

GENETICS AND BREEDING

Rapid Inversion of Dominance Relationship Matrices for Noninbred Populations by Including Sire by Dam Subclass Effects

I. HOESCHELE

Department of Dairy Science
Virginia Polytechnic Institute and State University
Blacksburg 24061-6999

P. M. VanRADEN

Animal Improvement Programs Laboratory
Agricultural Research Service
US Department of Agriculture
Beltsville, MD 20705-2350

ABSTRACT

For estimation of dominance effects and dominance variance, the inverse of a dominance relationship matrix is required. Dominance effects can be partitioned into sire \times dam or sire \times maternal grandsire subclass effects that are inherited and residuals within subclass that are not inherited. The subclass effects have immediate use in predicting performance of offspring from prospective matings. A rapid method for directly computing the inverse relationship matrix of subclass effects is presented. The procedure is similar to Henderson's simple method of computing an inverse additive genetic relationship matrix. The inverse relationship matrix among subclass effects consists of a contribution from each subclass of coefficients of a matrix of maximum size 9×9 . The algorithm can be modified to compute the inverse of the relationship matrix among sire \times dam or sire \times maternal grandsire subclasses and among individual dominance effects. Computing cost increases approximately linearly with dimensions of inverses. Dimensions could be several times the number of subclasses in the data because subclasses without records but providing relationship ties must be added. Computation of the inverse relationship matrix among 136,827 sire \times maternal grandsire subclasses in a population of 765,868

animals required 163 central processing unit seconds on an IBM 3090 and less than 4 Mbytes of memory.

(Key words: dominance relationships, matrix inversion, noninbred populations)

INTRODUCTION

Genetic evaluations have been largely restricted to additive genetic models in the form of sire or animal models. Nonadditive genetic variance contributes to the genetic covariance among relatives, which is well defined in randomly mating, noninbred populations with genetic effects composed of small contributions from many unlinked loci (3). Including nonadditive effects in genetic evaluation models might improve estimation of additive effects and also aid in planning matings to improve progeny performance.

Henderson (10) presented BLUP procedures (6) for mixed models including additive, dominance, additive by additive, and other genetic effects for randomly mating, noninbred populations. He proposed solving mixed model equations for total genetic merit (the sum of additive, dominance, additive by additive, and other genetic effects) to reduce number of simultaneous equations. Given estimates of total genetic merit, \hat{m} , estimates of components of m have to be computed; e.g., $d = \sigma_d^2 D M^{-1} \hat{m}$ for the vector of dominance effects d , where $\text{Var}(m) = M$ and $\text{Var}(d) = D \sigma_d^2$ with D , a matrix of dominance relationships, and σ_d^2 , dominance variance. For large data sets, however, no efficient algorithm exists for computing M^{-1} ; hence, mixed model equations should be formed for estimating components of m directly. The REML (13) estimation of variance components involves comput-

Received January 8, 1990.
Accepted September 19, 1990.

ing terms such as $\hat{d}'D^{-1}\hat{d}$, which also is required in approximations such as the tilde-hat approach (19).

To date, no methods were available to compute inverses of relationship matrices for nonadditive genetic effects as efficiently as the rapid method to compute A^{-1} , the inverse of the additive genetic relationship matrix (7, 8, 9). Rapid construction of an inverse is possible by first establishing recurrence or inheritance patterns among effects included. Dominance effects are partitioned into sire \times dam or sire \times maternal grandsire subclass effects that are inherited and uncorrelated residuals within subclass that are not inherited. An algorithm for rapidly computing the inverse relationship matrix among either subclass effects or among dominance and subclass effects in noninbred populations is presented.

MATERIALS AND METHODS

Components of Dominance Effect

Dominance effects result from interactions of pairs of genes at the same locus. An animal cannot transmit its dominance effect directly to its progeny because only one gene of each pair is transmitted to each progeny. Instead, inheritance of dominance effects can be traced through pairs of animals. For example, if progeny of a particular sire and dam have high average dominance effects, a mating of the dam to a close relative of the sire or a mating of the sire to a close relative of the dam would also be expected to yield progeny with high dominance effects.

Because dominance effects are not inherited through individuals but through pairs of animals, a partition of dominance effects into sire \times dam subclass effects and within subclass deviations is useful. Let an individual dominance effect d be partitioned as

$$d = f_{S,D} + \delta \quad [1]$$

where $f_{S,D}$ represents the average dominance effect of many hypothetical full sibs produced by the individual's sire S and dam D , and δ is the individual's deviation from the sire \times dam subclass effect. Deviation δ is due to Mendelian

sampling and has an expected value of zero. Thus, $E(d|f_{S,D}) = f_{S,D}$, and $f_{S,D}$ and δ are uncorrelated.

Let i be a hypothetical progeny of S and D and let j be a hypothetical progeny of K and L . With $d_i = f_{S,D} + \delta_i$ and $d_j = f_{K,L} + \delta_j$, covariance among dominance effects equals covariance among subclass effects or

$$\text{Cov}(d_i, d_j) = \text{Cov}(f_{S,D}, f_{K,L}) \quad [2]$$

because all covariances involving δ_i and δ_j are zero.

Covariances of dominance effects in noninbred populations can be computed from additive relationships (a_{XY}) among parents as $\text{Cov}(d_i, d_j) = (a_{SK}a_{DL} + a_{SL}a_{DK}) (.25)\sigma_d^2$. Variance of sire \times dam subclass effects σ_f^2 is equal to covariance among full sibs due to dominance, or $\sigma_f^2 = .25\sigma_d^2$. Hence,

$$\text{Cov}(f_{SD}, f_{KL}) = (a_{SK}a_{DL} + a_{SL}a_{DK})\sigma_f^2 \quad [3]$$

By subtraction, variance of Mendelian sampling δ in Equation [1] equals

$$\sigma_d^2 - \sigma_f^2 = .75\sigma_d^2$$

Recurrence Relationship

Inverses of variance matrices can be constructed easily if recurrence relationships exist among effects included (7, 14). Recurrence relationships for additive effects are simple because additive effects are averages of individual gene effects and each progeny received a sample half of each parent's genes. Thus,

$$a_i = .5 (a_S + a_D) + \phi_i \quad [4]$$

where a_i , a_S , and a_D are additive effects of animal i , its sire, and its dam, respectively, and ϕ represents Mendelian sampling.

