Estimation of genetic covariances with Method $\mathfrak{R}$[1]

T. Druet*,†,[2,3], I. Misztal‡, M. Duangjinda‡, A. Reverter§, N. Gengler*,†,[2]

*National Fund for Scientific Research, B-1000 Brussels, Belgium;
†Animal Science Unit, Gembloux Agricultural University, B-5030 Gembloux, Belgium;
‡Department of Animal and Dairy Science, University of Georgia, Athens 30602; and
§Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351, Australia

ABSTRACT: Method $\mathfrak{R}$ is a simple and computationally inexpensive method for estimating (co)variances. The objective of the study was to investigate properties of Method $\mathfrak{R}$ for estimation of (co)variance components with emphasis on covariance estimation. Theoretical Method $\mathfrak{R}$ formulas were developed for simplified single-variate and bivariate models. In single-trait models, the curve of the regression of Method $\mathfrak{R}$ was continuous and monotonic and its slope depended on the amount of information on each animal and on the variance ratio. The curve became steeper as the number of records per animal decreased. For covariance, the curve of the regression was monotonic but not continuous. However, a regression coefficient of 1 still corresponded to the correct covariance. Similar curves were observed in analyses of simulated data sets. Because of the observed discontinuity, algorithms implementing Method $\mathfrak{R}$ that require a continuous regression curve would not work in models with covariances. An alternative algorithm was based on a transformation matrix obtained by multiplying a matrix of numerators with the inverse of a matrix of denominators of the regression factors. Such an algorithm converged reliably for all models tested. Method $\mathfrak{R}$ can be modified to estimate covariances in models too large for other methods.

Key Words: Covariance, Statistical Analysis, Variance Covariance Matrix

Introduction

Method $\mathfrak{R}$ (Reverter et al., 1994) is a method for estimating (co)variances. It is simple and computationally inexpensive. Therefore, it may still be applicable when methods based on REML or MCMC are not computationally feasible. Low cost has made Method $\mathfrak{R}$ the method of choice for estimation of dominance variance in large models (e.g., Misztal, 1997; Misztal et al., 1997; Culbertson et al., 1998).

Theoretical properties of Method $\mathfrak{R}$ are unknown. Reverter et al. (1994) claimed that the estimates were equivalent to estimates of REML based on similar quadratics. Reverter (1994) also developed some empirical evidence of the optimality of Method $\mathfrak{R}$ estimates. However, Schenkel (1999) found empirically that with genetic selection based on BLUP breeding values, heritability estimates given by Method $\mathfrak{R}$ would be biased upward.

Extension of Method $\mathfrak{R}$ to two additive genetic effects is straightforward if these effects are uncorrelated. Gengler et al. (1998) used Method $\mathfrak{R}$ to estimate the dominance variance for postweaning gain in the U.S. Limousin population. Because maternal effects were assumed unimportant, their approach did not require correlated additive genetic effects. Ability to estimate covariances among correlated effects is necessary to implement models with direct and maternal effects as for multiple-trait models. Reverter (1994) developed formulas for covariances, but properties of these formulas have not been verified. The application of these formulas using the numerical approach by Misztal (1997) was not able to update covariances successfully. The objective of this study was to investigate the properties of Method $\mathfrak{R}$ with particular emphasis on formulas for covariances in order to apply the method for (co)variance estimation in beef cattle.

Materials and Methods

Method $\mathfrak{R}$ Theory and Existing Algorithms to Obtain (Co)Variances

Reverter et al. (1994) presented Method $\mathfrak{R}$ for a two-trait model, and similar methods can be used with corre-

lated effects. Let $\mathbf{u}_i \sim N(0, \mathbf{A}\sigma_i^2)$ and $\mathbf{u}_j \sim N(0, \mathbf{A}\sigma_j^2)$ be vectors of the additive genetic random effect for traits (or effects) i and j, and $\mathbf{A}$ is the numerator relationship matrix. If $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{u}}_j$ represent vectors of estimated solutions with a complete data set and $\hat{\mathbf{u}}_{ip}$ and $\hat{\mathbf{u}}_{jp}$, vectors of estimated solutions with a partial data set, regression factors (r_i) can be defined for variance of trait i as follows:

$$r_i = \frac{\text{Cov}(\hat{\mathbf{u}}_{ip}, \hat{\mathbf{u}}_i)}{\text{Var}(\hat{\mathbf{u}}_{ip})} = \frac{\hat{\mathbf{u}}_{ip}' \mathbf{A}^{-1} \hat{\mathbf{u}}_i}{\hat{\mathbf{u}}_{ip}' \mathbf{A}^{-1} \hat{\mathbf{u}}_{ip}} \quad [1]$$

For the covariance between both traits two regression factors can be defined:

$$r_{ij} = \frac{\text{Cov}(\hat{\mathbf{u}}_{ip}, \hat{\mathbf{u}}_j)}{\text{Cov}(\hat{\mathbf{u}}_{ip}, \hat{\mathbf{u}}_{jp})} = \frac{\hat{\mathbf{u}}_{ip}' \mathbf{A}^{-1} \hat{\mathbf{u}}_j}{\hat{\mathbf{u}}_{ip}' \mathbf{A}^{-1} \hat{\mathbf{u}}_{jp}}; \quad r_{ji} = \frac{\hat{\mathbf{u}}_{jp}' \mathbf{A}^{-1} \hat{\mathbf{u}}_i}{\hat{\mathbf{u}}_{jp}' \mathbf{A}^{-1} \hat{\mathbf{u}}_{ip}} \quad [2]$$

Reverter (1994) used the mean of the two regression factors presented in Eq. [2] as the regression factor for covariances. When (co)variances are underestimated and overestimated, regressions factors are expected to be greater than and less than 1, respectively. Method $\Re$ estimates are obtained when all regression factors equal 1 (Reverter et al., 1994).

Several computational algorithms were developed to find such estimates. First, Reverter et al. (1994) proposed a binary iteration strategy and a linear extrapolation strategy. Reverter (1994) also presented a multiplicative iterative algorithm to update (co)variances for traits i and j:

$$\mathbf{G}^{n+1} = \mathbf{G}^n * \begin{bmatrix} r_i^n & r_{ij}^n \\ r_{ji}^n & r_j^n \end{bmatrix} \quad [3]$$

where $\mathbf{G}^n$ is the estimate of the genetic covariance matrix (2×2) in round n, r_i^n is the regression factor for trait i obtained with $\mathbf{G}^n$ as prior genetic covariance matrix, and * is the element-by-element multiplication (Hadamard product). Many rounds of iterations were necessary for convergence. Misztal (1997) used the secant method, which is a linear extrapolation strategy in which derivatives were approximated by numerical differentiation. This method could achieve convergence in fewer rounds of iterations, but heuristics were required to avoid estimates outside of the parameter space. A multivariate form of formula [3], by splitting the components of the regression factors, results in the following formula:

