Unconstrained Procedures for the Estimation of Positive Definite Covariance Matrices Using Restricted Maximum Likelihood in Multibreed Populations[1]

M. A. Elzo

Animal Science Department, University of Florida, Gainesville 32611

ABSTRACT: Two unconstrained procedures to ensure that intrabreed and interbreed genetic and environmental covariance estimates from multibreed populations are computed within the permissible ranges were developed. These procedures were called Partial Scoring and Cholesky Maximization. The Partial Scoring procedure uses partial steps to keep estimates of covariance matrices positive definite at each expectation-maximization (EM) iteration, and the Cholesky Maximization procedure achieves the same goal by computing the elements of the Cholesky Decomposition of each intrabreed and interbreed genetic and environmental covariance matrix. Groups of small simulated data sets containing either direct genetic effects for two traits (90 bulls, 13,500 calves) or direct and maternal genetic effects for a single trait (135 bulls, 32,400 calves) were used to test the computational feasibility of these two procedures. The overall means (and ranges) of the numbers of expectation-maximization iterations, times to convergence, and accuracy of estimation were 10 (2 to 184), 26.2 min (4.1 to 773.2 min), and 40.1% (12.7 to 81.9%) for the Partial Scoring procedure and 7 (3 to 37), 16.7 min (9.5 to 64.6 min), and 37.8% (3.1 to 67.8%) for the Cholesky Maximization procedure. Although the overall accuracy of both procedures was similar, the Cholesky Maximization procedure should be preferred because it converged faster and its covariance estimates were less affected by the values of the covariance priors than those computed using the Partial Scoring strategy. Application to large unbalanced multibreed data sets will require an iterative version of these procedures.

Key Words: REML, Variance Components, Genetic Parameters, Multibreed Population

J. Anim. Sci. 1996. 74:317–328

Introduction

Interest in the evaluation of straightbred and crossbred animals in a multibreed population context has increased in recent years. A format for interbreed data exchange has already been proposed (Golden and Bourdon, 1994). This data recording system will provide the appropriate infrastructure to maintain the large database needed for national and international multibreed genetic evaluations.

Currently, there are 1) mixed models for predicting additive and nonadditive genetic effects (Elzo and Famula, 1985; Arnold et al., 1992), 2) procedures to directly compute the inverse of the additive (Elzo, 1990a) and nonadditive (Elzo, 1990b) multibreed genetic covariance matrices, and 3) multibreed REML (**MREML**) procedures to compute additive, nonadditive, and environmental covariances using data from straightbred and crossbred animals (Elzo, 1994). Unfortunately, the expectation-maximization (**EM**) algorithm used to compute the MREML covariances (**MREMLEM** procedure) does not guarantee that 1) estimates of all genetic and environmental variances will be positive (which implies that heritability estimates will be greater than 0 and less than 1), and 2) all correlation estimates will be greater than −1 and less than 1 (Elzo, 1994). These two conditions will be met if all matrices of covariance estimates are symmetric positive definite.

The approaches to be used here to ensure that all estimated genetic and environmental covariance matrices are positive definite are unconstrained (i.e., no inequality constraints are used in the maximization of the likelihood). Thus, the objective of this research was to develop two unconstrained strategies to ensure that covariance matrices estimated using MREMLEM procedures are positive definite.

[1]Florida Agricultural Experiment Station journal series no. R-04478. I thank L. D. Van Vleck, R. L. Quaas, and C. B. Moss for useful discussions, and R. C. Littell, T. A. Olson, and C. Vargas for reviewing the manuscript. I also thank the reviewers for their valuable suggestions and for pointing out the works of Lindstrom and Bates (1986), Groeneveld (1994), and Meyer (1994).

Received April 20, 1995.

Accepted October 19, 1995.

Development of the Unconstrained Strategies for the Estimation of Positive Definite Covariance Matrices

The terminology, assumptions, and the basic MREMLEM methodology used for the computation of covariance matrices were the same as those described in Elzo (1994). Traits are assumed to be determined by a large number of unlinked loci, there are random segregation and assortment of alleles and absence of inbreeding, and covariance values do not change over time. The basic MREMLEM procedure uses a bull model and an EM algorithm to simultaneously compute genetic additive, genetic nonadditive, and environmental covariance matrices.

The main objective of the unconstrained strategies is the modification of the set of equations to be solved in the maximization step (M-step) of the basic MREMLEM algorithm, so that the estimates of all genetic and environmental covariance matrices at each EM iteration are positive definite. To achieve this objective, it is important that the MREMLEM equations be well behaved. However, this may not always be the case. The MREMLEM equations have an *intrinsic* degree of multicollinearity. The term *intrinsic* is used because products of matrices whose elements have known relationships among themselves are used to compute the matrices whose traces are the elements of the MREMLEM equations. These matrices are 1) the matrices of derivatives of the multibreed genetic and environmental matrices that are formed by elements (expected fractions of breeds) that have known relationships among themselves (e.g., the expected breed fraction in a progeny is the mean of the corresponding values in the parents), and 2) the multibreed genetic and environmental covariance matrices themselves, which are computed as linear combinations of intrabreed and interbreed genetic and environmental covariances, respectively. The more similar the intrabreed and interbreed genetic and environmental covariance matrices are, the higher the degree of multicollinearity of the MREMLEM equations. The most extreme case would be if all covariance matrices were equal, which would cause complete confounding among intrabreed and interbreed equations (in this case a single covariance matrix should be estimated).

The degree of multicollinearity of the MREMLEM equations will vary widely, depending on the structure of the multibreed data set and the values of the sets of covariances of the various intrabreed and interbreed additive genetic, nonadditive genetic, and environmental covariances. For some data sets the algorithm given in Elzo (1994) can be used to obtain multibreed REML covariance estimates with one additional step to ensure that estimates of covariance matrices are positive definite. In other cases, however, additional steps will be required to reduce the negative effects of the ill conditioning of the MREMLEM equations such that the covariance estimates are either REML or reasonable approximations to REML (**quasi-REML**).

The unconstrained strategies for the computation of MREMLEM covariance estimates presented next incorporate a check for multicollinearity (computation of the eigenvalues of the MREMLEM equations) and an approach to reduce the ill-conditioning it causes on the set of equations solved in the M-step. Small simulated data sets will be used to explore the ability of these methods to simultaneously estimate additive and nonadditive genetic as well as environmental covariances.

Unconstrained Strategies

The unconstrained strategies to compute symmetric positive definite covariance matrices studied here were 1) use of partial steps during the computation of the vector of intrabreed and interbreed covariances by Scoring iterations in the M-step of the i^{th} EM iteration of MREMLEM (*Partial Scoring Strategy*), and 2) estimation of the elements of the Cholesky Decomposition (Golub and Van Loan, 1989) of the intrabreed and interbreed additive genetic, nonadditive genetic and environmental matrices, followed by computation of the corresponding covariance matrices by multiplication of the Cholesky matrices by their transposes (*Cholesky Maximization Strategy*).