Dominance effects are not inherited as simply as additive effects because they result from pairs of genes. Simple recurrence formulas for dominance effects can be developed by including effects of pairs of animals rather than individual animals. Parent subclass effects for sire

× dam subclass $f_{S,D}$ are interactions of S with parents of D, interactions of D with parents of S, and interactions of parents of S with parents of D. Figure 1 illustrates these parent subclasses with SS and DS denoting parents of S and SD and DD parents of D. Same sex subclasses such as S,SD never have records but do provide relationship ties in the same way that sires provide ties among cows with records for a sex-limited trait.

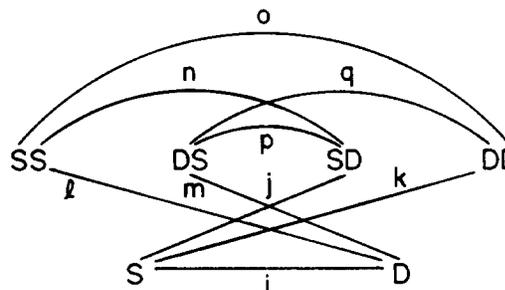


Figure 1. Parent subclasses of full sib family i.

Because dam D received half of her genes from SD and half from DD, half of the gene combinations contributing to $f_{S,SD}$ and $f_{S,DD}$ are expected to be identical to those in $f_{S,D}$. Hence, a simple recurrence to predict $f_{S,D}$ is $f_{S,D} = .5(f_{S,SD} + f_{S,DD}) + e_1$ where e_1 is a segregation residual. Another simple recurrence, using parents of the sire, is $f_{S,D} = .5(f_{SS,D} + f_{DS,D}) + e_2$ where e_2 is a different segregation residual. A new recurrence formula with smaller residual variance than that of e_1 or e_2 can be obtained by combining the two previous recurrences. Unfortunately, the subclasses involved have common contributions from their ancestor subclasses, for example, $f_{S,SD} = .5(f_{SS,SD} + f_{DS,SD}) + e_3$ and $f_{SS,DD} = .5(f_{SS,SD} + f_{SS,DD}) + e_4$ share $.5f_{SS,SD}$. Addition of the parent subclass contributions in the first two recurrences and subtraction of the redundant contributions of parent subclasses yields the final recurrence equation

$$f_{S,D} = .5(f_{S,SD} + f_{S,DD} + f_{SS,D} + f_{DS,D}) - .25(f_{SS,SD} + f_{SS,DD} + f_{DS,SD} + f_{DS,DD}) + e \quad [5]$$

where e is a segregation residual with smallest possible variance. Because $f_{S,D} - E(f_{S,D}|\text{parent subclass effects})$ equals e , parent subclass effects and residual e are uncorrelated.

Recurrence [5] can be derived more formally by regressing $f_{S,D}$ on its parent subclass effects:

$$f_{S,D} = b'f_{\text{par}} + e \quad [6]$$

where f_{par} is a vector of the eight parent subclass effects and b is a vector of corresponding partial regression coefficients with

$$b' = \text{Cov}(f_{S,D}, f_{\text{par}}) [\text{Var}(f_{\text{par}})]^{-1} \quad [7]$$

and

$$\text{Var}(e) = \sigma_f^2 - b' \text{Var}(f_{\text{par}}) b. \quad [8]$$

Use of Equations [6], [7], and [8] requires obtaining variances of and covariances among $f_{S,D}$ and f_{par} . Variances equal σ_f^2 and covariance between any two subclass effects $f_{S,D}$ and $f_{K,L}$ equals $(a_{SKADL} + a_{SLADK}) \sigma_f^2$, as given in Equation [3].

Let the nine subclasses in the order as they appear in Equation [5] be identified by (i, j, k, l, m, n, o, p, q). Relationship matrix among parent subclass effects is $\text{Var}(f_{\text{par}})/\sigma_f^2$, or

| | | | | | | | | | | |
|---|---|-----|-----|-----|-----|----|----|----|----|---|
| | | j | k | l | m | n | o | p | q | |
| j | [| 1 | .0 | .25 | .25 | .5 | .0 | .5 | .0 |] |
| k | | .0 | 1 | .25 | .25 | .0 | .5 | .0 | .5 | |
| l | | .25 | .25 | 1 | .0 | .5 | .5 | .0 | .0 | |
| m | | .25 | .25 | .0 | 1 | .0 | .0 | .5 | .5 | |
| n | | .5 | .0 | .5 | .0 | 1 | .0 | .0 | .0 | |
| o | | .0 | .5 | .5 | .0 | .0 | 1 | .0 | .0 | |
| p | | .5 | .0 | .0 | .5 | .0 | .0 | 1 | .0 | |
| q | | .0 | .5 | .0 | .5 | .0 | .0 | .0 | 1 | |

[9]

Vector of relationships between $f_{S,D}$ and its parent subclass effects is $\text{Cov}(f_{S,D}, f_{\text{par}})/\sigma_f^2$, or

$$i \begin{bmatrix} \text{S,SD} & \text{S,DD} & \text{SS,D} & \text{DS,D} & \text{SS,SD} & \text{SS,DD} & \text{DS,SD} & \text{DS,DD} \\ j & k & l & m & n & o & p & q \\ \text{.5} & \text{.5} & \text{.5} & \text{.5} & \text{.25} & \text{.25} & \text{.25} & \text{.25} \end{bmatrix} \quad [10]$$

Postmultiplication of [10] with the inverse of $\text{Var}(f_{\text{par}})/\sigma_f^2$ as in [7] yields regression coefficients

$$b' = [.5 \ .5 \ .5 \ .5 \ -.25 \ -.25 \ -.25 \ -.25]$$

which are identical to the coefficients in Equation [5]. Also, from Equation [8],

$$\text{var}(e) = .25\sigma_f^2.$$

If subclasses involving more remote ancestors of S and D are added, partial regression coefficients always equal zero, which can be verified using Equation [7]. This is similar to the situation with additive effects in which partial regression coefficients on grandparents are zero if parents are included.