$$\mathbf{G}^{n+1} = \begin{bmatrix} \hat{\mathbf{u}}_{ip}' \mathbf{A}^{-1}\hat{\mathbf{u}}_i & \hat{\mathbf{u}}_{jp}' \mathbf{A}^{-1}\hat{\mathbf{u}}_i \\ \hat{\mathbf{u}}_{ip}' \mathbf{A}^{-1}\hat{\mathbf{u}}_j & \hat{\mathbf{u}}_{jp}' \mathbf{A}^{-1}\hat{\mathbf{u}}_j \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}}_{ip}' \mathbf{A}^{-1}\hat{\mathbf{u}}_{ip} & \hat{\mathbf{u}}_{jp}' \mathbf{A}^{-1}\hat{\mathbf{u}}_{ip} \\ \hat{\mathbf{u}}_{ip}' \mathbf{A}^{-1}\hat{\mathbf{u}}_{jp} & \hat{\mathbf{u}}_{jp}' \mathbf{A}^{-1}\hat{\mathbf{u}}_{jp} \end{bmatrix}^{-1} \mathbf{G}^n \quad [4]$$

where $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{u}}_j$ represent vectors of estimated solutions with a complete data set and $\hat{\mathbf{u}}_{ip}$ and $\hat{\mathbf{u}}_{jp}$ represent vec-

tors of estimated solutions with a partial data. Solutions were estimated with $\mathbf{G}^n$ as prior genetic covariance matrix. Use of this formula would result in the $\mathbf{G}$ matrix being asymmetric during iteration. However, at convergence, the $\mathbf{G}$ matrix should be symmetric, and during iteration the $\mathbf{G}$ matrix can be forced to be symmetric by using average of the off-diagonal elements.

Theoretical Developments of Method $\Re$

Properties of Method $\Re$ were studied theoretically and by simulation. In the theoretical approach, estimates of Method $\Re$ were derived analytically for simplified models. Method $\Re$ is based on regression of predictions derived from full data and partial data. In partial data, estimated solutions are more influenced by prior (co)variances because less information is available. In mixed models, additional information could come from different sources: additional records, additional pedigree information and correlated traits, or effects information. Therefore, three case studies were analyzed for the theoretical approach based on the following: 1) changes in the number of records for the same animals, 2) changes in the pedigree information, and 3) two traits.

Change in Number of Records. Assume a single-trait random model for observations from n_a unrelated animals (Model 1):

$$y_{ij} = u_i + e_{ij} \quad [5]$$

where y_{ij} was record j of animal i, u_i is the animal additive genetic effect for animal i, and e_{ij} is the residual effect. In matrix notation,

$$\mathbf{y} = \mathbf{Zu} + \mathbf{e}. \quad [6]$$

Expectations and variances are

$$E(\mathbf{y}) = 0 \quad [7]$$

$$E(\mathbf{u}) = 0 \quad [8]$$

$$E(\mathbf{e}) = 0 \quad [9]$$

$$\text{var}(\mathbf{u}) = \mathbf{I}\,\sigma_a^2 \quad [10]$$

$$\text{var}(\mathbf{e}) = \mathbf{I}\,\sigma_e^2 \quad [11]$$

$$\text{var}(\mathbf{y}) = (\mathbf{ZZ}'\,\sigma_a^2 + \mathbf{I}\,\sigma_e^2) \quad [12]$$

and the mixed model equations are

$$[\mathbf{Z}'\mathbf{Z} + \mathbf{I}\,\hat{\lambda}]\,\hat{\mathbf{u}} = \mathbf{Z}'\mathbf{y} \quad [13]$$

or

$$\begin{bmatrix} n_r + \hat{\lambda} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & n_r + \hat{\lambda} \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \vdots \\ \hat{u}_{n_a} \end{bmatrix} = \begin{bmatrix} n_r u_1 + \sum_{j=1}^{n_r} e_{1j} \\ \vdots \\ n_r u_{n_a} + \sum_{j=1}^{n_r} e_{n_a j} \end{bmatrix} \tag{14}$$

where n_a is the number of animals, $\hat{\lambda} = \dfrac{\hat{\sigma}_e^2}{\hat{\sigma}_a^2}$ (i.e., the estimated residual to additive genetic variance ratio), $\hat{u}_i$ is the estimate of the effect of animal i, u_i is true value of the effect of animal i, and e_{ij} is the residual effect associated with the record j of animal i ($e_{ij} = y_{ij} - u_i$). The solution for animal i in Eq. [14] is

$$\hat{u}_i = \frac{n_r u_i + \sum_{j=1}^{n_r} e_{ij}}{n_r + \hat{\lambda}} \tag{15}$$

When the number of records per animal is reduced from n_r to αn_r (where α is the proportion of records selected), then the estimate of animal j in Eq. [14] becomes (the subscript p indicating partial data)

$$\hat{u}_{ip} = \frac{\alpha n_r u_i + \sum_{j=1}^{\alpha n_r} e_{ij}}{\alpha n_r + \hat{\lambda}} \tag{16}$$

When $n_a \to \infty$, and after detailed calculations provided in Appendix A, the expected regression factor becomes

$$E(r) = E\left(\frac{\hat{u}_p' A^{-1} \hat{u}}{\hat{u}_p' A^{-1} \hat{u}_p}\right) = \frac{(\alpha n_r + \hat{\lambda})(n_r \sigma_a^2 + \sigma_e^2)}{(n_r + \hat{\lambda})(\alpha n_r \sigma_a^2 + \sigma_e^2)} = 1 + \frac{(1 - \alpha)n_r(\hat{\lambda} - \lambda)}{(n_r + \hat{\lambda})(\alpha n_r + \lambda)} \tag{17}$$

where σ_a^2 is the true additive variance, σ_e^2 is the true variance for residual effects, λ is the true variance ratio, and $\hat{\lambda}$ is the variance ratio based on prior variances.

Change in Information from Relatives. To study properties of Method $\mathfrak{R}$ in models with information on relatives, repeated parent-progeny pairs were used with the model similar to the previous one except that relationships were introduced: $\mathbf{u} \sim N(0, \mathbf{A}\sigma_a^2)$. In the full data set, each animal and its sire had one record each. In the reduced data set, records for parents were eliminated. The mixed-model equations were (Model 2) as follows:

$$\begin{bmatrix} 1 + 4/3\hat{\lambda} & -2/3\hat{\lambda} & \dots & 0 & 0 \\ -2/3\hat{\lambda} & 1 + 4/3\hat{\lambda} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 + 4/3\hat{\lambda} & -2/3\hat{\lambda} \\ 0 & 0 & \dots & -2/3\hat{\lambda} & 1 + 4/3\hat{\lambda} \end{bmatrix} \begin{bmatrix} \hat{u}_{s1} \\ \hat{u}_{p1} \\ \dots \\ \hat{u}_{s\,n_p} \\ \hat{u}_{p\,n_p} \end{bmatrix} = \begin{bmatrix} u_{s1} + e_{s1} \\ u_{p1} + e_{p1} \\ \dots \\ u_{s\,n_p} + e_{s\,n_p} \\ u_{p\,n_p} + e_{p\,n_p} \end{bmatrix} \tag{18}$$

where n_p is the number of pairs of animals ($2n_p$ is the number of animals), subscripts s denotes sires, and p denotes progenies. When only the progenies had records, the mixed models equations were as follows:

$$\begin{bmatrix} 4/3\hat{\lambda} & -2/3\hat{\lambda} & \dots & 0 & 0 \\ -2/3\hat{\lambda} & 1 + 4/3\hat{\lambda} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 4/3\hat{\lambda} & -2/3\hat{\lambda} \\ 0 & 0 & \dots & -2/3\hat{\lambda} & 1 + 4/3\hat{\lambda} \end{bmatrix} \begin{bmatrix} \hat{u}_{s1} \\ \hat{u}_{p1} \\ \dots \\ \hat{u}_{s\,n_p} \\ \hat{u}_{p\,n_p} \end{bmatrix} = \begin{bmatrix} 0 \\ u_{p1} + e_{p1} \\ \dots \\ 0 \\ u_{p\,n_p} + e_{p\,n_p} \end{bmatrix} \tag{19}$$

