Rapid Inversion of Additive by Additive Relationship Matrices by Including Sire-Dam Combination Effects

P. M. VanRADEN
Animal Improvement Programs Laboratory
Agricultural Research Service
United States Department of Agriculture
Beltsville, MD 20705-2350

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

ABSTRACT

Inverses of relationship matrices are useful for prediction of individual additive or nonadditive genetic merits and for estimation of variance components. An algorithm to form inverses of additive by additive relationship matrices rapidly from lists of individuals and their parents was developed. The algorithm uses simple recurrences among additive by additive and sire-dam combination effects to construct inverses for noninbred or inbred populations. Dimensions of matrices produced may be several times the number of individuals in the population because combination effects for sire-dam subclasses must be included in matrices. Rules of inheritance of sire-dam combination effects are the same as for dominance combination effects. Cost of forming inverses increases linearly with number of individuals. Each individual contributes 36 or fewer nonzero coefficients, and each sire-dam subclass contributes an additional 81 or fewer nonzero coefficients to the matrix. Computation of inverse of the relationship matrix due to 1003 sires and maternal grandsires of 765,868 cows required forming a matrix of order 137,830 and 4 Mbytes of memory.

(**Key words**: additive by additive relationships, recurrence relationships, matrix inversion)

Abbreviation key: A × A = additive by additive.

INTRODUCTION

Genetic variation is often partitioned into additive and nonadditive components. Nonadditive genetic variation results from interactions of genes. Interactions of genes at the same locus result in dominance variance and interactions of two genes at different loci result in additive by additive (A × A) variation. Interactions of more than two genes may also contribute nonadditive variance through terms such as additive by dominance, dominance by dominance, and additive by additive by additive variation.

Covariances between individuals due to any of these sources of genetic variation are well defined in randomly mating, noninbred populations if each source of variation consists of small contributions from many unlinked loci (3). More complex inheritance such as major genes and linkage also affect covariance among relatives but are beyond the scope of this paper and are not addressed further.

Genetic effects of individual members of a population can be predicted by use of mixed model equations (4, 8). These equations require inverses of relationship matrices among all additive and nonadditive effects in the model and knowledge of variance ratios. Unknown variances may be estimated by techniques such as REML (12), which also require inverses of relationship matrices.

Henderson (5, 6) presented rules for rapidly computing inverses of additive relationship matrices for noninbred populations. Hoeschele and VanRaden (9) present rules for rapidly comput-

---

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

ing inverses of dominance relationship matrices for noninbred populations. Chang et al. (2) presented methods to reduce times required to invert A × A relationship matrices. This paper presents an algorithm combining procedures of Hoeschele and VanRaden (9) and Chang et al. (2) that allows inverses of A × A relationship matrices to be computed rapidly for large populations.

MATERIALS AND METHODS

Recurrence Relationships for Noninbred Populations

Transmission of a sample half of genes from parents to offspring in each generation allows genetic effects to be described by simple recurrence relationships. These recurrence relationships allow rapid construction of relationship matrices and their inverses (7, 9). Inverses of additive and dominance relationship matrices can be formed rapidly from lists of animals and their parents (5, 7) and lists of sire × dam subclasses (9), respectively.

The A × A relationship matrix for any population can be formed rapidly by first forming the additive relationship matrix using a simple recursive procedure (7) and then squaring each element (3, 8, 10). Standard matrix inversion techniques could then be used to obtain the A × A inverse, but this approach could be extremely expensive for a large population. A more appealing strategy is to form the A × A inverse directly by using a simple recurrence relationship among A × A effects.

Let a be an animal's A × A effect. Then a can be described in terms of the animal's sire A × A effect (s), dam A × A effect (d), a combination effect of the sire and dam (s,d), and a Mendelian sampling effect (m):

$$a = .25(s) + .25(d) + (s,d) + m. \quad [1]$$

In the absence of inbreeding, the four terms on the right are mutually uncorrelated. Proof is that s represents interlocus interactions among the sire's genes, d represents interlocus interactions among the dam's genes, and (s,d) represents interactions of genes in the sire with genes at other loci in the dam. No covariance exists among these three terms unless the sire and dam share common genes.

Further, m represents the deviation of a from its full-sib family mean (f) defined as f = .25(s) + .25(d) + (s,d). Therefore, E(a|f) = f, and f and m are uncorrelated. The advantage of recurrence [1] over that used by Chang et al. (2) is that sire-dam combination effects, which have fairly simple inheritance (9), are now separated from Mendelian sampling effects, which are uncorrelated.