Inverse of Relationship Matrix of Subclass Effects

Let f represent a vector of sire \times dam subclass effects for a population. Define the relationship matrix for these subclass effects to be $F = \text{Var}(f)/\sigma_f^2$. Elements of F are relationships among subclass effects such as $f_{S,D}$ and $f_{K,L}$ with numerical values ($a_{SK}a_{DL} + a_{SL}a_{DK}$) as given in Equation [3].

Quaas (14) provided an alternative derivation of the rules for finding A^{-1} , which is valuable in understanding the structure of A and A^{-1} . A similar approach can be used to derive rules for computing F^{-1} .

A matrix representation of recurrences [5] and [6] is

$$f = Qf + \epsilon \quad [11]$$

in which row i of Q has nonzero elements only in columns pertaining to parent subclasses of subclass i . If all parent subclasses are known as in [5], nonzero entries are .5 for parent subclasses of the type $f_{S,SD}$ and $-.25$ for parent subclasses of the type $f_{SS,SD}$. More generally,

with any number of missing parent subclasses, row i of Q contains the elements of b from Equation [7] in the columns pertaining to identified parent subclasses of subclass i . Rearrangement of [11] yields

$$f = [I - Q]^{-1} \epsilon. \quad [12]$$

The variance-covariance matrix of f is

$$\text{Var}(f) = F\sigma_f^2 = [I - Q]^{-1} R[I - Q]^{-1} \sigma_f^2$$

where $R\sigma_f^2 = \text{Var}(\epsilon)$. Finally,

$$F^{-1} = [I - Q] R^{-1} [I - Q]. \quad [13]$$

Diagonal elements of $\text{Var}(\epsilon)$ can be calculated from Equation [8]. Offdiagonal elements of $\text{Var}(\epsilon)$ are zero if all ancestor subclasses providing relationship ties are included in f . One method to ensure that all such subclasses are present is for each filled subclass to include all combinations of the sire and its ancestors with the dam and its ancestors. Many of these ancestor subclasses do not actually provide ties and can be treated as missing. Parent subclasses are actually missing if some SS, DS, SD, or DD are unknown or are treated as missing when they provide no ties.

Two conditions are helpful in deciding which subclasses should remain known: 1) a subclass should remain in f if any of its parent subclasses remain in f , and 2) a subclass should remain in f if f contains two or more of its immediate progeny subclasses, i.e., progeny subclasses related to it by .5.

Use of these two conditions seems to guarantee that R will be diagonal. Other, less restrictive conditions can also produce diagonal R but are more difficult to verify. If the above conditions are met, matrix R can be written as

$$R = Q_b F_b Q_b' + .25I$$

where F_b is the relationship matrix among unknown parent subclasses (f_b), and Q_b contains

partial regression coefficients from recurrence [5] pertaining to unknown parent subclasses.

Unknown subclasses in f_b can have more than one known progeny subclass in f , but contributions from unknown parents to each of these are independent, causing $Q_b F_b Q_b'$ to be diagonal. A subclass in f_b may have a partial regression coefficient in Q_b of .5 for only one progeny subclass because of condition 2. Then, that particular row of Q_b will also contain regressions of -.25 on the immediate parent subclasses of the parent subclass, which are unknown because of condition 1. Several progeny subclasses in f may have regressions of -.25 on the same unknown subclass in f_b , but each such row of Q_b would contain a regression of .5 on a different intermediate unknown subclass because of condition 2. Thus, each row of Q_b constructs uncorrelated differences among effects of unknown parent subclasses, provided that membership in f and f_b follows the two conditions given.

Rules for Computing F^{-1}

The following procedure for computing F^{-1} is suggested:

1. Begin with a list of animals and their parents (sire S and dam D). Parents not in the list of animals and with only one progeny may be treated as unknown. Parents with more than one progeny should be added to the list and assigned parent values of unknown. Step 1 is identical to that for A^{-1} .

2. Create a list of all filled (S,D) subclasses. Ancestor subclasses that provide ties should be added to this list. First, all ancestor subclasses can be identified by listing subclasses for sire with parents of the dam and dam with parents

of the sire for each filled subclass and then repeating this process for the subclasses just added until no further ancestors are known. When listing ancestor subclasses, same sex subclasses of animal 1 with animal 2 and of animal 2 with animal 1 should be treated as identical. The list including all filled and all ancestor subclasses should be sorted so that each progeny subclass precedes its parent subclasses. Then, beginning with oldest ancestor subclasses, subclasses may be declared unknown if they are not filled, have no known parents, and do not tie two or more filled descendant subclasses.

Tie status of ancestor subclasses can be determined approximately from counts formed when ancestor subclasses originally are identified. Progeny subclass $f_{S,D}$ would contribute + 1 to parent subclasses of type $f_{S,SD}$ and $f_{SS,D}$ but -1 to parent subclasses of type $f_{SS,SD}$. Subtraction of 1 is necessary because $f_{S,SD}$ and $f_{SS,D}$ are both counted as progeny subclasses of $f_{SS,SD}$ but both may have originated from just one filled subclass $f_{S,D}$ (see Figure 1). Some ancestor subclasses eliminated for a count of less than two may be needed to preserve diagonal R. These can be declared known in a final pass of the subclass list by adding back subclasses of type $f_{SS,SD}$ if $f_{S,SD}$ and $f_{SS,D}$ are both known.