Partial Scoring Strategy

This strategy is based on one described by Jennrich and Schluchter (1986), which in turn is a version of a generalized EM (**GEM**) algorithm (Dempster et al., 1977). An important characteristic of a GEM algorithm is that an increase in the value of the expected log-likelihood of the complete data in a given step guarantees an increase in the log-likelihood of the complete data in that step (Jennrich and Schluchter, 1986; Laird et al., 1987). Thus, because the Scoring iterations within an M-step are a series of GEM steps, then an increase in the value of the expected log-likelihood of the complete data rather than the log-likelihood of the complete data could be checked at each Scoring iteration.

The Partial Scoring algorithm used here differs from that of Jennrich and Schluchter (1986) in that 1) partial steps are not only used to increase the value of the expected log-likelihood of the complete data (Equation [3], Appendix, Elzo, 1994) during the Scoring iterations in the M-step but also to ensure that all estimated covariance matrices are positive definite, 2) solutions to the MREMLEM equations are computed by Scoring iterations, using partial steps as needed, so that the expected log-likelihood of the complete data is maximized within each EM iteration, rather than by the first approximation that increases the expected log-likelihood of the complete data (Scoring step), and 3) ridge regression procedures

(Hoerl and Kennard, 1970) are used to improve the condition of the MREMLEM equations when the reciprocal of the condition number (the ratio of the smallest to the largest eigenvalue) is close to or smaller than the machine's floating point precision (e.g., 10^{-6} for single precision). The strategy used here is 1) to check the reciprocal of the condition number (**RCN**) each time new estimates of covariances were computed, and 2) to add the value of the smallest eigenvalue to the diagonal of the MREMLEM equations as many times as needed to increase RCN above single machine precision (e.g., 10^{-5}). If nothing were added to the diagonal of the MREMLEM equations this procedure should converge to a set of REML covariance estimates, provided that the MREMLEM equations were well behaved. If small positive numbers were added to the diagonal elements of the MREMLEM equations at convergence, then the set of covariance estimates will be biased and not REML. However, these quasi-REML covariance estimates will probably be closer to the population values than those obtained without the use of ridge regression procedures applied to the M-step.

The steps of the partial scoring algorithm are as follows:

Step 0.

1) Input an initial set of intrabreed and interbreed additive and nonadditive genetic covariances and environmental covariances (ϕ).

2) Compute the matrices of derivatives of the multibreed additive genetic, nonadditive genetic, and residual covariance matrices with respect to ϕ.

3) Compute the expected log-likelihood of the complete data (Equation [3], Appendix, Elzo, 1994), or, if feasible, the log-likelihood of the complete data (Equation [2], Appendix, Elzo, 1994), using the initial covariance values.

E-Step.

1) Compute the additive genetic, nonadditive genetic, and residual covariance matrices to be used in the mixed-model equations (**MME**) of the bull model.

2) Compute the predicted values of the additive genetic, nonadditive genetic, and residual effects and their respective error variances of prediction (**EVP**).

3) Compute the sums of the products of the vectors of predicted additive and nonadditive genetic effects per bull, and of predicted residuals per calf, times their transposes plus their corresponding EVP matrices.

M-Step.

1) Compute the vector of covariance estimates ϕ for the current EM iteration by Scoring iterations. Within each Scoring iteration perform the following steps.

a) Check that the condition number of the MREMLEM equations is larger than the minimum value chosen (e.g., 10^{-5}); if so, go to M-Step 1b), or else add the value of the smallest eigenvalue to the diagonal elements of the MREMLEM equations. Repeat this step as many times as needed.

b) Check that the estimates of all covariance matrices are positive definite; if not, recompute the estimate of ϕ for the current Scoring iteration using a partial step. Decrease the length of the partial steps until all estimates of covariance are positive definite.

c) Compute the expected log-likelihood of the complete data (or the log-likelihood of the complete data) and compare it with the one from the previous Scoring iteration; if the current value of the expected log-likelihood of the complete data is larger than the one from the previous Scoring iteration, then go to the next Scoring iteration, else use partial steps to increase the value of the expected log-likelihood of the complete data until this condition is met.

d) Check that the absolute difference between vectors of covariance estimates from two consecutive Scoring iterations is less than a vector of small numbers; if so, go to M-Step 2, else continue with the Scoring iterations until convergence is achieved.

2) Check that the vector of absolute differences between the estimates of ϕ from the previous and the current EM iterations is less than a vector of small numbers; if so, stop, else go back to the E-Step of the next EM iteration.

A variation of this algorithm that used a single Scoring step instead of Scoring iterations in the M-Step was also tested. Test runs showed little difference in the values of the covariance estimates between them. However, more EM iterations were required with a single Scoring step than with Scoring iterations.

Cholesky Maximization Strategy

This strategy involves a reparameterization (Bard, 1974; Harville, 1977) of the vector of covariances that introduces a built-in restriction for positive definiteness of the estimates of the covariance matrices in multibreed data sets. This is accomplished by maximizing the expected log-likelihood of the complete data with respect to the elements of the Cholesky Decomposition of the intrabreed and interbreed genetic and environmental covariance matrices (γ). The form of the equations of the M-Step (Equations [9] through [12], Appendix, Elzo, 1994) needs to be modified as follows: 1) $\mathbf{Q_a}$, $\mathbf{Q_n}$, and $\mathbf{Q_v}$ are differentiated with respect to γ_i, and $\mathbf{G}_{0ag}$, $\mathbf{G}_{0nk}$, and $\mathbf{R}^*_{0m}$ are differentiated with respect to either γ_i or γ_j, as

required, in Equations [9] through [12]; 2) the matrices $\mathbf{G}_{0ag}^{-1}\mathbf{G}_{0ag}$, $\mathbf{G}_{0ag}$ written as

$$\sum_{j=1}^{N_a} \frac{\partial G_{0ag}}{\partial \gamma_j} \ (.5 \ \gamma_j),$$

are inserted in the first term of Equation [9], and similar terms are inserted in Equations [10] and [11]; and 3) γ is substituted for ϕ in the first term of Equations [9] through [12]. The resulting nonlinear set of equations is $.5\{B_{ij}\}\ \gamma = \{d_i\}$. A consequence of maximizing with respect to γ is that the matrices of derivatives of Equations [9] through [12] are formed by Cholesky elements from the previous Scoring step; thus, they need to be recomputed at each M-Step.