Full sibs have an additive relationship of .5 and an A × A relationship of $(.5)^2 = .25$. Let σ^2_{aa} represent variance of A × A effects. Then

$$\begin{aligned}
\mathrm{Var}(a) &= \mathrm{Var}(s) = \mathrm{Var}(d) = \sigma^2_{aa}\\
\mathrm{Var}(f) &= .25\ \sigma^2_{aa}\\
\mathrm{Var}(s,d) &= \mathrm{Var}(f) - (.25)^2\mathrm{Var}(s)\\
&\quad - (.25)^2\mathrm{Var}(d) = .125\ \sigma^2_{aa}\\
\mathrm{Var}(m) &= \mathrm{Var}(a) - \mathrm{Var}(f) = .75\ \sigma^2_{aa}
\end{aligned}$$

The covariance matrix for a and components of f is

$$\mathrm{Var}\begin{bmatrix} a\\ s\\ d\\ (s,d)\end{bmatrix} = \begin{bmatrix} 1 & .25 & .25 & .125\\ .25 & 1 & 0 & 0\\ .25 & 0 & 1 & 0\\ .125 & 0 & 0 & .125\end{bmatrix}\sigma^2_{aa}$$

In Equation [1], an individual's A × A effect is expressed in terms of its sire and dam A × A effects and a combination effect of sire and dam. A similar recurrence equation to express combination effects in terms of ancestor combination effects is needed. Let ss and ds refer to sire and dam of the sire; let sd and dd refer to sire and dam of the dam. Then

$$(s,d) = \begin{cases} .5(s,sd) + .5(s,dd) + e_1\\ .5(ss,d) + .5(ds,d) + e_2\end{cases}$$

These expressions both state that the combination effect of sire and dam equals the average of the combinations of one with the parents of the other plus Mendelian sampling terms. Combinations (s,sd) and (ds,d) are of animals of the same sex, and such combinations may not be possible in mammals (11). Nevertheless, including same-sex combination effects in the matrix is the easiest way to tie together more distantly related combinations of opposite sex that may occur.

The e_1 and e_2 do *not* have smallest variance possible. An e with smallest possible variance is obtained by combining the two (s,d) expressions as shown by Hoeschele and VanRaden (9):

$$\begin{aligned}(s,d) = {} & .5[(s,sd) + (s,dd) + (ss,d) \\ & + (ds,d)] - .25[(ss,sd) + (ss,dd) \\ & + (ds,sd) + (ds,dd)] + e \qquad [2]\end{aligned}$$

This recurrence equation for A × A combination effects is identical to that for dominance combination effects (9). Sire × dam subclass effects for dominance and A × A combination effects are both interactions of genes in the sire with genes in the dam and are inherited in the same way. For example, correlation of combination effects (s,d) and (s,o), where o is an offspring of d from a sire unrelated to s or d, is .5 whether (s,d) and (s,o) represent dominance or A × A combination effects. This is because d and o are expected to share half the same genes whether one is measuring within-locus or between-locus interactions of those genes with s.

Computation of Additive by Additive Inverse for Noninbred Populations

Inverses of relationship matrices among A × A effects can be constructed by combining algebra of Quaas (13) with recurrence equations [1] and [2]. Each A × A effect and each sire-dam combination effect is described in terms of ancestor effects. Some animals and subclasses must have unknown ancestors because pedigrees must end eventually. Let **u** contain known A × A effects ($\mathbf{u}_a$) and known sire-dam combination effects ($\mathbf{u}_c$) and let $\mathbf{u}_b$ contain effects of unknown ancestors and ancestor combinations.
Then

$$\begin{bmatrix} \mathbf{u}_a \\ \mathbf{u}_c \end{bmatrix} = \mathbf{u} = \begin{bmatrix} \mathbf{S} & \mathbf{W} \\ \mathbf{O} & \mathbf{Q} \end{bmatrix} \mathbf{u} + \begin{bmatrix} \mathbf{S}_b & \mathbf{W}_b \\ \mathbf{O} & \mathbf{Q}_b \end{bmatrix} \mathbf{u}_b + \begin{bmatrix} \mathbf{m}_a \\ \mathbf{m}_c \end{bmatrix} \qquad [3]$$

where $\mathbf{m}_a$ and $\mathbf{m}_c$ are vectors of Mendelian sampling effects associated with A × A and combination effects, respectively, and matrices **S**, **W**, **Q**, $\mathbf{S}_b$, $\mathbf{W}_b$, and $\mathbf{Q}_b$ link elements of **u** to ancestor effects.