3. Proceed through the list of all subclasses and write certain coefficients to disk or tape for each. From Equation [13], the contribution of subclass i to F^{-1} is

$$r^{ii} c_i c_i' \tag{14}$$

where r^{ii} is diagonal element i of R^{-1} , and c_i is row i of $I - Q$. If all eight parent subclasses of subclass i are known, $r^{ii} = 4$, and the contribution of i to F^{-1} is

| | S,D | S,SD | S,DD | SS,D | DS,D | SS,SD | SS,DD | DS,SD | DS,DD |
|---|-----|-----------|------|------|------|-------|-------|-------|-------|
| | i | j | k | l | m | n | o | p | q |
| i | 4 | -2 | -2 | -2 | -2 | 1 | 1 | 1 | 1 |
| j | | 1 | 1 | 1 | 1 | -.5 | -.5 | -.5 | -.5 |
| k | | | 1 | 1 | 1 | -.5 | -.5 | -.5 | -.5 |
| l | | | | 1 | 1 | -.5 | -.5 | -.5 | -.5 |
| m | | | | | 1 | -.5 | -.5 | -.5 | -.5 |
| n | | | | | | .25 | .25 | .25 | .25 |
| o | | symmetric | | | | | .25 | .25 | .25 |
| p | | | | | | | | .25 | .25 |
| q | | | | | | | | | .25 |

Fewer coefficients are contributed to F^{-1} by subclass i if any of its parent subclasses are unknown. The vector c_i contains nonzero coefficients equal to 1 for subclass i and equal to $-b$ for known parent subclasses, with b computed from Equation [7]. Coefficients of b will always equal .5 and $-.25$ if unknown status was determined by the two conditions stated previously. Diagonal element i of R^{-1} is obtained as $r^{ii} = 1/\text{Var}(e)$ with $\text{Var}(e)$ computed as in Equation [8].

4. Sort the coefficients by columns in rows and sum those with identical row and column to obtain the inverted relationship matrix F^{-1} among sire \times dam subclasses.

In populations where sires have many but dams have few progeny, relationships through sires and maternal grandsires rather than sires and dams may be sufficient. If subclass is defined as a sire \times maternal grandsire instead of sire \times dam combination, the model for statistical analysis of Allaire and Henderson (1) including sire, maternal grandsire, and sire \times maternal grandsire interaction can be used, except that relationships among interactions can now be accounted for by an F^{-1} as in the Appendix.

Prediction of Dominance Effects

Inheritance of dominance effects is not from dominance effects of ancestors but rather from ancestor subclass effects. Thus, the algorithm presented computes inverse of the relationship matrix among only the subclass effects. The vector of dominance effects d can be predicted by solving for predictions of f and backsolving for predictions of Mendelian sampling or by computing inverse of the relationship matrix among both dominance and subclass effects. Let W be the incidence matrix relating d to f and let δ equal d minus Wf . A recurrence equation analogous to Equation [11] is

$$\begin{bmatrix} d \\ f \end{bmatrix} = \begin{bmatrix} 0 & W \\ 0 & Q \end{bmatrix} \begin{bmatrix} d \\ f \end{bmatrix} + \begin{bmatrix} \delta \\ \varepsilon \end{bmatrix} \quad [15]$$

The dominance relationship matrix including subclass effects is D^* with

$$\text{Var} \begin{bmatrix} d \\ f \end{bmatrix} = D^* \sigma_d^2$$

$$\text{Var} \begin{bmatrix} \delta \\ \varepsilon \end{bmatrix} = \begin{bmatrix} .75I & 0 \\ 0 & .25R \end{bmatrix} \sigma_d^2$$

The inverse can then be computed rapidly using

$$D^{*-1} = \begin{bmatrix} I & 0 \\ -W & I - Q \end{bmatrix} \begin{bmatrix} (4/3)I & 0 \\ 0 & 4R^{-1} \end{bmatrix} \begin{bmatrix} I & -W \\ 0 & I - Q \end{bmatrix}$$

Matrix D^{*-1} is similar to F^{-1} except that coefficients contributed by subclasses are multiplied by 4 and additional coefficients of $4/3$ on diagonals and $-4/3$ on offdiagonals link dominance effects to the appropriate subclass effect.

Inbred Populations

In all livestock populations, inbreeding exists at low to moderate levels (20). Algorithms for computing inverses of relationship matrices in noninbred populations can give incorrect results if applied to populations with some inbreeding. Inverses of relationship matrices for additive (14) and dominance genetic effects can generally be written as:

$$\text{Inverse} = (I - P')R^{-.5}R^{-.5}(I - P) \quad [16]$$

where R is a diagonal matrix of variances of Mendelian segregation residuals, and nonzero elements of P are partial regression coefficients in recurrence equations, such as [4], [11], and [15] relating genetic effects to their parents. When formed using rules for noninbred populations, inverses for inbred populations are at least positive definite because $R^{-.5}(I - Q)$ is nonsingular.

In the presence of dominance variation, mean performance often decreases with inbreeding. The change in mean is linear in the inbreeding coefficient if there is no epistasis of higher order than additive by additive variation (12). Genetic covariances among relatives in inbred populations include components due to additive (σ_a^2) and dominance (σ_d^2) variances in an infinite randomly mating reference popula-

tion, a covariance between breeding values and dominance effects of inbred animals, and three additional quadratic components in the absence of epistasis (2, 4). Thus, inverses of additive genetic and dominance relationship matrices can no longer be formed separately.

