}) \end{pmatrix}^{-1} \begin{pmatrix} \hat{\mathbf{a}}_p' \mathbf{A}_p^{-1} \mathbf{Z}_p' \mathbf{T} \mathbf{y} + \hat{\boldsymbol{\varepsilon}}' \tilde{\mathbf{S}}^{-1} \mathbf{D} \tilde{\mathbf{S}}^{-1} \mathbf{y} \\ \hat{\boldsymbol{\varepsilon}}' \tilde{\mathbf{S}}^{-1} \tilde{\mathbf{S}}^{-1} \mathbf{y} \end{pmatrix}, \quad (19)$$

assuming that $\mathbf{T}$ is equal to $\tilde{\mathbf{S}}^{-1}$.

4. Preliminary numerical analysis

The conditions mentioned in Section 3 are fulfilled in the data of beef carcass traits, since slaughtered fattening animals have one record each and never become a parent that produces the progeny. Here, considering a typical data-structure on Wagyu cattle, a total of thirty animals were simulated, of which fourteen animals including full- and half-sibs were fattening animals and slaughtered in either of two consecutive years. Parents of these slaughtered animals were composed of five sires and seven dams with no records, and the remaining four animals were ancestors in the pedigree that have no records.

Assuming that both additive genetic value and environmental deviation follow normality, records of 'marbling score' were generated. The true values used for overall mean, heritability, and additive genetic variance were 2.5, 0.3, and 0.4, respectively. The proportion of the variation due to the fixed effects, or slaughter-year effects, to the total variation was supposed to be 30 %. Ten replicates were conducted, and each sample was analyzed by the procedures using equations (18) and (19) and the EM-REML procedure, using the initial value of 2.3333 (the true value) for the variance ratio and the additive relationship matrix constructed according to Quaas (1976). Convergence was assumed when the change of the variance ratio between two successive rounds was less than 10^{-6} %.

As expected, the current procedures actually performed numerically, and for all the samples the completely converged estimates of animal and residual variances were obtained. With the same numbers of iterations, the difference between the corresponding estimates based on the use of equations (18) and (19) was quite negligible for each sample. Table 1 shows the numbers of iterations required to attain convergence with the current procedure using equation (18) (or (19)) and the REML procedure and the degrees of similarity of the current estimates to the

Table 1. The numbers of iterations required in the current and the EM-REML procedures and the degree of similarity of the current estimate to the REML estimate

Sample #	No. of iterations		Degree of similarity*	
	Current PE	EM-REML	$\hat{\sigma}_a^2$	$\hat{\sigma}_e^2$
1	114	855	+3.0	-4.0
2	242	1,456	+4.9	-15.6
3	257	3,706	-0.8	-0.4
4	118	841	+0.1	+1.5
5	115	749	+31.4	-29.2
6	95	1,072	+0.5	+1.6
7	112	917	-1.6	+1.0
8	194	1,662	+36.9	-20.7
9	353	1,548	+0.7	-2.3
10	117	778	-9.2	+7.3

*{(current estimate - REML estimate)/REML estimate} $\times$ 100(%).

corresponding REML estimates. Obviously, the convergence in the current procedure was quite faster than that in EM-REML. With certain samples, the current, final estimates of the two components of variance were very akin to the REML estimates, while it was observed with some samples that the current estimates were about 20 to 30 % different from the REML estimates.

5. Discussion

In this article, we report the results of a preliminary study on the PE procedures using animal models for variance component estimation. First, it is proved that no PE approach using IAM is able to be defined under the setting where each animal has only one record. The present finding is not available in the literature before this work. Incorporating the information such as the additive relationship matrix for animals into the mixed model equations enables us to conduct the estimation of variance components with the (RE)ML procedures using IAM. Contrary to this, however, it is revealed here that even if such information is used, the PE procedure of Thompson (1969) and Schaeffer (1986) does not perform under the situation mentioned above, since the estimate of the variance ratio is never renovated from one to the next round of iteration with the PE scheme.

Then, in a particular case where the data analyzed are carcass records, we have attempted to utilize RAM, instead of IAM, for establishing a PE procedure. This attempt would be justified, noting that RAM is an equivalent model to IAM under the use of the specific data like carcass records. The current PE approach is a simple application of Thompson's (1969) and Schaeffer's (1986) PE algorithm.

Actually, the current procedure worked numerically, and the convergence of the estimates was successfully attained in all the replicates. Also, relative to EM-REML, rapid convergence was realized, as would be expected. Dealing with the balanced case, Yamagishi and Iwaisaki (1998) showed that with EM-REML the function representing the relationship between two consecutive estimates of the variance ratio in the iteration is non-linear, but the function is definitely linear in the PE estimation. Using this theoretical result, they also proved that faster convergence always occurs with the PE approach compared to EM-REML under the use of balanced data.