Nonzero coefficients of **S** and $\mathbf{S}_b$ are values of .25 linking individuals in $\mathbf{u}_a$ to known parents in $\mathbf{u}_a$ and unknown parents in $\mathbf{u}_b$. Nonzero coefficients of **Q** and $\mathbf{Q}_b$ are values of .5 and −.25 linking sire-dam combinations to known or unknown ancestor combination effects in **u** and $\mathbf{u}_b$ as in Equation [2]. Nonzero coefficients of **W** and $\mathbf{W}_b$ are values of 1 linking individuals in $\mathbf{u}_a$ to sire-dam combinations in $\mathbf{u}_c$ as described by Hoeschele and VanRaden (9). Sire-dam combinations containing only one observation that are related to no other combinations may be declared unknown and included in $\mathbf{u}_b$ rather than **u**.

Redefinition of [3] gives

$$\mathbf{u} = \mathbf{P}\mathbf{u} + \mathbf{P}_b\mathbf{u}_b + \mathbf{m} \qquad [4]$$

where

$$\mathbf{P} = \begin{bmatrix} \mathbf{S} & \mathbf{W} \\ \mathbf{O} & \mathbf{Q} \end{bmatrix}, \ \mathbf{P}_b = \begin{bmatrix} \mathbf{S}_b & \mathbf{W}_b \\ \mathbf{O} & \mathbf{Q}_b \end{bmatrix}, \ \mathbf{m} = \begin{bmatrix} \mathbf{m}_a \\ \mathbf{m}_c \end{bmatrix}$$

Algebraic rearrangement of [4] gives

$$\mathbf{u} = (\mathbf{I} - \mathbf{P})^{-1}(\mathbf{P}_b\mathbf{u}_b + \mathbf{m}) \qquad [5]$$

Mean and variance of **u** are given by

$$\begin{aligned} E(\mathbf{u}) &= (\mathbf{I} - \mathbf{P})^{-1}\mathbf{P}_b E(\mathbf{u}_b), \text{ and} \\ Var(\mathbf{u}) &= (\mathbf{I} - \mathbf{P})^{-1}[\mathbf{P}_b Var(\mathbf{u}_b)\mathbf{P}_b' \\ &\quad + Var(\mathbf{m})]\ (\mathbf{I} - \mathbf{P}')^{-1} \end{aligned}$$

Var (**m**) is a diagonal matrix because Mendelian sampling effects are uncorrelated. Diagonals are .75 σ^2_{aa} corresponding to elements of $\mathbf{m}_a$ or .03125 σ^2_{aa} for elements of $\mathbf{m}_c$. If the conditions stated by Hoeschele and VanRaden (9) for determining unknown status of ancestor combinations are followed, $\mathbf{P}_b$ Var($\mathbf{u}_b$)$\mathbf{P}_b'$ is diagonal even though Var($\mathbf{u}_b$) is not. Diagonals of P_b Var(u_b)P_b' are sums of variance contributions from unknown ancestors and unknown ancestor combinations with an adjustment for the fact that ancestor combinations may be correlated.

Usually E(u_b) will be assumed null. If distinctly different base populations such as different breeds exist or if selection on A × A effects has caused differences in base members across time, procedures analogous to those in (13, 15) might be developed to account for these differences.

Let **R** represent the diagonal matrix [P_bVar(u_b)P_b' + Var(m)] ($1/\sigma^2_{aa}$), and let U represent Var(u) ($1/\sigma^2_{aa}$).

Then

$$\begin{aligned} U &= (I - P)^{-1}\, R(I - P')^{-1} \\ U^{-1} &= (I - P')R^{-1}\,(I - P) \qquad [6] \end{aligned}$$

A convenient algorithm for rapid construction of inverses of relationship matrices among A × A effects is provided by Equation [6].

Dimensions of U and U^{-1} can be several times the number of individuals in the population because ancestor combination effects are included. This increased size of U and U^{-1} is more than offset by the resulting sparseness and short time necessary to form U^{-1}, particularly for large populations. Cost of computing U^{-1} increases only linearly with number of individuals vs. cubic or possibly quadratic increases with algorithms involving matrix or sparse matrix inversion techniques.

Algorithm for Computing the Additive by Additive Inverse for Noninbred Populations

The following procedure for obtaining the A × A inverse is recommended.

1. Begin with a list of individuals and their parents. Parents not in the list of individuals and with only one progeny may be treated as unknown. Parents with more than one progeny should be added to the list of individuals and assigned parent values of unknown. Step 1 is identical to that for the additive inverse.

2. Create a list of filled sire-dam subclasses. Add to this list any ancestor subclasses that provide relationship ties. Filled subclasses that contain only one observation and are tied to no relative subclasses may be treated as unknown. Steps 1 and 2 are the same as for the dominance inverse (9).