The dominance relationships among inbred individual i with parents S and D and individual j with parents K and L , d_{ij} , and of i with itself, d_{ii} , must be computed from probabilities of gene identity by descent (5, 12). Let \equiv denote identity by descent, i_S and i_D represent alleles in i inherited from parents S and D , respectively, and j_K and j_L be defined similarly. Then, the covariance between i and j due to dominance variance in an infinite randomly mating reference population is $d_{ij}\sigma_d^2$ where [e.g., (5, 12)]

$$d_{ij} = \text{Prob}(i_S \equiv j_K \neq i_D \equiv j_L) \\ + \text{Prob}(i_S \equiv j_L \neq i_D \equiv j_K).$$

Algorithms that fully account for effects of inbreeding on genetic covariances in populations with additive and dominance variation add considerable complexity and computing expenses (16, 17). Further, inbreeding levels in livestock populations are often lower than under random mating (20); therefore, data may not permit estimation of the additional components of covariance due to inbreeding with sufficient accuracy.

When additive genetic and dominance relationship inverses are formed using algorithms for noninbred populations, relationship coefficients for inbred individuals and their close relatives may be computed incorrectly, and additional covariance components due to inbreeding are ignored. When [16] represents the inverse of the additive genetic relationship matrix, diagonals of R are equal to $1 - .25(1 + F_S) - .25(1 + F_D)$ (15) where F_S and F_D are inbreeding coefficients of sire and dam, respectively. When inbreeding is ignored by assuming that all F_S and F_D are 0, too much variance is attributed to offspring of inbred parents. If all inbred individuals in a population do not have offspring, the algorithm for noninbred populations produces the correct additive inverse.

The algorithm for computing the inverse of the relationship matrix among sire \times dam subclass effects, F^{-1} , for noninbred populations does not assign the correct variance to inbred animals. Diagonals of D should equal 1 minus the animal's inbreeding coefficient (5), but values of 1 or greater are assigned to inbred animals by the algorithm presented. When sire and dam are inbred by any amount but the resulting progeny are noninbred, dominance covariance among progeny in the same subclass (variance of sire \times dam subclass effect) equals $(1 + F_S)(1 + F_D)\sigma_f^2$ (11) and dominance covariance among progeny in different subclasses (covariance among sire \times dam subclass effects) equals $(a_{SK}a_{DL} + a_{SL}a_{DK})\sigma_f^2$ for subclass effects $f_{S,D}$ and $f_{K,L}$. These covariances are assigned correctly by the algorithm presented but dominance variance assigned to progeny is too large because Mendelian sampling variance is less than the $.75\sigma_d^2$ assumed. Modifications of the algorithm to account for the effects of inbreeding may be very difficult.

RESULTS

Small Example

To illustrate the algorithm, F^{-1} was computed from 16 animals with sires and dams as listed in Table 1. Table 2 contains the list of sire \times dam subclasses. List of filled subclasses was created in pass 1 by processing the list of animals and their parents in Table 1 in reverse order so that subclasses of younger sires and dams precede those of older sires and dams. Passes 2 and 3 through this list identified all ancestor subclasses. Parent subclasses of the types $f_{S,SD}$ and $f_{SS,SD}$ received counts of 1 and -1, respectively, from each progeny subclass $f_{S,D}$. For example, subclass $f_{H,A}$ received 1 from each of $f_{H,G}$, $f_{H,F}$, and $f_{N,A}$, and -1 from $f_{N,F}$. Known status was determined by proceeding through the subclass list in Table 2 in reverse order for the ancestor (not filled) subclasses (from $f_{E,B}$ to $f_{N,A}$). Those with a count of 1 and no parent subclasses known were treated as unknown.

Then, F^{-1} computed by the algorithm presented is

$$\begin{array}{r}
 f_{S,D} = \\
 \text{No.} =
 \end{array}
 \begin{array}{cccccccc}
 f_{N,F} & f_{H,I} & f_{H,G} & f_{H,F} & f_{C,E} & f_{A,B} & f_{N,A} & f_{H,A} \\
 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8
 \end{array}
 \left[\begin{array}{cccccccc}
 1.778 & .0 & .0 & -.889 & .0 & .0 & -.889 & .44 \\
 .0 & 1.33 & .0 & -.667 & .0 & .0 & .0 & .0 \\
 .0 & .0 & 1.33 & .0 & .0 & .0 & .0 & -.667 \\
 -.889 & -.667 & .0 & 2.11 & .0 & .0 & .44 & -.889 \\
 .0 & .0 & .0 & .0 & 1. & .0 & .0 & .0 \\
 .0 & .0 & .0 & .0 & .0 & 1. & .0 & .0 \\
 -.889 & .0 & .0 & .44 & .0 & .0 & 1.778 & -.889 \\
 .44 & .0 & -.667 & -.889 & .0 & .0 & -.889 & 2.11
 \end{array} \right]$$

As an example, contribution of subclass effect $f_{N,F}$ (number 1) to the inverse given above using [14] is

$$\begin{array}{r}
 \text{No.} =
 \end{array}
 \begin{array}{cccc}
 1 & 4 & 7 & 8
 \end{array}
 \left[\begin{array}{cccc}
 1. & -.5 & -.5 & .25 \\
 -.5 & .25 & .25 & -.125 \\
 -.5 & .25 & .25 & -.125 \\
 .25 & -.125 & -.125 & .0625
 \end{array} \right] 1.778$$

where $c'_1 = [1 \ b'_1]$ and $b'_1 = [.5 \ .5 \ -.25]$, computed using Equation [7] and F_1 , the relationship matrix among known parent subclasses 4, 7, and 8

$$\begin{array}{r}
 \text{No.} =
 \end{array}
 \begin{array}{ccc}
 4 & 7 & 8
 \end{array}
 \left[\begin{array}{ccc}
 1. & .25 & .5 \\
 .25 & 1. & .5 \\
 .5 & .5 & 1.
 \end{array} \right], \quad r^{11} = \frac{1}{1 - b'_1 F_1 b_1} = 1.778$$