With $n_p \to \infty$, the expected regression coefficient is as follows:

$$E(r) = \frac{(1 + \lambda)((3.75 + 7\lambda)\sigma_a^2 + (3 + 5\lambda)\sigma_e^2)}{(4\lambda^2 + 8\lambda + 3)1.25(\sigma_a^2 + \sigma_e^2)} = 1 + \frac{(\hat{\lambda} - \lambda)(8\hat{\lambda} + 3)}{5(2\hat{\lambda} + 1)(2\hat{\lambda} + 3)} \qquad [20]$$

Two-Trait Model Based on Number of Records. Because regression formulas for a maternal model were complicated, a bivariate model was used to investigate properties of regressions in models with covariances. The model was similar to Model 1 except that it was extended to two traits (k and l with σ_k^2 and σ_l^2, the genetic variance of trait k and l, respectively, and σ_{kl}, the genetic covariance), the number of records changed from n_r for both traits to α n_r, and $\mathbf{u} \sim N(0, \mathbf{G} \otimes \mathbf{A})$, where

$$\mathbf{G} = \begin{bmatrix} \sigma_k^2 & \sigma_{kl} \\ \sigma_{kl} & \sigma_l^2 \end{bmatrix}; \quad \mathbf{G}^{-1} = \begin{bmatrix} g^{11} & g^{12} \\ g^{12} & g^{22} \end{bmatrix} \qquad [21]$$

For simplification of the model, residual variances were assumed equal to σ_e^2 for both traits, and residual covariance was set to 0. After factoring the residual variances out and defining $\hat{\lambda}_k = \hat{g}^{11}\sigma_e^2$, $\hat{\lambda}_l = \hat{g}^{22}\sigma_e^2$, $\hat{\lambda}_{kl} = \hat{g}^{12}\sigma_e^2$ (where $\hat{g}^{11}$, $\hat{g}^{22}$ and $\hat{g}^{12}$ were defined as in Eq. [21] but with the inverse of the prior $\mathbf{G}$ matrix), the mixed-model equations were (Model 3) as follows:

$$\begin{bmatrix} n_r + \hat{\lambda}_k & \hat{\lambda}_{kl} & \cdots & 0 & 0 \\ \hat{\lambda}_{kl} & n_r + \hat{\lambda}_l & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & n_r + \hat{\lambda}_k & \hat{\lambda}_{kl} \\ 0 & 0 & \cdots & \hat{\lambda}_{kl} & n_r + \hat{\lambda}_l \end{bmatrix} \begin{bmatrix} \hat{u}_{kl} \\ \hat{u}_{ll} \\ \cdots \\ \hat{u}_{kn_a} \\ \hat{u}_{ln_a} \end{bmatrix} = \begin{bmatrix} n_r u_{kl} + \sum_{j=1}^{n_r} e_{k\,lj} \\ n_r u_{ll} + \sum_{i=j}^{n_r} e_{llj} \\ \cdots \\ n_r u_{kn_a} + \sum_{\substack{i=j \\ n_r}}^{n_r} e_{kn_a j} \\ n_r u_{ln_a} + \sum_{i=j} e_{ln_a j} \end{bmatrix} \qquad [22]$$

where u_{ki} is the effect of animal i for trait k, $\hat{u}_{ki}$ is the estimate u_{ki}, and e_{kij} is the residual effect associated with record j of animal i for trait k. To create partial data, the same mixed-model equations were applied but with α n_r records for each animal. Four regression factors had to be computed, one for each variance and two for covariance. The expected regression factors for variances are as follows:

$$E(r_k) = \frac{\dfrac{(\alpha n_r + \hat{\lambda}_l)(n_r + \hat{\lambda}_l)(\alpha n_r^2\sigma_k^2 + \alpha n_r \sigma_e^2) - ((1 + \alpha)n_r + 2\hat{\lambda}_l)(\hat{\lambda}_{kl}\alpha n_r^2)\sigma_{kl} + \hat{\lambda}_{kl}^2(\alpha n_r^2\sigma_l^2 + \alpha n_r \sigma_e^2)}{(n_r + \hat{\lambda}_k)(n_r + \hat{\lambda}_l) - \hat{\lambda}_{kl}^2}}{\dfrac{(\alpha n_r + \hat{\lambda}_l)^2(\alpha^2 n_r^2\sigma_k^2 + \alpha n_r \sigma_e^2) - (2\alpha n_r + 2\hat{\lambda}_l)(\hat{\lambda}_{kl}\alpha^2 n_r^2)\sigma_{kl} + \hat{\lambda}_{kl}^2(\alpha^2 n_r^2\sigma_l^2 + \alpha n_r \sigma_e^2)}{(\alpha n_r + \hat{\lambda}_k)(\alpha n_r + \hat{\lambda}_l) - \hat{\lambda}_{kl}^2}} \qquad [23]$$

and the expected regression factor for covariance is (derived from first equation for covariance [see Eq. (2)]; with the second equation for covariance similar patterns were observed) as follows:

$$E(r_{kl}) = \frac{(\alpha n_r + \hat{\lambda}_k)(\alpha n_r + \hat{\lambda}_l) - \hat{\lambda}_{kl}^2}{(n_r + \hat{\lambda}_k)(n_r + \hat{\lambda}_l) - \hat{\lambda}_{kl}^2} * \qquad [24]$$

$$\frac{(n_r + \hat{\lambda}_k)(\alpha n_r + \hat{\lambda}_l)n_r\sigma_{kl} - (n_r + \hat{\lambda}_k)\hat{\lambda}_{kl}(n_r\sigma_l^2 + \sigma_e^2) - (\alpha n_r + \hat{\lambda}_l)\hat{\lambda}_{kl}(n_r\sigma_k^2 + \sigma_e^2) + \hat{\lambda}_{kl}^2 n_r\sigma_{kl}}{(\alpha n_r + \hat{\lambda}_k)(\alpha n_r + \hat{\lambda}_l)\alpha n_r\sigma_{kl} - (\alpha n_r + \hat{\lambda}_l)\hat{\lambda}_{kl}(\alpha n_r\sigma_l^2 + \sigma_e^2) - (\alpha n_r + \hat{\lambda}_l)\hat{\lambda}_{kl}(\alpha n_r\sigma_k^2 + \sigma_e^2) + \hat{\lambda}_{kl}^2\alpha n_r\sigma_{kl}}$$

where σ_{kl} is additive genetic covariance between trait k and l.