The Cholesky Maximization strategy used a GEM algorithm (i.e., a single Scoring step that increases the expected log-likelihood of the complete data is required at each M-Step). A check for ill conditioning of the MREMLEM equations and a ridge regression approach to reduce its effects are also included in the algorithm. An algorithm with Scoring iterations in the M-Step was also tested and later dropped, because it was unstable.

The steps of the Cholesky Maximization algorithm are as follows:

Step 0.

1) Input an initial covariance vector ϕ.

2) Compute the expected log-likelihood, or the likelihood, of the complete data using ϕ.

E-Step.

Same as E-Step of the Partial Scoring algorithm.

M-Step.

1) Compute the matrices of derivatives of the multibreed additive and nonadditive genetic and residual covariances with respect to the vector of Cholesky elements γ.

2) Check that the condition number of the MREMLEM set of equations is larger than the chosen minimum value (e.g., 10^{-5}); if not, add the value of the smallest eigenvalue to the diagonal elements of MREMLEM and check again. Repeat this step as needed.

3) Compute the vector of Cholesky elements γ for the Scoring step of the current EM iteration.

4) Compute the vector of covariances ϕ using the Cholesky elements in vector γ.

5) Check that this is a step of a GEM algorithm. Thus, a) compute the expected log-likelihood, or the log-likelihood, of the complete data using the current estimate of vector ϕ; and b) if the current value of the expected log-likelihood of the complete data is larger than the one from the previous EM iteration, then go to M-Step 6, else go back to M-Step 1 and use a partial step to increase its value. Repeat these steps as needed.

6) Check if EM algorithm converged (e.g., by comparing the absolute difference between estimates of vector ϕ in two consecutive iterations); if convergence is achieved, then stop, else go back to the E-Step and continue with the EM iterations.

Independently, Lindstrom and Bates (1986), Groeneveld (1994), and Meyer (1994), also suggested maximizing the log-likelihood with respect to the elements of the Cholesky decomposition of a covariance matrix. These three papers dealt with models for a single population (i.e., all individuals had the same covariance matrix). Two of them (Groeneveld, 1994; Meyer, 1994) applied the Cholesky reparameterization to Newton-Raphson (or Quasi-Newton) and Downhill Simplex (or Derivative Free) algorithms. Lindstrom and Bates (1986) devised Newton-Raphson and EM algorithms based on a QR decomposition (Golub and Van Loan, 1989) of the design matrices of the linear model and on the optimization of the log-likelihood with respect to the Cholesky elements of the covariance matrix of random effects.

Analyses of Small Simulated Data Sets

The purpose of these analyses was to obtain knowledge of the computational behavior of the Partial Scoring and Cholesky Maximization algorithms. All multibreed data sets used here required partial steps and(or) ridge regression during (some of) the EM iterations.

Computations were conducted in an IBM RS6000 workstation, model 580. The computer program used in the computations was an upgraded version of the MREMLEM program (Elzo, 1994), which incorporated the two unconstrained algorithms. The MREMLEM computer program was compiled by the AIX XL FORTRAN Compiler/6000 and was not optimized for either speed or memory and contained numerous checks in various subroutines. The FSPAK routines (M. Perez-Enciso, Univ. of Wisconsin, USA, and UdL-IRTA, Spain; I. Misztal, Univ. of Wisconsin, USA; and M. A. Elzo, Univ. of Florida, USA, personal communication) were included in the MREMLEM program to obtain the inverse of the left-hand side of the MME.

Simulation of Data

Two groups of five non-inbred multibreed data sets were generated: Group 1 had records for two traits and only direct genetic effects, and Group 2 had records for one trait and both direct and maternal genetic effects.

The structure of the data sets was similar in Groups 1 and 2. Two base breeds (A and B) and five breed groups (at 20% intervals) were defined. A total of 90 unrelated bulls and 13,500 calves were generated in each data set of Group 1, and 135 unrelated bulls and 32,400 calves per data set in Group 2. There were six bulls per breed-group combination in the data

sets of Group 1 and nine bulls in the data sets of Group 2. Bulls in both groups were mated to dams of all breed-group combinations to generate calves.

The same set of genetic and environmental effects was simulated in the data sets of Groups 1 and 2. The only fixed effect simulated was sex of calf. All other effects were random. The random effects were three additive direct (intrabreed A, intrabreed B, and interbreed AB), one nonadditive genetic (interbreed AB at one locus), and three environmental effects (intrabreed A, intrabreed B, and interbreed AB).

Covariance Priors

Covariances for each data set were estimated using the Partial Scoring and the Cholesky Maximization procedures. For each procedure, three sets of preliminary covariance values were used: higher than, equal to, and smaller than the covariance values used to simulate the data. Because of multicollinearity, different preliminary covariance values were expected to converge to somewhat different covariance estimates. Given a data set, small differences among covariance estimates across MREMLEM runs started with different preliminary covariance values are an indication that the set of equations solved in the M-step were well behaved and(or) the computational strategy used to account for multicollinearity was effective. The total number of runs was 60, 15 per procedure and group of data sets.

The covariance values used in the simulation of data sets in Group 1 are shown in Table 1, and those used for data sets in Group 2 are in Table 2. For each group of data sets, one pair of covariance matrices was defined for the low-prior runs (one genetic covariance matrix and one environmental covariance matrix), and another pair was defined for the high-prior runs. Each genetic covariance so defined (low or high) was used as the prior for all intrabreed and interbreed additive genetic covariance matrices as well as the interbreed nonadditive genetic covariance matrix. Similarly, each chosen low (or high) environmental covariance was used as prior for all intrabreed and interbreed environmental covariance matrices. For data sets in Group 1, 1) the low preliminary genetic covariance matrix was (1.8, 3.0, 3.0, 18.0), 2) the low environmental covariance matrix was (4.0, 3.0, 3.0, 60.0), 3) the high preliminary genetic covariance matrix was (11.0, 25.0, 25.0, 110.0), and 4) the high preliminary environmental covariance matrix was (28.0, 25.5, 25.5, 235.0). The corresponding four preliminary covariance matrices for data sets of Group 2 were 1) (38.0, 10.0, 10.0, 68.0), 2) (78.0), 3) (155.0, 110.0, 110.0, 320.0), and 4) (400.0).