3. For each individual in the expanded list created in step 1, write to disk or tape certain coefficients pertaining to the individual and its sire, dam, and sire-dam subclass effect. Number of coefficients written will vary from 1 to 16, and coefficients will differ depending on unknown status of sire, dam, and subclass. Four situations may be encountered:

If all sources are known, add

$$\begin{array}{c|cccc|} & a & s & d & (s,d) \\ a & 16 & -4 & -4 & -16 \\ s & -4 & 1 & 1 & 4 \\ d & -4 & 1 & 1 & 4 \\ (s,d) & -16 & 4 & 4 & 16 \end{array}\;(1/12)$$

If sire and dam are known but subclass effect is treated as unknown, add

$$\begin{array}{c|ccc|} & a & s & d \\ a & 16 & -4 & -4 \\ s & -4 & 1 & 1 \\ d & -4 & 1 & 1 \end{array}\;(1/14)$$

If one parent is known and subclass effect is unknown, add

$$\begin{array}{c|cc|} & a & \text{parent} \\ a & 16 & -4 \\ \text{parent} & -4 & 1 \end{array}\;(1/15)$$

If all sources are unknown, add 1 to the individual's diagonal.

4. For each sire-dam subclass in the expanded list of subclasses created in step (2), write to disk or tape certain coefficients pertaining to that sire-dam subclass and its ancestor subclasses. Number of coefficients will vary from 1 to 81, and coefficients will differ depending on which ancestor subclasses are unknown. Coefficients are the same as for the dominance inverse (9) except multiplied by eight.

5. Sort coefficients by row and by column within row, and sum coefficients with identical row and column to obtain U^{-1}.

For some populations, U^{-1} may be formed from a list of sires and maternal grandsires instead of sires and dams. Required changes in the algorithm are in the Appendix.

Computation of Additive by Additive Inverse for Inbred Populations

In a pedigree contains inbreeding, contributions of sire (s), dam (d), and sire-dam subclass (s,d) in recurrence [1] are correlated. With the most extreme inbreeding, selfing, sire equals dam and these two effects are perfectly correlated. With less extreme inbreeding such as from mating paternal halfsibs, a parent interaction of subclass (s,d) is the interaction of sire of sire (= sire of dam) with himself (ss,ss). Because all inbreeding results from an ancestor appearing on both sides of an animal's pedigree, the list of all ancestor interactions for an inbred animal always includes interaction of the common ancestor with itself.

To establish the relationship between an individual's interaction with itself and its A × A effect, suppose an individual with A × A effect i is selfed to produce two offspring with A × A effects j and k. With F_i denoting inbreeding coefficient of i,

$$\mathrm{Var}\begin{bmatrix} i \\ j \\ k \end{bmatrix} = \begin{bmatrix} (1 + F_i)^2 & (1 + F_i)^2 & (1 + F_i)^2 \\ (1 + F_i)^2 & (1.5 + .5F_i)^2 & (1 + F_i)^2 \\ (1 + F_i)^2 & (1 + F_i)^2 & (1.5 + .5F_i)^2 \end{bmatrix} \sigma^2_{aa}.$$

Elements of this matrix are of course squares of the additive relationship elements. Variance of the A × A effect common to j and k (f) equals covariance of j and k or

$$\mathrm{Var}(f) = \mathrm{Cov}(j,k) = (1 + F_i)^2\sigma^2_{aa}.$$

Let m_j be Mendelian sampling in j such that $j = f + m_j$. Because m_j is independent of i,

$$\mathrm{Cov}(i,f) = \mathrm{Cov}(i,j) = (1 + F_i)^2\sigma^2_{aa}.$$

Correlation of i and f is given by

$$\mathrm{Corr}(i,f) = \mathrm{Cov}(i,f)/\sqrt{\mathrm{Var}(i)\mathrm{Var}(f)} = 1.$$

Because i and f are perfectly correlated, have equal variances, and also have equal expectations, i = f. From Equation [1], f = .25(i) + .25(i) + (i,i). Substitution of i for f gives (i,i) = .5(i). Consequently, an individual's interaction with itself equals half its A × A effect.