Validation Study by Simulation Approach

Simulated Data Sets. First, in order to confirm the theoretical curves, three data sets were simulated using the animal model and six generations without selection. The first data set (data set SIM-1) was simulated with the model

$$y_{ijk} = \mu + cg_i + sex_j + a_k + e_{ijk} \qquad [25]$$

where y_{ijk} is performance of animal k of sex j in contemporary group i, μ is the mean effect, cg is the contemporary group fixed effect, a is the additive genetic effect, and e is the residual. The number of contemporary groups was 60, the number of animals was 11,254, and the number of records was 10,704. Variances used in the simulation were $\sigma^2_{cg} = 10$, $\sigma^2_a = 40$, and $\sigma^2_e = 60$. The second data set (data set SIM-2) was similar to the first one except that it included the maternal effect m:

$$y_{ijk} = \mu + cg_i + sex_j + a_k + m_l + e_{ijk} \qquad [26]$$

with all variances as in the first model except $\sigma^2_m = 20$, $\sigma_{am} = -10$, and $\sigma^2_e = 50$. The last data set (data set SIM-3) was identical to the previous one but the covariance was positive: $\sigma_{am} = 10$ and $\sigma^2_e = 30$.

Three supplementary data sets were simulated for comparing the solving algorithms. They were created with the same model as data sets SIM-2 and SIM-3 but with different random seeds and with the following (co)variances structure: data set SIM-4, $\sigma^2_a = 20$, $\sigma^2_m = 15$, $\sigma_{am} = -5$ and $\sigma^2_e = 70$; data set SIM-5, $\sigma^2_a = 25$, $\sigma^2_m = 25$, $\sigma_{am} = 0$, and $\sigma^2_e = 50$; data set SIM-6, $\sigma^2_a = 40$, $\sigma^2_m = 20$, $\sigma_{am} = -15$, and $\sigma^2_e = 55$.

Analysis of Simulated Data with Method $\Re$. The computer program for this study was derived from program BLUP90IOD, which was a rewrite of BLUPF90 (Misztal, 1999) to iteration-on-data with preconditioned-conjugate-gradient (Stranden and Lidauer, 1999) by S. Tsuruta. Initially, Method $\Re$ was implemented as described by Misztal (1997); solutions were computed for the complete and partial data sets, and the partial data set was a random 50% selection of the complete data set. After each round of preconditioned gradient method for both data sets, regression factors were computed with Eq. [1] for variance and Eq. [2] for covariance (the mean of both regression factors). This process was repeated until the regression factors changed less than 10^{-6} for three consecutive rounds.

Results and Discussion

Theoretical Developments

The discussion below focuses mainly on estimation of animal variance with Method $\Re$. However, the discussed properties can be extended to estimation of vari-
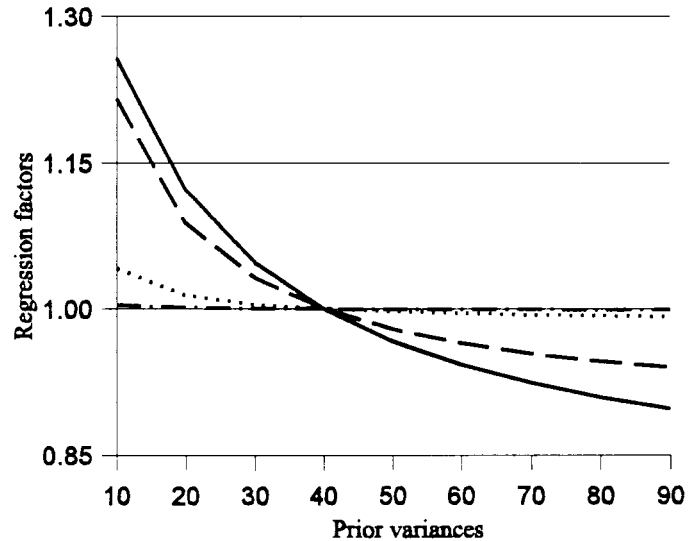

Figure 1. Values of regression factors for different numbers of records per animal (——, 4 records; — —, 10 records; ·······, 100 records; — · — ·, 1,000 records) in the first theoretical model (Model 1).

ance for any random effect (e.g., a permanent environment effect or a random herd-year-season effect). For some types of effects, consequences of the properties could be more important than for an animal effect because, for instance, there is no relationship among different levels or the number of records per level could be much larger.

Change in Number of Records. Equation [17] is in agreement with the theory presented by Reverter et al. (1994). When exact variance ratios were used, regressions factors would be equal to 1. Whether the regression factor was lower or higher than 1 depended on the differences between the used and true variance ratios. According to Reverter et al. (1994), too high a variance ratio (too low a heritability) resulted in a regression factor higher than 1, whereas when animal variances were underestimated, regression factors were lower than 1. At convergence, when the regression factor is equal to 1, the estimated variance ratio is equal to the true variance ratio ($\hat{\lambda} = \lambda$).

Derivation of equation [17] (Appendix B) reveals that regression curves are continuously decreasing ($\alpha - 1$ is always negative) and that their slope is steeper when the number of records decreases or when λ decreases. Shown in Figure 1 are regression factors from equation [17] when the number of records was fixed to 4, 10, 100, or 1,000 for each animal, α was 0.5 (selection of 50% of the records), $\sigma^2_a = 40$, and $\sigma^2_e = 60$. With few records per animal, the curves were continuously decreasing with relatively steep slopes. Finding the intersection with 1 should not be a major challenge in most cases. With many records per animal, the curves were almost horizontal, which could lead to numerical problems. In extreme cases, the regression factors could be numerically computed as 1 for all variance ratios.

Formulas [15] and [16] show clearly that if, through sampling, all records for a specific animal are kept ($\alpha = 1$), then the solutions with partial and complete data sets are identical (if the relationship matrix is not considered). As a consequence, this animal does not contribute to the regression factor deviating from 1 even if incorrect prior variance is used. This is also true when $\alpha = 0$ (e.g., when, through sampling, all records for a specific animal are discarded). Without a relationship matrix, only animals for which some but not all records are eliminated are informative for Method $\Re$ estimation. Therefore, the sampling procedure is important, especially for the estimation of variance for random effects with independent levels. For instance, a sampling by herds could result in problems for estimating a permanent environment variance with Method $\Re$, because in selected herds the same number of records would be used for the estimation of solutions, but in discarded herds solutions for permanent environment of animals would be 0. A random sampling or a sampling by year of record would achieve better results. Finally, the proportion of sampling should be such that a maximum number of animals have their number of records reduced without being equal to 0. For random sampling, 50% sampling seems, therefore, a better sampling proportion than 10% or 90%. With large data sets, sampling should result in a sufficient number of informative animals. Poor sampling or too few informative animals would be reflected in a large sampling distribution between multiple Method $\Re$ estimates. Multiple sampling with Method $\Re$ offers, therefore, an additional security.