Convergence, Accuracy, and Number of Expectation-Maximization Iterations

The convergence criterion used in the computer runs was slightly different from the one described for the two unconstrained strategies above. Instead of checking each covariance estimate individually, a single number was computed. This number was **CCONV**, the ratio of the sum of the squares of the absolute differences between covariance estimates from the previous and the current EM iterations to the sum of squares of the covariance estimates of the previous EM iteration. Thus, the convergence criterion was to check whether CCONV was less than a small number (e.g., 10^{-4}). The main advantage of this criterion is its simplicity (a single comparison between unitless numbers). Another ratio, **CCSIM**, the ratio of sum of the squares of the absolute differences between the covariance estimates of the current EM iteration and the covariance values to the sum of squares of the covariance values, was computed to measure the degree of separation between estimates and actual covariance values. Finally, a measure of the degree of closeness between estimated and actual covariance values (**ACSIM**) was computed (mimicking the BIF accuracy formula, BIF, 1990) as (1.0 – square root [CCSIM]) × 100.0. As for CCONV, the main advantage of ACSIM is its simplicity: a single number expressed as percentage. Table 3 shows the values of ACSIM for the Partial Scoring and the Cholesky Maximization procedures within and across prior covariance values and groups of data sets. To illustrate the computation of CCONV, CCSIM, and ACSIM, consider the estimation of three covariances whose actual values are (10, 20, 30). Let their estimates in the previous and the current EM iterations be (8.5, 16.0, 34.6) and (9.0, 18.4, 31.2), respectively. Thus, CCONV = 17.6/1525.4 = .0115, CCSIM = 39.4/1400.0 = .0281, and ACSIM = 83.2%.

The distribution of the number of EM iterations needed to achieve convergence was severely skewed within several covariance prior × estimation procedure subclasses. Thus, the median, rather than the mean, of the number of EM iterations and the time until convergence within covariance prior × estimation procedure subclass are given in Table 3. Also, the means of the number of EM iterations and the times to convergence per computational procedure within and across groups of data sets are the means of the appropriate medians of the high, equal, and low covariance prior × estimation procedure subclasses.

The overall mean of the number of EM iterations needed to achieve convergence was nine, with numbers of EM iterations ranging from 2 to 184. Forty-seven out of 60 analyses (78%) converged in less than 12 EM iterations. Four out of the 60 runs (7%) required more than 20 EM iterations to achieve convergence (three of them used the Partial Scoring method and the remaining one the Cholesky Maximization procedure). However, in all 60 runs at least one of the methods achieved convergence (CCONV < .0001) in less than 20 iterations. Thus, although both

Table 1. Relative means and ranges of covariance estimates for direct genetic effects of two traits from five simulated data sets using two unrestricted procedures and three sets of preliminary covariance estimates

			Partial Scoring			Cholesky Maximization		
Covariance	Trait pair	Value[a]	Low[b]	Equal[c]	High[d]	Low	Equal	High
Additive								
Intrabreed A	(1, 1)	4.0	.6[e] (.2, .8)[f]	.8 (.2, 1.2)	1.2 (.2, 2.3)	.9 (.2, 1.8)	.8 (.2, 1.7)	.9 (.2, 1.8)
	(1, 2)	5.0	.9 (.5, 1.4)	.9 (.5, 1.4)	1.8 (.5, 3.7)	.9 (.5, 1.5)	.9 (.5, 1.6)	.9 (.5, 1.5)
	(2, 2)	40.0	.9 (.6, 1.3)	1.0 (.8, 1.3)	1.4 (.8, 2.2)	1.0 (.7, 1.2)	1.0 (.7, 1.5)	1.0 (.7, 1.3)
Intrabreed B	(1, 1)	8.0	.9 (.4, 1.4)	1.2 (1.0, 1.4)	1.3 (1.1, 1.4)	1.1 (.8, 1.5)	1.0 (.8, 1.3)	1.0 (.7, 1.3)
	(1, 2)	17.0	1.0 (.3, 1.6)	1.3 (1.0, 1.8)	1.3 (1.1, 1.7)	.9 (.8, 1.1)	.9 (.8, 1.0)	.8 (.7, 1.0)
	(2, 2)	60.0	1.0 (.5, 1.6)	1.3 (1.0, 1.6)	1.4 (1.2, 1.6)	.8 (.5, 1.2)	.8 (.5, 1.2)	.8 (.5, 1.2)
Interbreed AB	(1, 1)	2.0	1.5 (.1, 2.6)	1.5 (.4, 2.6)	1.9 (.6, 2.6)	1.8 (.1, 3.9)	1.8 (.2, 3.9)	1.9 (.1, 4.6)
	(1, 2)	4.0	1.5 (.6, 2.6)	1.6 (.9, 2.5)	2.1 (.9, 2.5)	2.3 (.0, 5.1)	2.1 (−.1, 5.2)	2.3 (−.1, 6.2)
	(2, 2)	20.0	1.9 (.3, 3.1)	2.0 (.6, 2.9)	2.4 (.9, 3.2)	2.7 (.0, 4.9)	2.7 (.1, 5.0)	2.8 (.1, 5.5)
Nonadditive								
Interbreed AB (1 locus)	(1, 1)	6.0	.7 (.4, 1.2)	.9 (.7, 1.2)	1.1 (.7, 1.4)	1.0 (.8, 1.5)	1.0 (.8, 1.4)	1.0 (.8, 1.4)
	(1, 2)	20.0	.7 (.3, 1.2)	.9 (.8, 1.2)	1.0 (.8, 1.2)	1.0 (.8, 1.5)	1.0 (.8, 1.4)	1.0 (.8, 1.5)
	(2, 2)	80.0	.7 (.4, 1.1)	.9 (.8, 1.1)	1.0 (.8, 1.2)	1.0 (.8, 1.3)	1.0 (.8, 1.3)	1.0 (.8, 1.3)
Environmental								
Intrabreed A	(1, 1)	6.0	1.0 (.7, 1.2)	1.1 (1.0, 1.2)	1.5 (1.1, 2.0)	1.0 (.7, 1.2)	1.1 (.8, 1.2)	1.1 (.7, 1.4)
	(1, 2)	4.0	1.0 (.8, 1.4)	1.1 (.9, 1.4)	1.5 (1.0, 2.2)	1.2 (.7, 1.4)	1.2 (.7, 1.4)	1.4 (.6, 1.8)
	(2, 2)	90.0	.9 (.8, 1.0)	1.0 (1.0, 1.0)	1.2 (1.0, 1.4)	1.0 (1.0, 1.1)	1.0 (.9, 1.1)	1.1 (1.0, 1.1)
Intrabreed B	(1, 1)	22.0	.7 (.4, 1.0)	.9 (.9, 1.0)	1.0 (.9, 1.1)	1.0 (.8, 1.1)	1.0 (.8, 1.1)	1.0 (.8, 1.1)
	(1, 2)	45.0	.6 (.3, .9)	.9 (.8, 1.0)	.8 (.8, .9)	1.0 (.9, 1.1)	1.0 (.9, 1.0)	1.0 (.9, 1.1)
	(2, 2)	240.0	.7 (.5, .9)	.9 (.8, 1.0)	.9 (.8, 1.0)	1.0 (.9, 1.1)	1.0 (.9, 1.0)	1.0 (.9, 1.0)
Interbreed AB	(1, 1)	14.0	.7 (.5, 1.0)	1.0 (.6, 1.2)	1.1 (.6, 1.7)	1.0 (.5, 1.6)	1.0 (.5, 1.6)	1.0 (.4, 1.6)
	(1, 2)	10.0	.8 (.4, 1.5)	1.0 (.4, 1.5)	1.3 (.4, 2.1)	1.0 (−.5, 2.0)	.9 (−.6, 2.0)	.9 (−.6, 2.1)
	(2, 2)	70.0	.9 (.7, 1.3)	1.0 (.7, 1.3)	1.3 (.7, 2.2)	.9 (.2, 1.6)	.9 (.3, 1.8)	.9 (.2, 1.6)