The recurrence equation for an animal's interaction with itself (a,a) can be obtained by replacing .5a, .5s, and .5d with (a,a), (s,s), and (d,d), respectively, in Equation [1]:

$$\begin{aligned}(a,a) &= .25(s,s) + .25(d,d) \\ &\quad + .5(s,d) + .5m \qquad [7]\end{aligned}$$

From the inheritance of combination effects (9),

$$\begin{aligned}(a,a) &= .5(a,s) + .5(a,d) + e_1 \qquad [8] \\ &= (a,s) + (a,d) - .5(a,s) - .5(a,d) \\ &\quad + e_1, \\ (a,s) &= .5(s,s) + .5(s,d) + e_2 \qquad [9] \\ (a,d) &= .5(s,d) + .5(d,d) + e_3 \qquad [10]\end{aligned}$$

Substituting [9] and [10] for the terms subtracted off in Equation [8] produces another recurrence equation for (a,a) that has smallest possible residual variance:

$$\begin{aligned}(a,a) &= (a,s) + (a,d) - .25(s,s) \\ &\quad - .25(d,d) - .5(s,d) + e \qquad [11]\end{aligned}$$

Interestingly, Equation [11] can also be obtained by applying the general recurrence [2] to interaction effect (a,a).

For situations in which some parent effects in [2] and [11] may be unknown, recurrence equations for combination effects of the form (a,a) and (s,d) can generally be written (9) as

$$(s,d) = \mathbf{b}'\mathbf{c} + e \qquad [12]$$

where **c** is a vector of known parent interactions and **b** is a vector of partial regression coefficients with

$$\mathbf{b} = \mathrm{Cov}[(s,d), \text{known parent interactions}]\,[\mathrm{Var}\,(\text{known parent interactions})]^{-1} \quad [13]$$

and

$$\mathrm{Var}\,(e) = \mathrm{Var}(s,d) - \mathbf{b}'[\mathrm{Var}(\text{known parent interactions})]\mathbf{b}. \quad [14]$$

Computing the covariance matrix of known parent interaction effects requires a general expression for covariances among interactions. Sire-dam interactions are A × A effects of animals minus direct contributions of sire and dam and minus Mendelian sampling, or $(s,d) = a - .25\ (s + d) - m$, where a and m belong to a randomly chosen progeny of s and d. Let a′ be the A × A effect of another randomly chosen member of a different full-sib family with sire s′ and dam d′. Then

$$\mathrm{Cov}[(s,d),(s',d')] = \mathrm{Cov}[(a - .25s - .25d - m), (a' - .25s' - .25d' - m')].$$

Because m is uncorrelated with a′, s′, d′ and m′ and m′ is uncorrelated with a,s,d and m

$$\mathrm{Cov}[(s,d),(s',d')] = \mathrm{Cov}[(a - .25s - .25d), (a' - .25s' - .25d')].$$

Terms on the right side of the last equation are each functions of additive relationships among animals and parents. Let A_{ij} denote additive relationship of individuals with A × A effects i and j, respectively. Relationships involving animals can be expressed in terms of A_{ij} coefficients of parents, e.g., $A_{as'} = .5(A_{ss'} + A_{ds'})$ and $A_{aa'} = .25(A_{ss'} + A_{ds'} + A_{sd'} + A_{dd'})$. Writing covariance of (s,d) with (s′,d′) in terms of A_{ij} coefficients of parents and further algebra yields

$$\mathrm{Cov}[(s,d),(s',d')] = .125\ (A_{ss'}A_{dd'} + A_{sd'}A_{ds'}).$$

Covariance of subclasses (s,d) and (s′,d′) simplifies to variance of (s,d) when s = s′ and d = d′, yielding

$$\mathrm{Var}(s,d) = .125(A_{ss}A_{dd} + A^2_{sd}).$$

These results allow computing Var(known parent interactions) and evaluating [12] for any interaction (s,d) or (a,a). Proof of recurrence [11] can be obtained by applying Equation [12] to (a,a) with parent combinations (a,s), (a,d), (s,s), (d,d), and (s,d) known.

Algorithm for Inbred Populations

An algorithm for computing A × A inverses for inbred populations includes all of the same steps as the algorithm for noninbred populations. Interactions of animals with themselves must be excluded from the list of interactions created in step 2. Inclusion of an animal's A × A effect and its interaction with itself would result in a singular matrix because a = 2(a,a). Interactions of an animal with its sire and dam must be included if these provide ties between more than 1 descendant interaction.

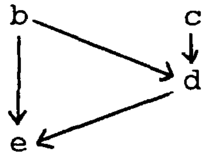
Coefficients written in steps 3 and 4 are computed according to Equation [6]. Diagonal elements of **R** are computed using [14], and nonzero coefficients of **P** are regression coefficients computed using [13]. Known parent interactions may include some subclasses of type (a,a). Because the vector **u** includes effects 2(a,a) = a rather than (a,a), covariances and variances of (a,a) ancestors must be multiplied by 2 and 4, respectively, when using Equation [13].