Pedigree Information. The value of r as a function of used variances with Eq. [20] is shown in Figure 2. As before, r = 1 when the correct variance ratio was used and the equation is in agreement with Method $\Re$ theory. Information from relatives improved Method $\Re$ estima-
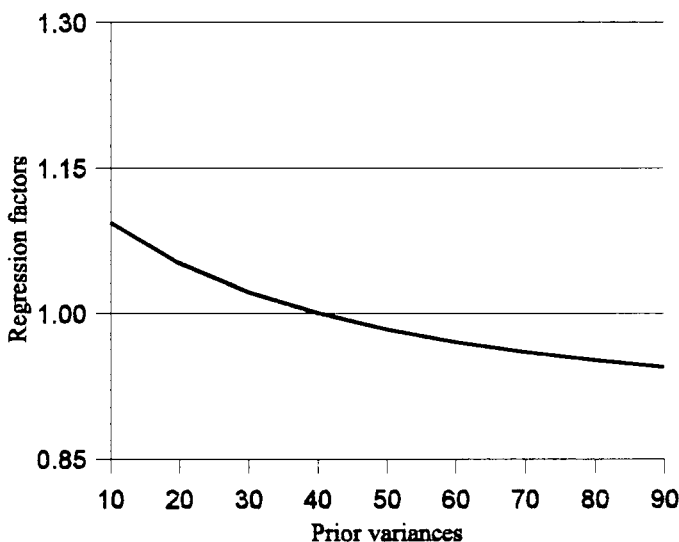

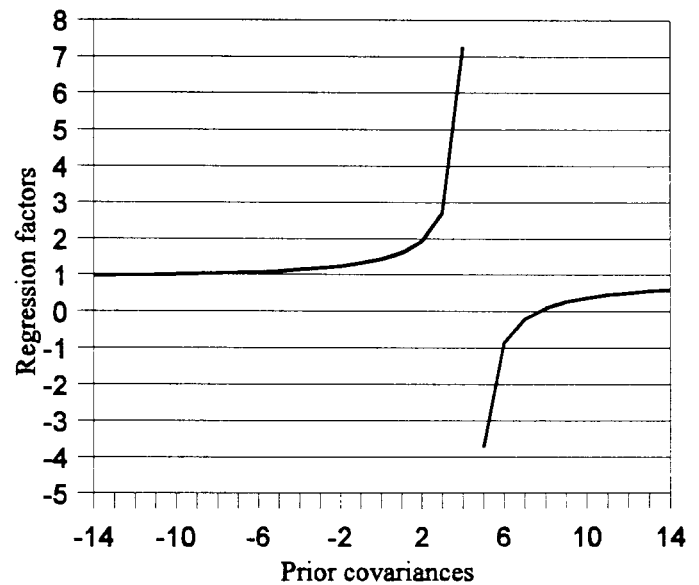**Figure 2**. Values of regression factors in the second theoretical model (Model 2).

Figure 3. Values of regression factors for various values of covariances used in the third theoretical model (Model 3).

tion. For additive genetic effects, individuals are related to each other through the relationship matrix. Animals for which the number of records stays equal or drops to 0 can still be informative for Method $\Re$ through the records of their relatives. The relationship matrix therefore improves the sampling of the records and distributes it over more animals.

Correlated Effects or Traits Information. If the used covariance is equal to 0, then $\hat{\lambda}_{kl} = 0$ and Eq. [23] reduces as expected to Eq. [17] because the bivariate model is then reduced to two independent univariate models. Both Eq. [23] and [24] show that in multivariate cases regression factors for one prior (co)variance component can be influenced by the other prior (co)variance components. Therefore, all regression factors have to be equal to 1 simultaneously and a multivariate converging algorithm is better adapted. Figure 3 shows the regression factor for the prior covariance when correct variances were used using Eq. [24] and when the (co)variances were set to $\sigma_k^2 = 40$, $\sigma_l^2 = 20$, $\sigma_{kl} = -10$, $\sigma_e^2 = 60$, and $\alpha = 0.5$. This curve is different from that for variances because it is monotonic but has a discontinuity. Although the factor was 1 with the correct covariance, the rule of adjusting the covariance downward or upward when the factor was too low or too high would no longer work. Figure 4 shows the numerator and the denominator of the factor. Both curves are almost linear, and they cross at the value of the correct covariance. Figure 5 shows the values of the numerator and the denominator of the regression factor for the variance (with the single-trait model) when different values of prior variances are used. Although the curve for the variance is less linear, the general trend is the same. The difference of regression curves is due to the fact that in a single-
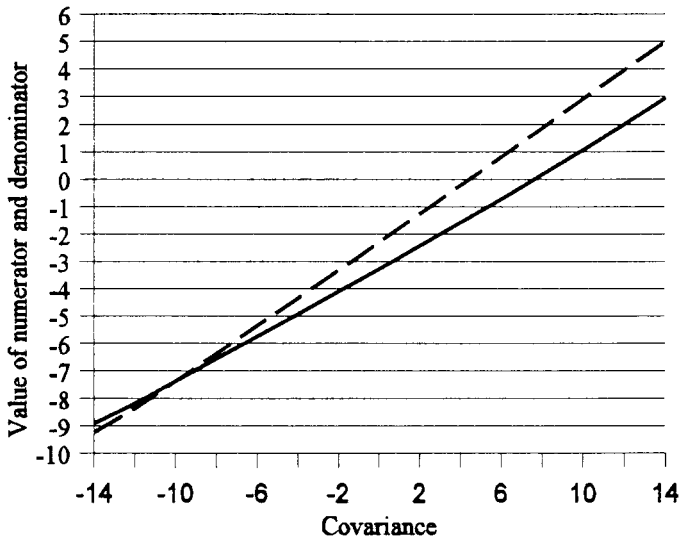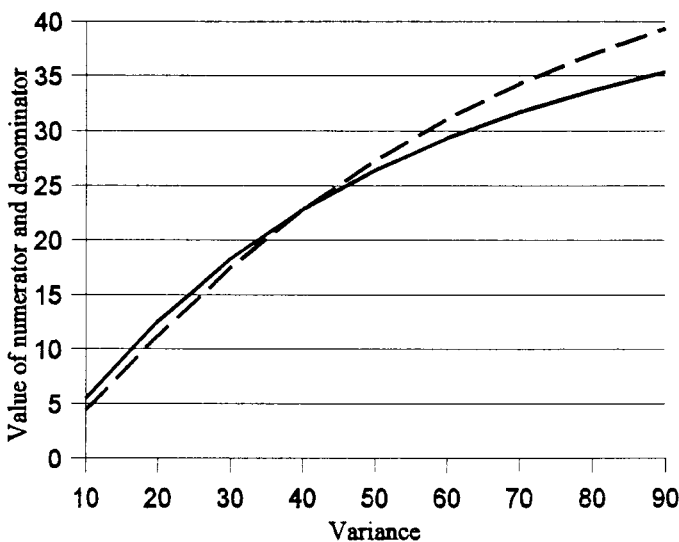
Figure 4. Theoretical values (divided by number of animals) of the numerator (——) and the denominator (— —) of regression factors of covariance when covariances used ranged from −14 to 14 in Model 3.
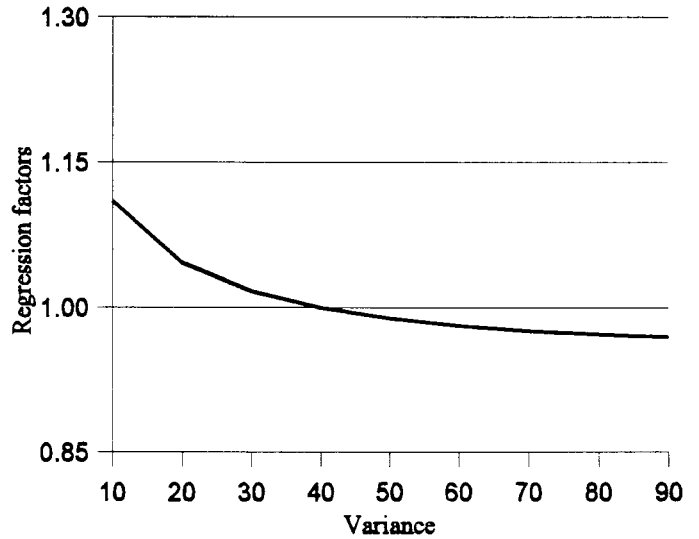