[a]Values of covariances used to simulate data.
[b]Covariance priors smaller than covariances used to simulate data.
[c]Covariance priors equal to covariances used to simulate data.
[d]Covariance priors larger than covariances used to simulate data.
[e]Mean of five (covariance estimate/covariance value) ratios.
[f](Smallest, largest) value among five (covariance estimate/covariance value) ratios.

methods controlled positive definiteness and increased the value of the log-likelihood of the complete data, in some cases their control of multicollinearity was insufficient to achieve a fast rate of convergence.

The mean of the number of EM iterations needed to achieve convergence with the Partial Scoring procedure (10) was similar to that for the Cholesky Maximization procedure (9). However, their ranges

Table 2. Relative means and ranges of covariance estimates for direct and maternal genetic effects of a single trait from five simulated data sets using two unrestricted procedures and three sets of preliminary covariance estimates

Covariance	Trait pair[a]	Value[b]	Partial Scoring			Cholesky Maximization		
			Low[c]	Equal[d]	High[e]	Low	Equal	High
Additive								
Intrabreed A	(d, d)	40.0	2.1[f] (1.1, 3.0)[g]	2.0 (1.1, 4.9)	3.6 (2.7, 4.9)	3.8 (2.0, 6.6)	1.7 (.5, 2.8)	2.6 (2.0, 3.4)
	(d, m)	30.0	.0 (−.3, .3)	.8 (.1, 1.3)	1.0 (−.9, 2.1)	−.5 (−2.2, .2)	.0 (−.3, .4)	−.8 (−1.2, −.1)
	(m, m)	80.0	.6 (.5, .7)	.9 (.7, 1.0)	1.7 (.7, 2.4)	1.2 (.6, 1.6)	.9 (.4, 1.7)	1.4 (1.2, 1.5)
Intrabreed B	(d, d)	100.0	1.5 (.6, 3.0)	1.2 (.5, 1.7)	1.5 (.5, 2.0)	1.2 (.6, 1.7)	1.2 (.5, 1.6)	1.4 (.7, 1.8)
	(d, m)	90.0	.5 (.3, .6)	.9 (.6, 1.4)	1.0 (.6, 1.3)	.5 (.1, .9)	.7 (.3, 1.5)	.9 (.4, 1.8)
	(m, m)	140.0	.7 (.1, 1.1)	1.0 (.7, 1.4)	1.4 (.7, 1.8)	1.0 (.7, 1.2)	1.1 (.9, 1.4)	1.7 (1.4, 2.2)
Interbreed AB	(d, d)	50.0	.4 (.0, .8)	1.0 (.5, 2.1)	1.1 (.5, 2.6)	2.1 (.2, 5.0)	1.8 (.4, 3.5)	1.4 (.8, 1.9)
	(d, m)	20.0	.0 (−2.9, 2.5)	−2.9 (−5.5, 2.5)	−2.8 (−6.1, 2.5)	−3.6 (−11.0, 1.8)	−5.0 (−9.2, −1.0)	−4.5 (−9.1, .7)
	(m, m)	150.0	.6 (.3, .8)	.8 (.4, 1.1)	.8 (.4, 1.0)	1.6 (.9, 2.9)	1.6 (.7, 3.7)	1.6 (.5, 2.6)
Nonadditive								
Interbreed AB (1 locus)	(d, d)	70.0	1.8 (1.1, 2.5)	2.1 (1.8, 2.5)	2.2 (1.9, 2.8)	1.9 (1.7, 2.3)	2.5 (1.9, 3.4)	2.4 (1.9, 3.4)
	(d, m)	60.0	.2 (−.2, .3)	.8 (.2, 1.3)	.8 (.2, 1.3)	.2 (.1, .3)	.5 (.4, .8)	.6 (.4, .7)
	(m, m)	120.0	.5 (.1, .6)	.7 (.1, .9)	1.3 (.1, 1.9)	.5 (.4, .6)	.6 (.4, .7)	1.2 (1.2, 1.3)
Environmental								
Intrabreed A	(e, e)	200.0	1.2 (.8, 1.4)	1.2 (.8, 1.4)	.8 (.6, 1.2)	1.3 (1.2, 1.4)	1.4 (1.1, 1.6)	1.1 (.9, 1.1)
Intrabreed B	(e, e)	240.0	1.1 (.4, 1.4)	1.2 (1.0, 1.3)	.9 (.6, 1.2)	1.2 (1.1, 1.3)	1.3 (1.0, 1.7)	1.1 (.8, 1.2)
Interbreed AB	(e, e)	180.0	1.0 (.6, 1.3)	1.0 (.9, 1.2)	1.0 (.8, 1.4)	1.1 (.9, 1.5)	1.4 (1.1, 1.8)	.9 (.7, 1.3)

[a]d = direct, m = maternal, and e = environmental.
[b]Values of covariances used to simulate data.
[c]Covariance priors smaller than covariances used to simulate data.
[d]Covariance priors equal to covariances used to simulate data.
[e]Covariance priors larger than covariances used to simulate data.
[f]Mean of five (covariance estimate/covariance value) ratios.
[g](Smallest, largest) value among five (covariance estimate/covariance value) ratios.

were substantially different: 2 to 184 for the Partial Scoring procedure and 2 to 37 for the Cholesky Maximization procedure. This was reflected in the overall mean and range of the time needed to achieve convergence: 26.2 min (4.1 to 773.2 min) for the Partial Scoring method and 16.7 min (9.5 to 64.6 min) for the Cholesky Maximization procedure. The overall accuracy (ACSIM) was higher for the Partial Scoring (42.3%) than for the Cholesky Maximization (37.8%) procedure (Table 3). However, this advantage of the Partial Scoring over the Cholesky Maximization procedure was not uniform across groups of data sets. The Partial Scoring method was 1.3% less accurate for data sets in Group 1 and 9.9% more accurate in data sets of Group 2 than the Cholesky Maximization procedure.