Individuals with inbred offspring may contribute a maximum of 36 rather than 16 coefficients because the list of known ancestor effects for such individuals may now include interactions of the animal with its sire and dam. Maximum coefficients contributed by any (s,d) subclass is still 81.

RESULTS

Small Example

Consider the following pedigree with individuals identified by letters:



Because e is inbred, the algorithm for inbred populations is used to form U^{-1}. The vector of

A × A and combination effects is $\mathbf{u}' = [b,c,d,e,(b,d),(b,c)]$. Effects b, c, and (b,c) have no known ancestors. Ancestor interactions (c,d), (b,e), and (d,e) have only one descendant each and, hence, were not included in **u**. Contributions to U^{-1} for each effect in **u** were computed from Equation [6] using [13] and [14]. In order, contributions were

for b add $\begin{array}{c} \text{b} \\ [1.0] \end{array}$,

for c add $\begin{array}{c} \text{c} \\ [1.0] \end{array}$,

for d add

$$\begin{array}{cccc} \text{b} & \text{c} & \text{d} & \text{(b,d)} \\ \left[\begin{array}{cccc} .125 & -.125 & .5 & -1 \\ -.125 & .125 & -.5 & 1 \\ .5 & -.5 & 2 & -4 \\ -1 & 1 & -4 & 8 \end{array}\right] \end{array},$$

for e add

$$\begin{array}{cccc} \text{b} & \text{d} & \text{e} & \text{(b,d)} \\ \left[\begin{array}{cccc} .0625 & .0625 & -.25 & .25 \\ .0625 & .0625 & -.25 & .25 \\ -.25 & -.25 & 1 & -1 \\ .25 & .25 & -1 & 1 \end{array}\right] \end{array},$$

for (b,d) add

$$\begin{array}{ccc} \text{b} & \text{(b,d)} & \text{(b,c)} \\ \left[\begin{array}{ccc} 1 & -4 & 2 \\ -4 & 16 & -8 \\ 2 & -8 & 4 \end{array}\right] \end{array},$$

and for (b,c) add $\begin{array}{c} \text{(b,c)} \\ [8.0] \end{array}$.

Summation of contributions from all six effects gives

$$U^{-1} = \left[\begin{array}{cccccc} \text{b} & \text{c} & \text{d} & \text{e} & \text{(b,d)} & \text{(b,c)} \\ 2.1875 & -.125 & .5625 & -.25 & -4.75 & 2 \\ & 1.125 & -.5 & 0 & 1 & 0 \\ & & 2.0625 & -.25 & -3.75 & 0 \\ & & & 1 & -1 & 0 \\ & & & & 25 & -8 \\ \text{symmetric} & & & & & 12 \end{array}\right]$$

Inversion of U^{-1} gives

$$U = \left[\begin{array}{cccccc} \text{b} & \text{c} & \text{d} & \text{e} & \text{(b,d)} & \text{(b,c)} \\ 1 & 0 & .25 & .5625 & .25 & 0 \\ & 1 & .25 & .0625 & 0 & 0 \\ & & 1 & .5625 & .25 & .125 \\ & & & 1.5625 & .28125 & .09375 \\ & & & & .15625 & .0625 \\ \text{symmetric} & & & & & .125 \end{array}\right]$$

The upper left 4 × 4 submatrix is the A × A relationship matrix and the remaining elements are covariances for sire-dam interactions.

Actual Population

A FORTRAN program was developed to implement the algorithm for large populations. This program computes coefficients of the inverse of the A × A relationship matrix either due to sires and dams or due to sires and maternal grandsires. The program was applied to a population of 765,868 Holstein cows. Number of sires plus maternal grandsires was only 1003 because of edits used for variance component estimation (14). Using the sire-maternal grandsire option and the algorithm for noninbred populations, order of the matrix constructed was 137,830. Number of nonzero coefficients required to form the inverse was 5,042,444. Among rows and columns of the inverse created, 136,827 pertained to sire-maternal grandsire combination effects. Total computing time was 165 CPU seconds on an IBM 3090, and memory requirements were 4 Mbytes.

Due to use of sire-maternal grandsire option, ratio of order of the A × A inverse to order of the additive inverse (1003) was approximately 100 to 1. With the sire-dam option, this ratio would reduce to approximately 4 or 5 to 1.

CONCLUSIONS

In noninbred populations, A × A effects are composed of independent contributions of sire, dam, sire-dam combination, and Mendelian sampling. Inheritance of A × A sire-dam combination effects is identical to inheritance of sire × dam subclass effects for dominance (9). Inverses of A × A relationship matrices for noninbred populations can be formed rapidly from lists of animals and their parents and lists of filled subclasses and parent subclasses.