Figure 6. Values of the regression factors with the first simulated data set (SIM-1) for changes in the number of records when the additive variance used varied from 10 to 90 and residual variance was fixed at 60.

trait model the denominator is always positive (Eq. [17]), whereas for covariance it can be negative or equal to 0, generating the discontinuity (Eq. [24]). Therefore, with covariance, $\mathrm{Cov}(\hat{u}_{ip},\hat{u}_j) > \mathrm{Cov}(\hat{u}_{ip},\hat{u}_{jp})$ is no longer equivalent to $r_{ij} > 1$, as in Reverter (1994). Method $\mathfrak{R}$ should rather be expressed as follows: if $\mathrm{Cov}(\hat{u}_{ip},\hat{u}_j) > \mathrm{Cov}(\hat{u}_{ip},\hat{u}_{jp})$, prior (co)variance is understimated and if $\mathrm{Cov}(\hat{u}_{ip},\hat{u}_j) < \mathrm{Cov}(\hat{u}_{ip},\hat{u}_{jp})$, prior (co)variance is overestimated.

Validation Study with Simulated Data

Single Trait with No Maternal Effect. The values of r obtained with the single-trait data set (SIM-1) are shown in Figure 6. The shape of r is very similar to that in the theoretical analyses, and the value of the additive variance is very close to the simulated one. It may be concluded that simple models used in the theoretical analyses were adequate for predicting properties of Method $\mathfrak{R}$ with more complicated models.

Single Trait with Maternal Effect. With the data set SIM-2, regression factors of 1 were obtained for one random sample at $\hat{\sigma}_k^2 = 40.1$, $\hat{\sigma}_l^2 = 18.1$, $\hat{\sigma}_{kl} = -8.6$. Figure 7 presents the regression factor as a function of the covariance when variances were set at their converged values. A subsequent graph in Figure 8 shows the numerator and the denominator of the last factor. Again, these figures are very similar to those obtained in theoretical studies.

With the data set SIM-3 (with the positive covariance), regression factors of 1 were obtained for one random sample at $\hat{\sigma}_k^2 = 40.6$, $\hat{\sigma}_l^2 = 18.1$, $\hat{\sigma}_{kl} = 10.2$. Figure 9 shows the regression factors for different prior covariances obtained with data set SIM-3. Again, r = 1 occurs close to the simulated value of 10; however, the shape of the curve is a mirror image of that in Figure 7, in which the covariance was negative. Both curves are monotonic with the exception of discontinuity, but whether they are increasing or decreasing depends on each case. Also, the discontinuity can correspond to a positive or negative prior correlation. Combination of these situations can result in different consequences on converging algorithms.
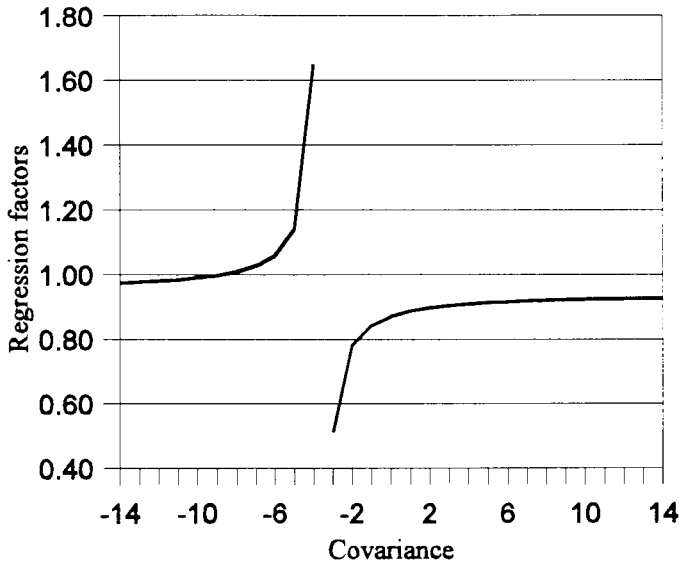

Figure 5. Theoretical values (divided by number of animals) of the numerator (——) and the denominator (— —) of regression factors of variance used when variances ranged from 10 to 90 in Model 1.

Figure 7. Values of regression factors with the second simulated data set (SIM-2, covariance = −10) when the covariance used varied from −14 to 14 (−0.52 to 0.52 genetic correlation) and variances used were the estimated ones for this sample.
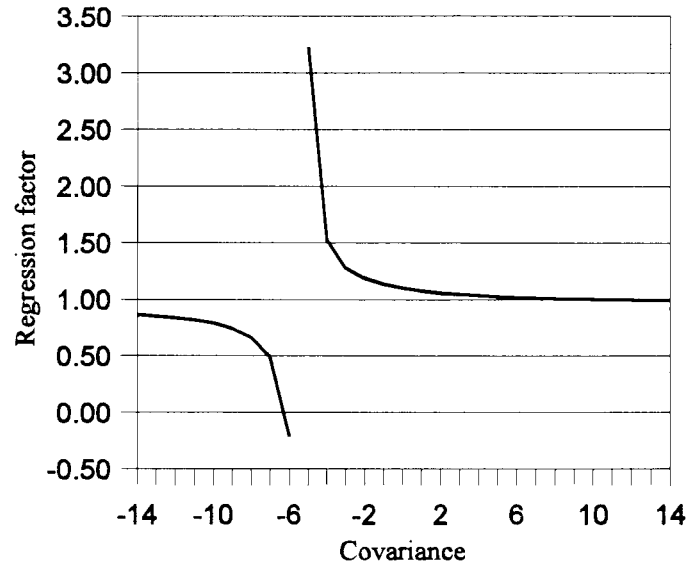

Figure 9. Values of regression factors with the third simulated data set (SIM-3, covariance = 10) when the covariances used varied from −14 to 14 (−0.52 to 0.52 genetic correlation) and the variances used were the estimated ones for this sample.

