# Accounting for Mutation Effects in the Additive Genetic Variance-Covariance Matrix and Its Inverse

Naomi R. Wray

Stable URL:

*Biometrics* is currently published by International Biometric Society.

# SHORTER COMMUNICATIONS

## EDITOR:
## NIELS KEIDING

# Accounting for Mutation Effects in the Additive Genetic Variance–Covariance Matrix and Its Inverse

## Naomi R. Wray*

Institute of Animal Genetics, University of Edinburgh, King's Buildings,
West Mains Road, Edinburgh EH9 3JN, Scotland

## SUMMARY

Procedures for calculating the additive genetic variance–covariance matrix and its inverse are adapted to accommodate the occurrence of mutations in the genome. The inverse matrix can be used in mixed model methodology for best linear unbiased prediction of breeding values or for variance component estimation under a genetic model that includes mutation effects.

## 1. Introduction

In the mixed model methodology (Henderson, 1973) used in animal breeding to predict breeding values of animals, the additive genetic variance–covariance matrix (or, in fact, its inverse) is included with the aim of increasing the accuracy of prediction by accounting for all the additive genetic variances and covariances between individuals. In this way, the reduction in additive genetic variance ($\sigma_a^2$) due to inbreeding and to selection is intrinsically accounted for (Sorenson and Kennedy, 1984). However, these decreases in additive genetic variance may be counterbalanced to some extent by the occurrence of mutations of the genome. Genomic mutations include base pair substitution, duplications, insertions, and inversions of segments of chromosomes. They may occur at neutral sites (e.g., at introns or inactive sites), at active sites but at loci that are neutral with respect to the quantitative trait of interest, or at loci that affect the trait. In the latter case the new mutations would serve to increase $\sigma_a^2$. Mutation has been offered as one explanation for the continued response in long-term selection experiments in which the attainment of a selection plateau might be expected, for example, more than 50 generations of selection for oil in maize (Dudley, 1977). Hill (1982) examined the effects of mutation on response to selection by simulation and concluded that although mutation events had little impact on the initial generations of selection programmes, they should not be ignored in the design and analysis of long-term experiments, nor in the analysis of data on many generations from small animal populations, e.g., pigs and poultry. Inclusion of the effects of mutation into an additive genetic mixed model therefore seems important for the analysis of some animal breeding data (Dempfle, 1987; Dempfle and Grundl, 1988) and requires that account be taken of variances and covariances between animals accrued by mutations in previous generations, yet not traceable to the base population.

---

*Currrent address*: The Edinburgh School of Agriculture, West Mains Road, Edinburgh EH9 3JG, Scotland.

*Key words:* Inverse of relationship matrix; Mutation variance; Numerator relationship matrix.

Few estimates of mutation variance exist; in a review of literature, Lynch (1988) found that estimates of mutation variance in laboratory species ranged from $10^{-4}$ to $5 \times 10^{-2}$ of the error variance. No estimates are available for traits of commercial importance in livestock. Incorporation of mutation effects into the additive genetic variance–covariance matrix used in a variance component estimation procedure may then provide useful estimates of the mutation variance.

When mutation effects are not included, $\mathbf{A}$ is the numerator relationship matrix, where $\mathbf{A} = \{a_{ij}\}$ such that $a_{ij}$ is the numerator of Wright's (1922) coefficient of relationship. Procedures exist for the construction of $\mathbf{A}$ (Emik and Terrill, 1949) and for construction of its inverse, $\mathbf{A}^{-1}$, without direct inversion (Henderson, 1976; Quaas, 1976; Thompson, 1977) from a list of animals and their sires and dams with the animals ordered by date of birth. The inverse additive genetic variance–covariance matrix used in mixed model methodology is $\mathbf{A}^{-1}\sigma_{\mathrm{a}}^{-2}$.

In this note these procedures are adapted to account for the random effects of mutation and their inheritance in the additive genetic variance–covariance matrix and its inverse.

## 2. The Additive Genetic Model Including Mutation Effects

The trait of interest is assumed to be under the control of many unlinked loci, each of small additive effect, the infinitesimal model. Individuals in the base population are assumed to have no mutation effects since any mutations in their genome contribute to the base additive genetic variance. New mutations are assumed to arise independently in individuals of the first and subsequent generations, and they are assumed to have a small additive effect [as justified by the simulations of Keightley and Hill (1988)] with mean zero and contributing a new variance of $\sigma_{\mathrm{m}}^2$ per individual. Mutations are inherited by descendants as infinitesimal additive genes.