Relationship matrix F , obtained by inverting F^{-1} given previously, is

$$\begin{array}{r}
 f_{S,D} = \\
 \text{No.} =
 \end{array}
 \begin{array}{cccccccc}
 f_{N,F} & f_{H,I} & f_{H,G} & f_{H,F} & f_{C,E} & f_{A,B} & f_{N,A} & f_{H,A} \\
 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8
 \end{array}
 \left[\begin{array}{cccccccc}
 1. & .25 & .125 & .5 & .0 & .0 & .5 & .25 \\
 .25 & 1. & .125 & .5 & .0 & .0 & .125 & .25 \\
 .125 & .125 & 1. & .25 & .0 & .0 & .25 & .5 \\
 .5 & .5 & .25 & 1. & .0 & .0 & .25 & .5 \\
 .0 & .0 & .0 & .0 & 1. & .0 & .0 & .0 \\
 .0 & .0 & .0 & .0 & .0 & 1. & .0 & .0 \\
 .5 & .125 & .25 & .25 & .0 & .0 & 1. & .5 \\
 .25 & .25 & .5 & .5 & .0 & .0 & .5 & 1.
 \end{array} \right]$$

Elements in F can be verified using Equation [3]. For further verification, the dominance relationship matrix among all animals in Table 1 with sire and dam known can be obtained as $D = (.25)WFW' + I(.75)$.

Actual Population

A FORTRAN program was developed to implement the algorithm for large populations. It computes the coefficients of the inverse relationship matrix among either sire \times dam or sire \times maternal grandsire subclasses. The program was applied to a data set including 765,868 daughters and granddaughters of 1003 popular Holstein bulls (18). Using the sire \times maternal grandsire option, there were 100,917 filled subclasses, and 35,910 ancestor subclasses were added to provide relationship ties. A total of 5,027,884 nonzero coefficients (required to form F^{-1}) were computed. This is about equal to the number of nonzero coefficients needed to form the inverse of the additive genetic relationship matrix for this population. Total computing time was 163 CPU seconds on an IBM

TABLE 1. List of animals and parents (noninbred pedigree).¹

| Animal | Sire | Dam |
|--------|------|-----|
| A | ... | ... |
| B | ... | ... |
| C | ... | ... |
| E | ... | ... |
| F | A | B |
| G | A | ... |
| H | C | E |
| I | ... | F |
| J | H | F |
| M | H | G |
| N | H | ... |
| O | H | I |
| P | N | F |
| Q | H | G |
| R | H | I |
| T | N | F |

¹Individuals are identified by letters.

3090, and memory requirements were less than 4 Mbytes. Computing times for two subsets with 1000 and 100,000 animals were 3 and 26 CPU seconds, respectively, which indicates that

TABLE 2. List of filled sire \times dam subclasses and ancestor subclasses.

| Sire \times dam subclass ¹ | | Grandparents | | | | Pass subclass was added ² | Count from progeny subclasses | Status ³ | Consecutive number of known subclasses | Known parent subclasses |
|---|---|--------------|-----|-----|-----|--------------------------------------|-------------------------------|---------------------|--|-------------------------|
| S | D | SS | DS | SD | DD | | | | | |
| N | F | H | ... | A | B | 1 | | K | 1 | 4, 7, 8 |
| H | I | C | E | ... | F | 1 | | K | 2 | 4 |
| H | G | C | E | A | ... | 1 | | K | 3 | 8 |
| H | F | C | E | A | B | 1 | 1 + 1 = 2 | K | 4 | 8 |
| C | E | ... | ... | ... | ... | 1 | | K | 5 | |
| A | B | ... | ... | ... | ... | 1 | | K | 6 | |
| N | A | H | ... | ... | ... | 2 | 1 | K | 7 | 8 |
| N | B | H | ... | ... | ... | 2 | 1 | TU | | |
| C | I | ... | ... | ... | F | 2 | 1 | TU | | |
| E | I | ... | ... | ... | F | 2 | 1 | TU | | |
| H | A | C | E | ... | ... | 2 | 1 + 1 + 1 - 1 = 2 | K | 8 | |
| C | G | ... | ... | A | ... | 2 | 1 | TU | | |
| E | G | ... | ... | A | ... | 2 | 1 | TU | | |
| H | B | C | E | ... | ... | 2 | 1 + 1 - 1 = 1 | TU | | |
| C | F | ... | ... | A | B | 2 | 1 + 1 - 1 = 1 | TU | | |
| E | F | ... | ... | A | B | 2 | 1 + 1 - 1 = 1 | TU | | |
| | | | | | | | 1 + 1 + 1 - 1 - 1 = | | | |
| A | C | ... | ... | ... | ... | 3 | 1 | TU | | |
| | | | | | | | 1 + 1 + 1 - 1 - 1 = | | | |
| A | E | ... | ... | ... | ... | 3 | 1 | TU | | |
| C | B | ... | ... | ... | ... | 3 | 1 + 1 - 1 = 1 | TU | | |
| E | B | ... | ... | ... | ... | 3 | 1 + 1 - 1 = 1 | TU | | |

¹S = Sire, D = dam, SS = sire of sire, DS = dam of sire, SD = sire of dam, DD = dam of dam.

²Pass 1 corresponds to list of filled sire \times dam subclasses.

³K = known, TU = treated unknown.

processing time increases approximately linearly with amount of data.