Consequences of Different Method $\Re$ Algorithms

For models used in this study, convergence was reached at regression factors numerically equal to 1. For covariance, discontinuity and whether the curve is increasing or decreasing has an important effect on computations. First, a binary iteration strategy would not work. The multiplicative iterative algorithm pre-

Figure 8. Values of the numerator (———) and the denominator (— —) of regression factors of covariance used when covariance varied from −14 to 14 (−0.52 to 0.52 genetic correlation) and the variances used were the estimated ones for this sample (data set SIM-2).

sented by Reverter (1994) as in Eq. [3] will not work reliably because the regression factor for covariances can be so large that the new (co)variance matrix may become nonpositive definite. Other problems can also result from application of this method. For instance, if the regression factor is always positive and smaller than 1 for positive prior covariance (e.g., Figure 7), application of the method would lead to a covariance of 0. Problems would also be present with the secant method (Misztal, 1997), in which the convergence would depend on the choice of starting values and heuristics built into that algorithm. One way the secant method and the binary iteration strategy could work would be to use differences between numerators and denominators instead of regression factors, thus avoiding division by 0. An attempt was made in this direction, but it was only partly successful because of the nonlinearity of the differences.
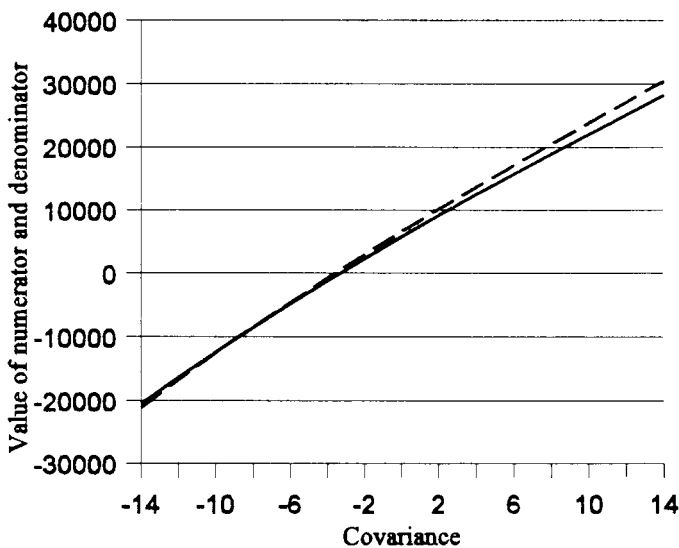
Table 1 shows the total number of preconditioned conjugate gradient rounds needed for convergence with the three different algorithms and 10 sets of starting values for SIM-4, -5, and -6. The multivariate multiplicative iterative algorithm consistently converged to the appropriate parameter estimates, whereas both the secant method (Misztal, 1997) and multiplicative iterative algorithm failed to consistently find a set of parameter estimates where all regression factors were equal to 1. This failure to converge appropriately was caused by the relationship of the initial value of the covariance component with the point of discontinuity in the curve of regression factors and the true value of the covariance. In SIM-5, for example, the multiplicative iterative algorithm moved the solution for the covariance compo-

Table 1. The number of total rounds of preconditioned conjugate gradient algorithm to reach convergence with Method $\Re$ with three different variance components updating algorithms for three samples (data sets SIM-4, -5 and -6) with 10 different starting values for prior (co)variances

Initial prior (co)variances				Data set SIM-4			Data set SIM-5			Data set SIM-6		
σ_e^2	σ_k^2	σ_{kl}	σ_l^2	MIA[a]	MMIA[b]	Secant[c]	MIA	MMIA	Secant	MIA	MMIA	Secant
55.00	40.00	−15.00	20.00	4,718	4,786	3,008	DIV[d]	3,972	DIV	8,124	4,604	DIV
60.00	20.00	5.00	15.00	DIV	2,186	DIV	DIV	5,314	DIV	4,558	5,048	2,112
50.00	30.00	−10.00	30.00	4,558	4,656	3,680	DIV	3,750	DIV	6,854	4,332	DIV
30.00	15.00	15.00	40.00	DIV	3,966	DIV	DIV	2,840	DIV	6,990	4,116	DIV
85.00	10.00	−5.00	10.00	DIV	4,136	DIV	DIV	6,440	DIV	7,040	5,992	DIV
45.01	40.00	−0.01	15.00	DIV	4,842	DIV	DIV	4,400	DIV	8,516	4,078	DIV
65.00	20.00	5.00	10.00	DIV	3,432	DIV	DIV	5,924	DIV	3,252	5,448	DIV
20.00	70.00	0.01	10.00	DIV	7,306	DIV	DIV	7,276	DIV	8,772	3,604	DIV
30.00	40.00	−0.01	30.00	DIV	5,820	DIV	DIV	5,756	DIV	DIV	2,944	DIV
70.00	20.00	−10.00	40.00	3,510	3,598	4,182	DIV	4,822	DIV	7,310	4,830	DIV

[a]MIA = multiplicative iterative algorithm as described in Eq. [3].
[b]MMIA = multivariate multiplicative iterative algorithm as described in Eq. [4].
[c]Secant = secant method (Misztal, 1997).
[d]DIV = divergence, the method was not able to find (co)variances for which regression factors were all equal to 1.

nent toward the point of discontinuity for all initial values <-1.29 and toward 0 for all initial values >-1.29. Therefore, the multiplicative iterative algorithm would not converge.

One disadvantage of the multivariate multiplicative iterative algorithm was its slower convergence compared with the secant method (Misztal, 1997) for models without covariances. Recently, Druet et al. (unpublished data) showed that the multivariate iterative algorithm can be accelerated by increasing updating steps with slight transformation of Eq. [4]:

$$\mathbf{G}^{n+1} = \left(\begin{bmatrix} \hat{\mathbf{u}}_{ip}\mathbf{A}^{-1}\hat{\mathbf{u}}_i & \hat{\mathbf{u}}_{jp}\mathbf{A}^{-1}\hat{\mathbf{u}}_i \\ \hat{\mathbf{u}}'_{ip}\mathbf{A}^{-1}\hat{\mathbf{u}}_j & \hat{\mathbf{u}}'_{jp}\mathbf{A}^{-1}\hat{\mathbf{u}}_j \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}}'_{ip}\mathbf{A}^{-1}\hat{\mathbf{u}}_{ip} & \hat{\mathbf{u}}'_{jp}\mathbf{A}^{-1}\hat{\mathbf{u}}_{ip} \\ \hat{\mathbf{u}}'_{ip}\mathbf{A}^{-1}\hat{\mathbf{u}}_{jp} & \hat{\mathbf{u}}'_{jp}\mathbf{A}^{-1}\hat{\mathbf{u}}_{jp} \end{bmatrix}^{-1} \right)^{\gamma} \mathbf{G}^n$$

[27]

where γ is a constant higher than 1 chosen in order to accelerate convergence. This factor increases the difference between successive updates so that the convergence is reached faster. The optimal value of γ is different for each data set. Too high a value could make the system diverge when changes become so big that the difference between successive updates and the real (co)variance components are increasing.

Implications

Method $\Re$ has to be used cautiously because it could result in poor estimates in several cases: low total number of records, suboptimal proportion of sampled records, poor distribution of sampling of records in effects, or extremely high number of records per level of a random effect. Despite the lack of known theoretical properties, Method $\Re$ was successful for estimation of variance components in studies in which large data sets were essential and other methods were too expensive. Fast implementation of Method $\Re$ was possible because

functions of regression factors for variances were continuous and monotonic. For covariances, functions of regression factors have a discontinuity point, and implementations based on regression factors do not work reliably. One solution is to use a matrix approach in which numerators and denominators are regrouped in a different matrix. One such formula converged reliably for all models tested.

Literature Cited

Culbertson, M. S., J. W. Mabry, I. Misztal, N. Gengler, J. K. Bertrand, and L. Varona. 1998. Estimation of dominance variance in purebred Yorkshire swine. J. Anim. Sci. 76:448–451.