Estimates of Covariances by Unconstrained Procedures

Tables 1, 2, 4, 5, and 6 present various means and ranges of estimates of intra- and interbreed additive genetic, interbreed genetic, and intra- and interbreed environmental covariances. For the sake of clarity, and to allow comparison of differences in estimation across covariances, these same tables present means and ranges of covariance estimated in relative terms,

Table 3. Number of expectation-maximization (EM) iterations, computing times, and accuracy of estimation of covariances using two unrestricted procedures

Data set	Priors[a]	Partial Scoring			Cholesky Maximization		
		Iter[b]	Time[c]	ACSIM[d]	Iter	Time	ACSIM
Two traits, direct genetic effects	Low	14 (2, 15)	29.9 (4.1, 33.6)	43.8 (27.8, 55.6)	9 (8, 17)	16.9 (15.0, 32.0)	53.7 (40.4, 65.5)
	Equal	11 (2, 13)	22.7 (4.1, 25.9)	63.7 (51.2, 81.9)	7 (6, 12)	13.3 (11.4, 22.8)	52.7 (40.7, 66.1)
	High	12 (2, 16)	24.8 (4.1, 35.7)	48.9 (39.2, 55.6)	10 (5, 37)	18.8 (9.5, 64.6)	53.7 (37.7, 67.8)
	All[e]	12 (2, 16)	25.8 (4.1, 35.7)	52.1 (27.8, 81.9)	9 (5, 37)	16.9 (9.5, 64.6)	53.4 (37.7, 67.8)
One trait, direct and maternal genetic effects	Low	9 (6, 23)	39.8 (20.8, 84.0)	32.5 (12.7, 41.2)	4 (4, 5)	14.9 (13.5, 16.7)	21.3 (8.9, 35.6)
	Equal	6 (2, 181)	21.5 (7.5, 756.5)	38.3 (26.9, 48.7)	7 (4, 12)	23.1 (14.0, 38.2)	19.9 (3.1, 38.8)
	High	5 (2, 184)	18.4 (7.1, 773.2)	25.2 (23.0, 28.7)	3 (3, 4)	11.2 (10.1, 12.6)	25.1 (12.5, 35.5)
	All[e]	7 (2, 184)	26.6 (7.1, 773.2)	32.0 (12.7, 48.7)	5 (3, 12)	16.4 (10.1, 38.2)	22.1 (3.1, 38.8)
All	All[e]	10 (2, 184)	26.2 (4.1, 773.2)	40.1 (12.7, 81.9)	7 (3, 37)	16.7 (9.5, 64.6)	37.8 (3.1, 67.8)

[a]Covariance priors smaller than, equal to, and larger than covariances used to simulate data.
[b]Iter = median and range of number of EM iterations until convergence.
[c]Time = median and range of minutes needed until convergence.
[d]ACSIM = Mean and range of the percentage of accuracy of estimated covariances over five simulated data sets.
[e]Mean and ranges of Iter, time, and ACSIM.

Table 4. Relative means and ranges across covariance estimates from two types of data sets using two unrestricted procedures

Data set	Covariance priors[a]	Partial Scoring	Cholesky Maximization
Two traits, direct genetic effects	Low	.9[b] (.1, 3.1)[c]	1.2 (−.5, 5.1)
	Equal	1.1 (.2, 2.9)	1.2 (−.6, 5.2)
	High	1.4 (.2, 3.7)	1.2 (−.6, 6.2)
	All	1.1 (.2, 3.7)	1.2 (−.6, 6.2)
One trait, direct and maternal genetic effects	Low	.8 (−2.9, 3.0)	.9 (−11.0, 6.6)
	Equal	.8 (−5.5, 4.9)	.8 (−9.2, 3.7)
	High	1.1 (−6.1, 4.9)	.9 (−9.1, 3.4)
	All	.9 (−6.1, 4.9)	.9 (−11.0, 6.6)
All	All	1.0 (−6.1, 4.9)	1.1 (−11.0, 6.6)

[a]Covariance priors smaller than, equal to, and larger than covariances used to simulate data.
[b]Mean of (covariance estimate/covariance value) ratios across covariances and simulated data sets.
[c](Smallest, largest) value (covariance estimate/covariance value) ratios across covariances and simulated data sets.

Table 5. Relative means and ranges of covariance estimates for direct genetic effects of two traits from five simulated data sets and three sets of preliminary covariance estimates using two unrestricted procedures

Covariance	Trait pair	Value[a]	Partial Scoring	Cholesky Maximization
Additive				
Intrabreed A	(1, 1)	4.0	.9[b] (.2, 2.3)[c]	.9 (.2, 1.8)
	(1, 2)	5.0	1.2 (.5, 3.7)	.9 (.5, 1.6)
	(2, 2)	40.0	1.1 (.6, 2.2)	1.0 (.7, 1.5)
Intrabreed B	(1, 1)	8.0	1.1 (.4, 1.4)	1.0 (.7, 1.5)
	(1, 2)	17.0	1.2 (.3, 1.8)	.9 (.7, 1.1)
	(2, 2)	60.0	1.2 (.5, 1.6)	.8 (.5, 1.2)
Interbreed AB	(1, 1)	2.0	1.6 (.1, 2.7)	1.8 (.1, 4.6)
	(1, 2)	4.0	1.7 (.6, 2.6)	2.2 (−.1, 6.2)
	(2, 2)	20.0	2.1 (.3, 3.1)	2.8 (.0, 5.5)
Nonadditive				
Interbreed AB (1 locus)	(1, 1)	6.0	.9 (.4, 1.4)	1.0 (.8, 1.5)
	(1, 2)	20.0	.9 (.3, 1.2)	1.0 (.8, 1.5)
	(2, 2)	80.0	.9 (.4, 1.2)	1.0 (.8, 1.3)
Environmental				
Intrabreed A	(1, 1)	6.0	1.2 (.7, 2.0)	1.1 (.7, 1.4)
	(1, 2)	4.0	1.2 (.8, 2.2)	1.2 (.6, 1.8)
	(2, 2)	90.0	1.0 (.8, 1.4)	1.0 (.9, 1.1)
Intrabreed B	(1, 1)	22.0	.9 (.4, 1.1)	1.0 (.8, 1.1)
	(1, 2)	45.0	.8 (.3, 1.1)	1.0 (.9, 1.1)
	(2, 2)	240.0	.8 (.5, 1.0)	1.0 (.9, 1.1)
Interbreed AB	(1, 1)	14.0	1.0 (.5, 1.7)	1.0 (.4, 1.6)
	(1, 2)	10.0	1.0 (.4, 2.1)	1.0 (−.6, 2.1)
	(2, 2)	70.0	1.0 (.7, 2.1)	.9 (.2, 1.8)

[a]Values of covariances used to simulate data.
[b]Mean of 15 (covariance estimate/covariance value) ratios.
[c](Smallest, largest) value among 15 (covariance estimate/covariance value) ratios.

i.e., each mean and range of a covariance estimate was divided by its actual value. To obtain actual means and ranges of covariance estimates, relative means and ranges within covariances need to be multiplied by the covariance values (Tables 1, 2, 5, and 6).