Inbred Example

For the inbred pedigree shown in Table 3, the relationship matrix among filled sire \times dam subclasses obtained by computing F^{-1} with the algorithm presented and inverting it is

$$f_{S,D} = \begin{bmatrix} f_{A,B} & f_{C,E} & f_{A,F} & f_{C,G} & f_{I,J} & f_{H,M} \\ 1 & 0 & .5 & 0 & 0 & 0 \\ 0 & 1 & 0 & .5 & 0 & 0 \\ .5 & 0 & 1 & 0 & 0 & 0 \\ 0 & .5 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.56 & .56 \\ 0 & 0 & 0 & 0 & .56 & 1.56 \end{bmatrix}$$

Dominance covariance among progeny of I and J is $\text{Var}(f_{I,J})$ which equals $1.56 \sigma_f^2$ or $.39 \sigma_d^2$. Dominance covariance of a progeny of I and J with a progeny of H and M is covariance of $f_{I,J}$ with $f_{H,M}$, which equals $.56 \sigma_f^2$ or $.14 \sigma_d^2$. Both covariances are correct because offspring were out of inbred but unrelated parents. Dominance variance assigned to an offspring of I and J or H and M is $[(1.56)(.25) + (.75)] \sigma_d^2 = 1.14 \sigma_d^2$. Dominance variance of an offspring of A and F or C and G is $[(1.0)(.25) + (.75)] \sigma_d^2 = 1.0 \sigma_d^2$. The correct coefficients of dominance variance are 1.0 and $1.0 - .5a_{AF} = .75$, respectively.

TABLE 3. List of animals and parents (inbred pedigree with parent-offspring matings).¹

| Animal | Sire | Dam |
|--------|------|-----|
| A | ... | ... |
| B | ... | ... |
| C | ... | ... |
| E | ... | ... |
| F | A | B |
| G | C | E |
| H | A | F |
| I | A | F |
| J | C | G |
| M | C | G |
| N | I | J |
| O | H | M |

¹Individuals are identified by letters.

Hence, the algorithm for noninbred populations attributes too much dominance variance to offspring of inbred but unrelated parents and also to offspring of related parents.

CONCLUSIONS

Dominance effects are composed of sire \times dam subclass effects that follow simple rules of inheritance plus independent residuals within subclass that are not inherited. Inverses of relationship matrices among sire \times dam subclass effects can be computed rapidly with cost proportional to number subclasses. Each subclass contributes coefficients of a matrix no larger than 9×9 (pertaining to interactions of sire and its parents with dam and its parents) to a larger matrix. Dimensions of inverses created could be several times the number of filled sire \times dam subclasses, because ancestor subclasses that provide ties must be added to the list of subclasses analogously to including sires and dams of more than one progeny in the list of animals when forming the additive relationship inverse (9).

Current knowledge about the magnitude of dominance variation in yield and nonyield (e.g., fertility, health, survival, type) traits is extremely limited. Dominance variance can be estimated as $\hat{\sigma}_d^2 = 4\hat{\sigma}_f^2$ by using the rules for computing F^{-1} and REML or an approximation. If additive by additive variation exists, estimates of $\hat{\sigma}_f^2$ would be biased upward. The techniques presented should allow dominance effects for noninbred populations to be included inexpensively in genetic evaluation models. Mixed model equations using such inverses predict which sires combine best with which dams, and predictions of individual dominance effects also can be obtained easily.

With low levels of inbreeding, inclusion of a linear regression on individual inbreeding coefficients in the model for data analysis while ignoring inbreeding in forming inverses of relationship matrices may be sufficient. This should allow an assessment of the variability of specific combining abilities beyond the average effects of inbreeding.

ACKNOWLEDGMENTS

This research was supported in part by grants from Eastern Artificial Insemination

Cooperative and the National Association of Animal Breeders.

REFERENCES

- 1 Allaire, F. R., and C. R. Henderson. 1965. Specific combining abilities among dairy sires. *J. Dairy Sci.* 48: 1096.
- 2 Chevalet, C., and M. Gillois. 1977. Estimation of genotypic variance components with dominance in small consanguineous populations. *Proc. Int. Conf. Quantitative Genet.*, Iowa State Univ. Press, Ames.
- 3 Cockerham, C. C. 1954. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* 39:859.
- 4 Cockerham, C. C., and B. S. Weir. 1984. Covariances of relatives stemming from a population undergoing mixed self and random mating. *Biometrics* 40:157.
- 5 Harris, D. L. 1964. Genotypic covariances between inbred relatives. *Genetics* 50:1319.
- 6 Henderson, C. R. 1973. Sire evaluation and genetic trends. Page 10 *in Proc. Anim. Breeding Genet. Symp. in Honor of J. L. Lush.* Am. Soc. Anim. Sci. Am. Dairy Sci. Assoc., Champaign, IL.
- 7 Henderson, C. R. 1975. Rapid method for computing the inverse of a relationship matrix. *J. Dairy Sci.* 58: 1727.
- 8 Henderson, C. R. 1975. Inverse of a matrix of relationships due to sires and maternal grandsires. *J. Dairy Sci.* 58:1917.
- 9 Henderson, C. R. 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32:69.
- 10 Henderson, C. R. 1985. Best linear unbiased prediction of nonadditive genetic merits in noninbred populations. *J. Anim. Sci.* 60:111.
- 11 Jacquard, A. 1974. The genetic structure of populations. Springer-Verlag, Berlin-Heidelberg, Germany.
- 12 Kempthorne, O. 1951. An introduction to genetic statistics. John Wiley & Sons, New York, NY.
- 13 Patterson, H. D., and R. Thompson. 1971. Recovery of interblock information when block sizes are unequal. *Biometrika* 58:545.
- 14 Quaas, R. L. 1988. Additive genetic model with groups and relationships. *J. Dairy Sci.* 71:1338.
- 15 Quaas, R. L. and E. J. Pollak. 1980. Mixed model methodology for farm and ranch beef cattle testing programs. *J. Anim. Sci.* 51:1277.
- 16 Smith, S. P. 1984. Dominance relationship matrix and inverse for an inbred population. Mimeo, Dep. Dairy Sci., The Ohio State Univ., Columbus.
- 17 Smith, S. P., and A. Maki-Tanila. 1990. Genotypic covariance matrices and their inverses for models allowing dominance and inbreeding. *Génét. Sél. Evol.* 22:65.
- 18 VanRaden, P. M. 1989. Estimates of nonadditive genetic variation for milk and fat yields of Holsteins. *J. Dairy Sci.* 72(Suppl. 1):59.(Abstr.)
- 19 VanRaden, P. M., and Y. C. Jung. 1988. A general purpose approximation to restricted maximum likelihood: the tilde-hat approach. *J. Dairy Sci.* 71:187.
- 20 Young, C. W., R. R. Bonczek, and D. G. Johnson. 1988. Inbreeding of and relationship among registered Holsteins. *J. Dairy Sci.* 71:1659.