Druet, T., I. Misztal, M. Duangjinda, A. Reverter, and N. Gengler. 2000. Covariance estimation with Method R. J. Dairy Sci. 83(Suppl. 1):56 (Abstr.).

Gengler, N., I. Misztal, J. K. Bertrand, and M. S. Culbertson. 1998. Estimation of the dominance variance for postweaning gain in the U.S. Limousin population. J. Anim. Sci. 76:2515–2520.

Misztal, I. 1997. Estimation of variance components with large-scale dominance models. J. Dairy Sci. 80:965–974.

Misztal, I. 1999. Complex models, more data: Simpler programming. In: Proc. Int. Workshop Comput. Cattle Breed. Tuusala, Finland. Interbull Bull. 20:33–42.

Misztal, I., T. J. Lawlor, and R. L. Fernando. 1997. Dominance models with Method R for stature of Holsteins. J. Dairy Sci. 80:975–978.

Reverter, A. 1994. Method R: A procedure for the estimation of variance and covariance components. Ph.D. dissertation, Colorado State Univ., Fort Collins.

Reverter, A., B. L. Golden, R. M. Bourdon, and J. S. Brinks. 1994. Method R variance components procedure: Application on the simple breeding value model. J. Anim. Sci. 72:2247–2253.

Schenkel, F. S. 1999. Effects of selection on animal genetic evaluation. In: Proc. Int. Symp. on Anim. Breed. and Genet., Viçosa, Brazil. pp 79–99.

Strandén, I., and M. Lidauer. 1999. Solving large mixed linear models using preconditioned conjugate gradient iteration. J. Dairy Sci. 82:2779–2787.

Appendix A

Derivations of regression factors for simplified models were reached using rules derived as follows. For a single trait, the upper part of the regression factors was as follows:

$$\hat{u}_p \, \hat{u} = \sum_{i=1}^{n_a} \hat{u}_{i\,p} \, \hat{u}_i \qquad [28]$$

After multiplication of solutions in Eq. [15] and [16],

$$\hat{u}_p \, \hat{u} = \sum_{i=1} \frac{n_a \alpha n_r^2 u_i u_i + \alpha n_r \sum_{j=1}^{n_r} e_{ij} u_i + n_r \sum_{j=1}^{\alpha n_r} e_{ij} u_i + \sum_{j=1}^{n_r} \sum_{n=1}^{\alpha n_r} e_{ij} \, e_{in}}{(n_r + \lambda)(\alpha n_r + \lambda)} \qquad [29]$$

Expected values of all the used quadratic forms are developed below:

$$E\left(\sum_{i=1}^{n_a} u_i \, u_i\right) = n_a \, \sigma_a^2 \qquad [30]$$

Because the mean for the additive effect is 0,

$$E\left(\sum_{i=1}^{n_a} \sum_{j=1}^{\alpha \, n_r} u_i \, e_{ij}\right) = E\left(\sum_{j=1}^{\alpha \, n_r} \sum_{i=1}^{n_a} u_i \, e_{i\,j}\right) = 0 \qquad [31]$$

and because the covariance between the residuals and the additive effects is 0,

$$E\left(\sum_{i=1}^{n_a} \left(\sum_{j=1}^{n_r} \sum_{n=1}^{\alpha \, n_r} e_{i\,j} e_{i\,n}\right)\right) = \alpha \, n_r \, n_a \, \sigma_e^2 \qquad [32]$$

if $j = n$, then

$$E\left(\sum_{i=1}^{n_a} e_{i\,j} \, e_{i\,n}\right) = n_a \, \sigma_e^2 \qquad [33]$$

but if $j \neq n$, then the same expression is expected to be equal to 0 because different residuals are independent.

Similar developments and rules can be applied for the lower part of the regression factors. For the model involving animal relationships, the same rules were applied, but new rules were added, with new expected values of quadratic forms:

$$E\left(\sum_{i=1}^{n_p} e_{s\,i} \, e_{p\,i}\right) = 0 \qquad [34]$$

because the residuals for each sire and its progeny were independent. In addition,

$$E\left(\sum_{i=1}^{n_p} u_{s\,i} \, u_{p\,i}\right) = n_p \, 0.5 \, \sigma_a^2 \qquad [35]$$

where 0.5 is the additive relationship between a sire and its progeny. Finally, in the multiple-trait model,

$$E\left(\sum_{i=1}^{n_a} e_{k\,i\,j} \, e_{l\,i\,j}\right) = 0 \qquad [36]$$

because residual covariance between trait k and l was assumed 0, and

$$E\left(\sum_{i=1}^{n_a} u_{k\,i}\, u_{l\,i}\right) = n_a\, \sigma_{kl}^2 \tag{37}$$

by definition of the additive covariance, and because means of both traits are assumed equal to 0. Developing Eq. [29] with the rules presented resulted in Eq. [17] or [38]:

$$E(r) = E\left(\frac{\hat{\mathbf{u}}_p'\mathbf{A}^{-1}\hat{\mathbf{u}}}{\hat{\mathbf{u}}_p'\mathbf{A}^{-1}\hat{\mathbf{u}}_p}\right) = \frac{\dfrac{n_a(\alpha n_r{}^2\sigma_a^2 + \alpha n_r\sigma_e^2)}{(n_r + \lambda)(\alpha n_r + \lambda)}}{\dfrac{n_a(\alpha^2 n_r{}^2\sigma_a^2 + \alpha n_r\sigma_e^2)}{(\alpha n_r + \lambda)(\alpha n_r + \lambda)}} = \frac{(\alpha n_r + \lambda)(\alpha n_r^2\sigma_a^2 + \alpha n_r\sigma_e^2)}{(n_r + \lambda)(\alpha^2 n_r^2\sigma_a^2 + \alpha n_r\sigma_e^2)} \tag{38}$$

Appendix B

Equation [17] can be derived with used variance as variable:

$$r(\hat{\sigma}_a^2) = \frac{(\alpha n_r + \hat{\lambda})(n_r\sigma_a^2 + \sigma_e^2)}{(n_r + \hat{\lambda})(\alpha n_r\sigma_a^2 + \sigma_e^2)} = \frac{\alpha n_r\hat{\sigma}_a^2 + \hat{\sigma}_e^2}{(n_r\hat{\sigma}_a^2 + \hat{\sigma}_e^2)} * \frac{(n_r\sigma_a^2 + \sigma_e^2)}{(\alpha n_r\sigma_a^2 + \sigma_e^2)} \tag{39}$$

$$r(\hat{\sigma}_a^2)' = \frac{(\alpha n_r^2\hat{\sigma}_a^2 + \alpha n_r\,\hat{\sigma}_e^2 - \alpha n_r^2\,\hat{\sigma}_a^2 - n_r\,\hat{\sigma}_e^2)}{(n_r\hat{\sigma}_a^2 + \hat{\sigma}_e^2)^2} * \frac{(n_r\sigma_a^2 + \sigma_e^2)}{(\alpha n_r\sigma_a^2 + \sigma_e^2)} = \frac{(\alpha - 1)n_r\,\hat{\sigma}_e^2}{(n_r\hat{\sigma}_a^2 + \hat{\sigma}_e^2)^2} * \frac{(n_r + \lambda)}{(\alpha n_r + \lambda)} \tag{40}$$