Let the additive genetic value or breeding value of animal $i$ be denoted $u_i$. Then

$$u_i = .5u_{\mathrm{s}} + .5u_{\mathrm{d}} + \phi_i + \gamma_i, \tag{1}$$

where $u_{\mathrm{s}}$ and $u_{\mathrm{d}}$ are the breeding values of the sire and dam of $i$; $\phi_i$ is the random Mendelian or within-family sampling effect unique to $i$, which is distributed with mean zero and var($\phi_i$); and $\gamma_i$ is the random mutation effect unique to $i$, which is distributed with mean zero and variance $\sigma_{\mathrm{m}}^2$. The variables $\phi_i$ and $\gamma_i$ are independent of each other and of $u_{\mathrm{s}}$ and $u_{\mathrm{d}}$.

Let the additive genetic variance–covariance matrix including mutation effects be $\mathbf{M}\sigma_{\mathrm{a}}^2$. $\mathbf{M}\sigma_{\mathrm{a}}^2$, up to time unit $t$, can be partitioned as

$$\mathbf{M}\sigma_{\mathrm{a}}^2 = \mathbf{A}_0\sigma_{\mathrm{a}}^2 + \sum_{k=1}^{t} \mathbf{A}_k\sigma_{\mathrm{m}}^2 = \left[\mathbf{A}_0 + \theta \sum_{k=1}^{t} \mathbf{A}_k\right]\sigma_{\mathrm{a}}^2, \tag{2}$$

where $\theta = \sigma_{\mathrm{m}}^2/\sigma_{\mathrm{a}}^2$. $\mathbf{A}_0\sigma_{\mathrm{a}}^2$ is the additive genetic variance–covariance matrix of additive effects attributed to the distribution of variance present in the base population. $\mathbf{A}_k\sigma_{\mathrm{m}}^2$ is the variance–covariance matrix of additive effects attributed to mutations arising in time unit $k$. For simplicity, time units hereafter will be called generations, which is appropriate for populations with discrete generations. For the situation of overlapping generations, time units can be considered as, for example, birth seasons. In fact, the only restriction on the allocation of time units is that parents must feature in an earlier time unit than their offspring. The elements of $\mathbf{A}_k$ are the additive genetic relationships if ancestors born in generations 0 to $k - 1$ are ignored. Elements of $\mathbf{A}_k$, like those of $\mathbf{A}_0$, have a maximum value of 2, which represents complete identity by descent. Diagonal elements of $\mathbf{A}_k$ (for $k \geq 0$) are $1 + f_{i,k}$, where $f_{i,k}$ is the inbreeding coefficient of individual $i$ ignoring common ancestors born in generations 0 to $k - 1$. The $f_{i,k}$ represent the proportional reduction of the segregation variance contributed by individual $i$ to its progeny and which

can be traced to effects arising in generation $k$ (or, if $k = 0$, effects which were present in the base population). Elements of $\mathbf{M}$ are similar to those of $\mathbf{A}_k$ in that they are the sums of relationships over generations weighted by $\theta$, but as such, they have no limiting value and therefore $\mathbf{M}\sigma_a^2$ is most simply interpreted as a variance–covariance matrix. In $\mathbf{M}$, terms denoted $h_i$ are defined which are similar to $f_{i,0}$ in $\mathbf{A}_0$ and to $f_{i,k}$ in $\mathbf{A}_k$, for $k > 0$, in that it represents the amount by which segregation variance is reduced, but $h_i$ may be greater than 1 since it may contain a term in $\theta$. An important assumption implicit in the infinitesimal model is that there is an infinitesimally small probability of a new mutation destroying the homology at an inbred locus. However, new mutations themselves may be subject to inbreeding.

For an individual in the base population, generation 0, $\text{var}(u_i) = \sigma_a^2$. For an individual in generation 1, $\text{var}(u_i) = \sigma_a^2 + \sigma_m^2$, because inbreeding is not possible in the first generation of a population in which selfing does not occur. In generation 2 and thereafter, the presence of mutation effects in the parent generations and the inbreeding of the offspring and also parents (in generation 3 and after) must be considered:

$$\text{var}(u_i) = .25\ \text{var}(u_s) + .25\ \text{var}(u_d) + .5\ \text{cov}(u_s, u_d) + \text{var}(\phi_i) + \sigma_m^2. \tag{3}$$

The first three terms of equation (3) comprise the between-family variance, which can be written as

$$(.25m_{ss} + .25m_{dd} + .5m_{sd})\sigma_a^2 = (.25m_{ss} + .25m_{dd} + h_i)\sigma_a^2, \tag{4}$$

where $m_{ij}$ are the elements of $\mathbf{M}$ and $h_i = .5m_{sd}$.