Covariance Means and Ranges Within Prior Values. One of the reasons for testing three different sets of prior covariance values was to determine whether, in the absence of reasonable priors (e.g., covariance estimates from intrabreed analyses), it would be

Table 6. Relative means and ranges of covariance estimates for direct and maternal genetic effects of a single trait from five simulated data sets and three sets of preliminary covariance estimates using two unrestricted procedures

Covariance	Trait pair[a]	Value[b]	Partial Scoring	Cholesky Maximization
Additive				
Intrabreed A	(d, d)	40.0	2.6[c] (1.1, 4.9)[d]	2.7 (.5, 6.6)
	(d, m)	30.0	.6 (−.9, 2.1)	−.4 (−2.2, .4)
	(m, m)	80.0	1.0 (.5, 2.4)	1.2 (.4, 1.7)
Intrabreed B	(d, d)	100.0	1.4 (.5, 3.0)	1.3 (.5, 1.8)
	(d, m)	90.0	.8 (.3, 1.3)	.7 (.1, 1.8)
	(m, m)	140.0 (.1, 1.8)	1.0 (.7, 2.2)	1.3
Interbreed AB	(d, d)	50.0	.8 (.0, 2.6)	1.8 (.2, 5.0)
	(d, m)	20.0	−1.9 (−6.1, 2.5)	−4.4 (−11.0, 1.8)
	(m, m)	150.0	.7 (.3, 1.1)	1.6 (.5, 3.7)
Nonadditive				
Interbreed AB (1 locus)	(d, d)	70.0	2.1 (1.5, 2.6)	2.3 (1.7, 3.4)
	(d, m)	60.0	.6 (−.2, 1.3)	.4 (.1, .8)
	(m, m)	120.0	.8 (.1, 1.9)	.8 (.4, 1.3)
Environmental				
Intrabreed A	(e, e)	200.0	1.0 (.6, 1.4)	1.3 (.9, 1.6)
Intrabreed B	(e, e)	240.0	1.0 (.4, 1.4)	1.2 (.8, 1.7)
Interbreed AB	(e, e)	180.0	1.0 (.6, 1.4)	1.1 (.7, 1.8)

[a]d = direct, m = maternal, and e = environmental.
[b]Values of covariances used to simulate data.
[c]Mean of 15 (covariance estimate/covariance value) ratios.
[d](Smallest, largest) value among 15 (covariance estimate/covariance value) ratios.

preferable to start the EM iterations with priors that underestimated covariances rather than with priors that overestimated the actual covariance values. The ACSIM values of Table 3 show that the Cholesky Maximization procedure was less affected by preliminary covariance values than the Partial Scoring procedure. Furthermore, comparison of covariance estimates from low, equal, and high priors (Tables 1, 2, and 4) suggests that the covariance estimates by the Partial Scoring method tended to follow the prior covariance values; low priors produced covariance estimates that were (on the average) lower than the covariance values, high priors produced estimates higher than covariance values, and priors equal to the covariance values yielded the best covariance estimates. This occurred because the Partial Scoring procedure frequently required a large number of partial steps to generate a complete set of positive definite estimates of covariance matrices; thus, the resulting step was too small to appreciably change the values of the estimates of the covariance matrices from the previous Scoring iteration. The Cholesky Maximization procedure, on the other hand, showed little or no influence of the prior covariances on the covariance estimates (Tables 1, 2, and 4). This made covariance estimates by the Cholesky Maximization procedure for individual data sets more credible than those estimated by the Partial Scoring approach. Thus, the Cholesky Maximization procedure should be preferred to the Partial Scoring procedure, particularly for traits without reliable covariance priors.

Intrabreed and Interbreed Covariance Means and Ranges. The ratios of individual intra- and interbreed genetic and environmental covariances are presented by procedure and covariance prior within procedure in Tables 1 and 2, and by procedure averaged across covariance priors in Tables 5 and 6, for data sets in Group 1 and Group 2, respectively.

Environmental covariances were estimated 20% more accurately than genetic covariances, and this occurred across procedures and groups of data sets. The absolute means of the relative environmental and genetic covariance estimates were 1.1 and 1.3, respectively. The poorest estimates were those for interbreed additive genetic effects, which, on average, were overestimated by 100% (absolute mean 2.0). These results were expected given the small number of bulls relative to the substantially larger number of calves in each sample.

Comparison of relative covariance values across groups of data sets indicates that covariances tended to be overestimated (absolute mean 1.2, Table 4) in data sets from Group 1 (two traits, direct genetic effects only) and underestimated (absolute mean .9, Table 4) in those from Group 2 (one trait, direct and maternal genetic effects). This occurred because most intra- and interbreed additive genetic as well as interbreed nonadditive genetic covariances between direct and maternal effects were underestimated by both the Partial Scoring and the Cholesky Maximization procedures (Table 2), particularly when low covariance values were used as priors.

Discussion

Computational Aspects

The Partial Scoring and Cholesky Maximization procedures managed to maintain positive definiteness of the estimated covariance matrices, and both procedures yielded similar overall means of relative covariance values and accuracies (Tables 3 and 4). However, their computational behavior with some data sets was substantially different, particularly in terms of the number of iterations required to achieve convergence. The Partial Scoring procedure tended to stay closer to the preliminary values used to start the algorithm, and convergence tended to proceed in short decreasing steps. The Cholesky Maximization procedure, on the other hand, tended to fluctuate more across EM iterations, and convergence tended to be achieved following a serrated path (not all EM iterations decreased the value of CCONV).

No relationship between speed of convergence and the reciprocal of the condition number MREMLEM equations, above the value of 10^{-5} allowed, was observed. In fact, the analyses that required the largest numbers of EM iterations (181 and 184) had a value of 10^{-2} for the reciprocal of the condition number.

The value of the accuracy of estimation ACSIM at convergence was largely determined during the first two to six EM iterations, regardless of the convergence value CCONV during these iterations. Thus, a smaller value of CCONV may be appropriate (e.g., 10^{-3} or even 10^{-2}), especially in data sets larger than the ones tested here.