An animal's A × A effect is twice the interaction (subclass) effect of the animal with itself. This identity allows correlations among sire, dam, and sire-dam subclass contributions to an animal's A × A effect to be accounted for in inbred populations. Inverses of A × A relationship matrices for inbred populations can be formed rapidly from lists of subclasses including interactions of animals with themselves and their parents.

Procedures to form inverses of A × A relationship matrices include many of the same steps used for dominance inverses (9). Inclusion of both inverses in mixed model equations should allow separation of dominance and A × A variation if both exist.

Inverses produced are very sparse but can have dimensions several times the number of animals because of addition of ancestor sire-dam subclass effects. Mixed model equations incorporating these inverses predict which sires combine best with which dams in addition to predictions of individual A × A effects. Time required to construct inverses is approximately proportional to number of animals. Results can be verified by comparing the inverse of the A × A inverse to the A × A relationship matrix.

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 Chang, H. L., R. L. Fernando, and D. Gianola. 1989. Inverse of an additive × additive relationship matrix due to sires and maternal grandsires. J. Dairy Sci. 72: 3023.

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 Henderson, C. R. 1973. Sire evaluation and genetic trends. Proc. Anim. Breeding Genet. Symp. in Honor of J. L. Lush. Am. Soc. Anim. Sci. Am. Dairy Sci. Assoc., Champaign, IL.

5 Henderson, C. R. 1975. Rapid method for computing the inverse of a relationship matrix. J. Dairy Sci. 58: 1727.

6 Henderson, C. R. 1975. Inverse of a matrix of relationships due to sires and maternal grandsires. J. Dairy Sci. 58:1917.

7 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.

8 Henderson, C. R. 1985. Best linear unbiased prediction of nonadditive genetic merits in noninbred populations. J. Anim. Sci. 60:111.

9 Hoeschele, I., and P. M. VanRaden. 1991. Rapid inversion of dominance relationship matrices for noninbred populations by including sire by dam subclass effects. J. Dairy Sci. 74:557.

10 Kempthorne, O. 1955. The theoretical values of correlations between relatives in random mating populations. Genetics 40:153.

11 McGrath, J., and D. Solter. 1984. Completion of mouse embryo genesis requires both the maternal and paternal genomes. Cell 37:179.

12 Patterson, H. D., and R. Thompson. 1971. Recovery of interblock information when block sizes are unequal. Biometrika 58:545.

13 Quaas, R. L. 1988. Additive genetic model with groups and relationships. J. Dairy Sci. 71:1338.

14 VanRaden, P. M. 1989. Estimates of nonadditive genetic variation for milk and fat yields of Holsteins. J. Dairy Sci. 72(Suppl. 1):59. (Abstr.)

15 Westell, R. A., R. L. Quaas, and L. D. VanVleck. 1988. Genetic groups in an animal model. J. Dairy Sci. 68:2267.

APPENDIX

Relationships due to common sires and maternal grandsires rather than sires and dams may sometimes be sufficient for studying nonadditive variation in dairy populations (1). Let a be an animal's A × A effect; s, its sire A × A effect; mgs, it maternal grandsire A × A effect; and (s,mgs), the combination effect of sire and maternal grandsire.

Recurrence Relationships for Noninbred Populations

$$\begin{aligned}
a &= .25(s) + .0625\,(mgs) + (s,mgs) + m \\
\mathrm{Var}(a) &= \mathrm{Var}(s) = \mathrm{Var}(mgs) = \sigma^2_{aa} \\
f &= .25(s) + .0625(mgs) + (s,mgs) \\
\mathrm{Var}(f) &= (.3125)^2\sigma^2_{aa} \\
\mathrm{Var}(s,mgs) &= .03125\sigma^2_{aa} \\
\mathrm{Var}(m) &= [1 - (.3125)^2]\sigma^2_{aa} = (231/256)\sigma^2_{aa} \qquad [A1]
\end{aligned}$$

Further,

$$\begin{aligned}
(s,mgs) = {} & .5[(s,smgs) + .5(s,mmgs) \\
& + (ss,mgs) + .5(mgss,mgs)] \\
& - .25[(ss,smgs) + .5(ss,mmgs) \\
& + .5(mgss,\ smgs) \\
& + .25(mgss,mmgs)] + e \qquad [A2]
\end{aligned}$$

where smgs is sire of mgs, mmgs is maternal grandsire of mgs, and mgss is maternal grandsire of s. In [3] through [6], nonzero coefficients of S and S_b are now .25 pertaining to sire and .0625 pertaining to maternal grandsire. Nonzero coefficients of Q and Q_b are .5, .25, −.25, −.125, and −.0625 as follows from [A2]. Var(m) is a diagonal matrix with diagonals equal to (231/256) σ^2_{aa} corresponding to elements of m_a or (121/256) (1/32) σ^2_{aa} for elements of m_c.