APPENDIX

Sire-Maternal Grandsire Subclass Effects

For the subclass of common sire and maternal grandsire (MGS), assume that animals are related to others only through sire and MGS. Then

$$d_i = f_{S,MGS} + \delta_i, \quad d_j = f_{K,MGK} + \delta_j$$

where S (K) denotes sire of i(j), and MGS(MGK) denotes maternal grandsire of i(j);

$$\begin{aligned} \text{Cov}(f_{S,MGS}, f_{K,MGK}) &= \text{Cov}(d_i, d_j) \\ &= .0625(a_{SK}a_{MGS,MGK} + a_{S,MGK}a_{K,MGS}) \sigma_d^2 \\ \text{Var}(f_{S,MGS}) &= .0625 \sigma_d^2 = \sigma_f^2 \\ \text{Var}(\delta) &= \sigma_d^2 - \sigma_f^2 = .9375\sigma_d^2 \\ f_{S,MGS} &= .5(f_{S,SMGS} + .5f_{S,MMGS} + f_{SS,MGS} + .5f_{MGSS,MGS}) - .25(f_{SS,SMGS} \\ &\quad + .5f_{SS,MMGS} + .5f_{MGSS,SMGS} + .25f_{MGSS,MMGS}) + e \end{aligned}$$

where SMGS is sire of MGS, MMGS is maternal grandsire of MGS, and MGSS is maternal grandsire of S.

Rules for computing F^{-1} are identical to those previously discussed except that sire \times dam subclass is replaced by sire \times MGS subclass, D by MGS, DS by MGSS, SD by SMGS, and DD by MMGS (see Figure A1).

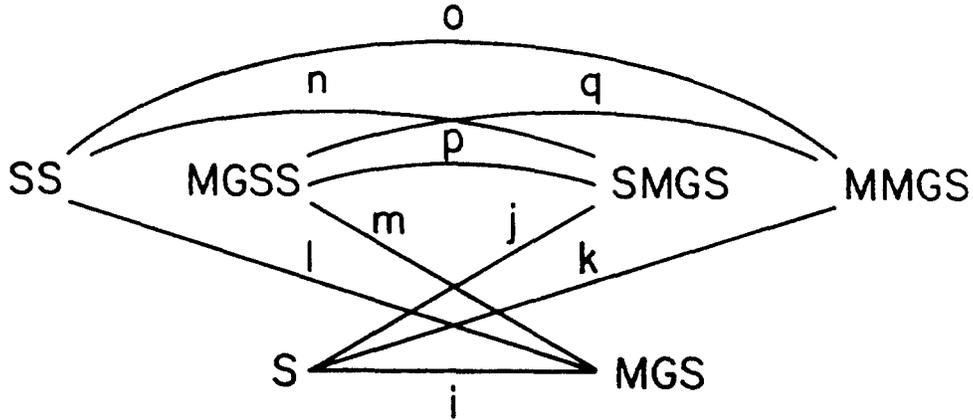


Figure A1. Parent subclasses of sire \times maternal grandsire subclass i.

Contribution of $f_{S,MGS}$ to F^{-1} is obtained using Equation [14] with

$$b' = \text{Cov}(f_{S,MGS}, f_{\text{par}}) [\text{Var}(f_{\text{par}})]^{-1},$$

and

$$\text{Var}(e) = \sigma_f^2 - b' \text{Var}(f_{\text{par}}) b.$$

Relationship matrix of the eight parent subclass effects, $\text{Var}(f_{\text{par}})/\sigma_f^2$, is

| | S, SMGS | S, MMGS | SS, MGS | MGSS, MGS | SS, SMGS | SS, MMGS | MGSS, SMGS | MGSS, MMGS |
|---|------------|------------|------------|--------------|-------------|-------------|---------------|---------------|
| | j | k | l | m | n | o | p | q |
| j | 1 | 0 | .25 | .125 | .5 | 0 | .25 | 0 |
| k | 0 | 1 | .125 | .0625 | 0 | .5 | 0 | .25 |
| l | .25 | .125 | 1 | 0 | .5 | .25 | 0 | 0 |
| m | .125 | .0625 | 0 | 1 | 0 | 0 | .5 | .25 |
| n | .5 | 0 | .5 | 0 | 1 | 0 | 0 | 0 |
| o | 0 | .5 | .25 | 0 | 0 | 1 | 0 | 0 |
| p | .25 | 0 | 0 | .5 | 0 | 0 | 1 | 0 |
| q | 0 | .25 | 0 | .25 | 0 | 0 | 0 | 1 |

The vector of relationships between sire \times MGS subclass i and its parent subclasses, $Cov(f_{s,mgs}, f_{par})/\sigma_f^2$, is

| | S, SMGS | S, MMGS | SS, MGS | MGSS, MGS | SS, SMGS | SS, MMGS | MGSS, SMGS | MGSS, MMGS |
|-------|------------|------------|------------|--------------|-------------|-------------|---------------|---------------|
| | j | k | l | m | n | o | p | q |
| S,MGS | .5 | .25 | .5 | .25 | .25 | .125 | .125 | .0625 |

Coefficients contributed to F^{-1} by sire \times MGS subclass i are then computed using [14].