The fact that the accuracy and the value of the covariance estimates at convergence were not exactly the same when different covariance priors were used (Tables 1, 2, 3, and 4) suggests that the set of covariance priors affected the level of multicollinearity of the MREMLEM equations and that this effect did not disappear with subsequent EM iterations. The method most affected was the Partial Scoring, particularly for covariance estimates from Group 1. On the other hand, results from individual runs using the Cholesky Maximization procedure were much more similar across covariance priors, especially for data sets from Group 1. From a practical standpoint, a possible alternative would be to make several runs per data set, each with a different set of covariance priors, and the final set of covariance estimates would be the set of average values across runs. Perhaps a minimum of three sets of starting covariances should be used. Given the tendency of the Partial Scoring procedure to follow the prior covariance values, the three sets of preliminary covariances for analyses of actual data sets should be close to the population values. A feasible set of covariance priors equivalent to the "equal" values utilized here would be the use of 1) estimates of intrabreed additive and environmental covariances from actual large population analyses as intrabreed additive and environmental priors, 2) the mean of the corresponding intrabreed covariances as interbreed additive and environmental covariance priors, and 3) covariances due to sire × breed group of dam interactions as interbreed nonadditive covariances at one locus, if available, else use values similar to those of the additive covariance priors. Next, two other vectors of covariance priors could be generated, one with values smaller than and another with values greater than the base vector. For the Cholesky Maximization procedure the choice of priors is less critical, although a good set of three vectors would probably give better estimates than a poor one.

Modeling and Programming Aspects

The main difference between the Partial Scoring and Cholesky Maximization procedures was the computation of the Cholesky elements of the intra- and interbreed genetic and environmental covariances at each EM iteration in the Cholesky Maximization procedure. This additional set of computations showed no increase in mean computing times per EM iteration (2.6 min for Partial Scoring and 2.4 min for Cholesky Maximization). This indicates that the partial steps used by the Partial Scoring procedure took a longer

time to control positive definiteness than the computation of the elements of the Cholesky Decomposition of the covariance matrices by the Cholesky Maximization procedure.

The expected log-likelihood and the actual log-likelihood of the complete data were computed at each Scoring and EM iteration for the Partial Scoring method and at each EM iteration for the Cholesky Maximization procedure. Both measures performed equally well to ensure that the log-likelihood of the complete data was increased at each step.

Covariances in data sets from Group 1 (two traits with direct genetic effects only) were estimated more accurately than those from Group 2 (a single trait with direct and maternal effects), even though data sets in Group 2 were constructed with a larger number of bulls than those in Group 1. This was expected because 1) a sire model, rather than a sire-maternal grandsire model, is actually used for the maternal component, and 2) a calf record supplies information for many more different covariances (direct and maternal) per trait than in analyses involving direct genetic effects only. For example, a record of an animal in Group 2 contributes to the estimation of 15 direct and maternal intrabreed and interbreed genetic and environmental covariances, as opposed to two records per animal jointly contributing to the estimation of 21 direct intra- and interbreed genetic and environmental covariances in Group 1.

The accuracies of estimation of the individual covariances obtained here were reasonable, given the small size of the simulated data sets used. Data sets containing tens of thousands of straightbred and crossbred bulls mated to dams of various breed groups would probably yield estimates of covariances of substantially higher accuracy. However, the computation of intrabreed and interbreed covariances in large unbalanced multibreed field data sets poses additional challenges not found in the small well-structured simulated data sets used here to test the Partial Scoring and Cholesky Maximization procedures. Some of these additional factors that must be considered are as follows: 1) multibreed data sets will probably contain more than two base breeds, thus increasing the number of covariances to be estimated, particularly the number of interbreed covariances; 2) a set of simplifying assumptions will be needed to help decide which intrabreed and(or) interbreed covariances will be considered to be equal, particularly in analyses of multibreed populations composed of more than two base breeds; 3) connectedness across multibreed contemporary groups will need to be checked, because data sets will be much less structured than the ones used here; 4) a much more complex model, particularly the part that accounts for group genetic effects and environmental fixed effects, will probably need to be used to describe a calf record; and 5) iterative procedures, rather than direct sparse procedures, will need to be used to solve for the MME.

Implications

The Partial Scoring and Cholesky Maximization procedures described here were successful in keeping estimates of intrabreed and interbreed variances and covariances in multibreed population within their permissible boundaries. However, because covariance estimates from the Cholesky Maximization procedure were much less affected by covariance priors than those from the Partial Scoring procedure, the Cholesky Maximization procedure should be preferred. Both procedures can be used in structured or in unstructured field data sets. Additional considerations (connectedness, simplifying assumptions, iterative strategies) will need to be taken into account when analyzing large unbalanced multibreed data sets.

Literature Cited

Arnold, J. W., J. K. Bertrand, and L. L. Benyshek. 1992. Animal model for genetic evaluation of multibreed data. J. Anim. Sci. 70:3322.

Bard, Y. 1974. Nonlinear Parameter Estimation. Academic Press, New York.

BIF. 1990. Guidelines for uniform beef improvement programs. Beef Improvement Federation, Oklahoma State Univ., Stillwater.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B 38:1.

Elzo, M. A. 1990a. Recursive procedures to compute the inverse of the multiple trait additive genetic covariance matrix in inbred and noninbred multibreed populations. J. Anim. Sci. 68:1215.

Elzo, M. A. 1990b. Covariances among sire by breed group of dam interaction effects in multibreed sire evaluation procedures. J. Anim. Sci. 68:4079.

Elzo, M. A. 1994. Restricted maximum likelihood procedures for the estimation of additive and nonadditive genetic variances and covariances in multibreed populations. J. Anim. Sci. 72:3055.

Elzo, M. A., and T. R. Famula. 1985. Multibreed sire evaluation procedures within a country. J. Anim. Sci. 60:942.

Golden, B. L., and R. M. Bourdon. 1994. Proposed interbreed data interchange format. Proc. Beef Improvement Fed. p 179.

Golub, G. H. and C. F. Van Loan. 1989. Matrix Computations (2nd Ed.). John Hopkins University Press, Baltimore, MD.

Groeneveld, E. 1994. A reparameterization to improve numerical optimization in multivariate REML (co)variance component estimation. Genet. Sel. Evol. 26:537.

Harville, D. A. 1977. Maximum likelihood approaches to variance component estimation and to related problems. J. Am. Stat. Assoc. 72:320.

Hoerl, E. A. and R. W. Kennard. 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12:55.

Jennrich, R. I., and M. D. Schluchter. 1986. Unbalanced repeated-measures models with structured covariance matrices. Biometrics 42:805.

Laird, N., N. Lange, and D. Stram. 1987. Maximum likelihood computations with repeated measures: Application of the EM algorithm. J. Am. Stat. Assoc. 82:97.

Lindstrom, M. J., and D. M. Bates. 1986. Newton-Raphson and EM algorithms for linear mixed models for repeated-measures data. J. Am. Stat. Assoc. 83:1014.

Meyer, K. 1994. Derivative-intense restricted maximum likelihood estimation of covariance components for animal models. Proc. 5th World Congr. Genet. Appl. Livest. Prod. 18:365, Univ. of Guelph, Guelph, Ontario, Canada.