Algorithm for Noninbred Populations

In the algorithm for computing U^{-1}, work with a list of individuals and their sires and maternal grandsires and with a list of sire-maternal grandsire subclasses. Contributions of an individual to the matrix are as follows.

1. If all sources are known, add

$$\begin{array}{r} \\ a \\ s \\ mgs \\ (s,mgs) \end{array}
\begin{array}{c} \begin{array}{cccc} a & s & mgs & (s,mgs) \end{array} \\ \left[\begin{array}{cccc} 256 & -64 & -16 & 256 \\ -64 & 16 & 4 & -64 \\ -16 & 4 & 1 & -16 \\ 256 & -64 & -16 & 256 \end{array}\right] \end{array} (1/231)$$

2. If sire and maternal grandsire are known but subclass effect is unknown, add

$$\begin{array}{r} \\ a \\ s \\ mgs \end{array}
\begin{array}{c} \begin{array}{ccc} a & s & mgs \end{array} \\ \left[\begin{array}{ccc} 256 & -64 & -16 \\ -64 & 16 & 4 \\ -16 & 4 & 1 \end{array}\right] \end{array} (1/239)$$

3. If sire is known but maternal grandsire and subclass effect are unknown, add

$$\begin{array}{r} \\ a \\ s \end{array}
\begin{array}{c} \begin{array}{cc} a & s \end{array} \\ \left[\begin{array}{cc} 256 & -64 \\ -64 & 16 \end{array}\right] \end{array} (1/240)$$

4. If maternal grandsire is known, but sire and subclass effect are unknown, add

$$\begin{array}{r} \\ a \\ mgs \end{array}
\begin{array}{c} \begin{array}{cc} a & mgs \end{array} \\ \left[\begin{array}{cc} 256 & -16 \\ -16 & 1 \end{array}\right] \end{array} (1/255)$$

5. If all sources are unknown, add 1 to the individual's diagonal.

Recurrence Relationships for Inbred Populations

With sire-maternal grandsire relationships, inbreeding occurs only when an animal's sire and mgs are related. An extreme case is when sire equals maternal grandsire. Then, algebra similar to that used for selfing gives the identity (a,a) = .25(a). This identity for sire-maternal grandsire combination effects differs from that for sire-dam combination effects because one of the animals involved now contributes 1/4 rather than 1/2 of its genes.

Dividing [A1] by 4 and replacing .25(a), .25(s) and .25(mgs) with (a,a), (s,s), and (mgs,mgs), respectively, yields

$$\begin{aligned}(a,a) &= .25(s,s) + .0625(mgs,mgs)\\ &\quad + .25(s,mgs) + .25m \qquad [A3]\end{aligned}$$

Other useful recurrences are

$$\begin{aligned}(a,a) &= .5(a,s) + .25(a,mgs) + e_1\\ &= (a,s) + .5(a,mgs) - .5(a,s)\\ &\quad - .25(a,mgs) + e_1 \qquad [A4]\end{aligned}$$

$$(a,s) = .5(s,s) + .25(s,mgs) + e_2 \quad [A5]$$

$$(a,mgs) = .5(s,mgs) + .25(mgs,mgs) + e_3 \quad [A6]$$

Substituting [A5] and [A6] for the terms subtracted off in recurrence [A4] yields a recurrence with smallest possible residual variance:

$$\begin{aligned}(a,a) &= (a,s) + .5(a,mgs) - .25(s,s)\\ &\quad - .0625(mgs,mgs)\\ &\quad - .25(s,mgs) + e \qquad [A7]\end{aligned}$$

Coefficients contributed by each subclass to the inverse are calculated using Equation [6] with [12], [13], and [14]. Covariance matrix of known parent interactions of sire-maternal grandsire subclasses is needed in [13] and [14]. This covariance matrix can be computed from the following general expression for covariance of (s,mgs) and (s′,mgs′) subclasses, derived analogously to that of (s,d) and (s′,d′) subclasses, $\mathrm{Cov}[(s,mgs),(s',mgs')] = .03125(A_{s,s'}A_{mgs,mgs'} + A_{s,mgs'}A_{mgs,s'})$.