In the unbalanced case, the relationship would be no longer linear with the PE approach, and the pattern of the convergence curve would be dependent on the factors such as the degree of unbalancedness, the number of observations per level of the random factor and how the PE procedure approximates to the algorithm of EM-REML. On this point, the previous simulation studies showed that the PE procedures converged considerably faster than EM-REML even under unbalanced data, as is the current case (e.g., Ouweltjes et al., 1988; Van Raden and Jung, 1988; Iwaisaki, 1992).

However, the results in the present numerical analysis strongly indicate that the current estimator could have a larger sampling variance. In addition, it is quite likely that the similarity of the current estimate to the REML estimate is definitely affected by the given data-structure,

suggesting that in spite of relatively fast convergence, the current PE procedure as it stands could not be a practical one.

Hence, when our purpose is to use a PE estimate as an initial value in the REML estimation, further investigation is obviously needed to improve the current computing strategy, or to establish another PE approach more mimicking REML. For the balanced case, on the other hand, it has been shown that the convergence curve of the PE estimate can be exactly represented using a result for the geometric series and consequently the PE convergence curve is equivalent to the convergence curve assumed underlyingly in the acceleration techniques like the Aitken extrapolation (Yamagishi and Iwaisaki, 1999). Therefore, further examination would also include the application of the extrapolation techniques to the REML estimation under unbalanced data.

REFERENCES

- Bertrand, J. K. and Kriese, L. A. (1990). Two methods for parameter estimation using multiple-trait models and beef cattle field data. *Journal of Animal Science* **68**, 2310-2318.
- Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* **72**, 320-338.
- Henderson, C. R. (1953). Estimation of variance and covariance components. *Biometrics* **9**, 226-252.
- Henderson, C. R. (1986). Estimation of variances in animal model and reduced animal model for single traits and single records. *Journal of Dairy Science* **69**, 1394-1402.
- Iwaisaki, H. (1992). A modified pseudoexpectation method for variance component estimation. *Journal of Animal Breeding and Genetics* **109**, 90-97.
- Khuri, M. G. and Sahai, H. (1985). Variance components analysis: A selective literature survey. *International Statistical Review* **53**, 279-300.
- Ouweltjes, W., Schaeffer, L. R., and Kennedy, B. W. (1988). Sensitivity of methods of variance component estimation to culling type of selection. *Journal of Dairy Science* **71**, 773-779.
- Patterson, H. D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**, 545-554.
- Quaas, R. L. (1976). Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* **32**, 949-953.
- Reverter, A., Golden, B. L., Bourdon, R. M., and Brinks, J. S. (1994). Method R variance component procedure: application on the simple breeding value model. *Journal of Animal Science* **72**, 2247-2253.
- Schaeffer, L. R. (1986). Pseudo expectation approach to variance component estimation. *Journal of Dairy Science* **69**, 2884-2889.

- Searle, S. R., Casella, G., and McCulloch, C. E. (1992). *Variance components*. New York: John Wiley & Sons.
- Shi, M. J., Laloë, D., Ménéssier, F., and Renand, G. (1993). Estimation of genetic parameters of preweaning performance in the French Limousin cattle breed. *Genetics Selection Evolution* **25**, 177-189.
- Thompson, R. (1969). Iterative estimation of variance components for non-orthogonal data. *Biometrics* **25**, 767-773.
- Van Raden, P. M. and Jung, Y. C. (1988). A general purpose approximation to restricted maximum likelihood: The tilde-hat approach. *Journal of Dairy Science* **71**, 187-194.
- Van Tassell, C. P., Casella, G., and Pollak, E. J. (1995). Effects of selection on estimates of variance components using Gibbs sampling and restricted maximum likelihood. *Journal of Dairy Science* **78**, 678-692.
- Wang, C. S., Rutledge, J. J., and Gianola, D. (1993). Marginal inferences about variance components in a mixed linear model using Gibbs sampling. *Genetics Selection Evolution* **25**, 41-62.
- Yamagishi, K. and Iwaisaki, H. (1998). Convergence property of the estimate of variance ratio in the tilde-hat procedure with balanced data. *Animal Science and Technology* **69**, 90-97.
- Yamagishi, K. and Iwaisaki, H. (1999). Characteristics of convergence curve of the PE estimate of variance ratio in the balanced case. *Animal Science Journal* **70**, 414-420.