$\text{Var}(\phi_i)$ is the variation within full-sib families caused by Mendelian sampling at segregating loci and can be separated into the sampling variance of male parent gametes and female parent gametes. If mutation effects are ignored, then the within-sire family variation is $.25\sigma_a^2$ in the base population. In later generations, this is reduced by a factor $f_s$ [$f_s$ is the inbreeding coefficient of the sire and $f_s\sigma_a^2 = .5\ \text{cov}(u_{ss}, u_{sd})$, where $u_{ss}$ and $u_{sd}$ are the breeding values of the paternal grandsire and grandam of individual $i$]. $f_s$ accounts for the proportion of loci which are identical by descent and therefore contribute no segregation sampling variance. In this circumstance,

$$\text{var}(\phi_i) = [.25(1 - f_s) + .25(1 - f_d)]\sigma_a^2,$$

which can be written as

$$\text{var}(\phi_i) = [.25(a_{ss} + a_{dd}) - .5f_s - .5f_d]\sigma_a^2. \tag{5}$$

When mutation effects are considered, the variability attributed to a non-inbred parent is greater than $.25\sigma_a^2$ because of the mutations which serve to increase the variability. Inbreeding of the parents will reduce the segregation variance in the same way as before, with $h_s\sigma_a^2 = .5\ \text{cov}(u_{ss}, u_{sd})$ representing the reduction of variance. The within-family variance analogous to equation (5) is

$$\text{var}(\phi_i) = (.25m_{ss} + .25m_{dd} - .5h_s - .5h_d)\sigma_a^2. \tag{6}$$

In follows that, in total, $\text{var}(u_i)$ is

$$\text{var}(u_i) = [.5(m_{ss} + m_{dd} - h_s - h_d) + h_i + \theta]\sigma_a^2 = m_{ii}\sigma_a^2,$$

which reduces to the more familiar equality $\text{var}(u_i) = (1 + f_i)\sigma_a^2$ when mutation effects are not included.

## 3. Procedure for Constructing the Additive Genetic Variance–Covariance Matrix $\mathbf{M}\sigma_a^2$

$\mathbf{M}$ can be computed recursively from a list of individual, sire, and dam identifications ordered by the age of the individuals (hereafter called the ordered pedigree file). Any sire or dam identified must previously have appeared in the list as an individual. Unknown

parents are identified by 0. Any individual with both parents unknown is assumed to be sampled from the base population and so has no accumulated mutation effect. The vector **h** stores the $h_i$ coefficients.

(i) If both parents of individual $i$ are known, say $p$ and $q$,

$$m_{ji} = m_{ij} = .5(m_{jp} + m_{jq}) \qquad \text{for } j = 1, \ldots, i - 1$$

$$m_{ii} = .5(m_{pp} + m_{qq} + m_{pq} - h_p - h_q) + \theta$$

$$h_i = .5m_{pq}$$

(ii) If only one parent is known, say $p$,

$$m_{ji} = m_{ij} = .5m_{jp} \qquad \text{for } j = 1, \ldots, i - 1$$

$$m_{ii} = .5(1 + m_{pp} - h_p) + \theta$$

$$h_i = 0$$

(iii) If neither parent is known,

$$m_{ji} = m_{ij} = 0 \qquad \text{for } j = 1, \ldots, i - 1$$

$$m_{ii} = 1$$

$$h_i = 0$$

## 4. Procedure for Constructing the Inverse of the Additive Genetic Variance–Covariance Matrix, $\mathbf{M}^{-1}\sigma_a^{-2}$

To form the mixed model equations applied to animal breeding data, $\mathbf{A}^{-1}\sigma_a^{-2}$ is added to the animal × animal block of the least squares equations. Henderson (1976) showed that $\mathbf{A}^{-1}$ can be constructed directly without the costly procedure of forming $\mathbf{A}$ itself and inverting this sparse matrix. $\mathbf{A}$, being a symmetric matrix, can be written as a product of a lower triangular matrix, $\mathbf{L}$, and its transpose $\mathbf{L}'$: $\mathbf{A} = \mathbf{LL}'$. In addition, $\mathbf{L}$ has elements $l_{ij}$ and let $\mathbf{L} = \mathbf{TD}$, where $\mathbf{D}$ is a diagonal matrix with diagonal elements $(d_i)$ identical to the diagonal elements of $\mathbf{L}$, and $\mathbf{T}$ is a lower triangular matrix with diagonal elements equal to 1. Then

$$\mathbf{A}^{-1} = (\mathbf{LL}')^{-1} = (\mathbf{TD}^2\mathbf{T}')^{-1} = (\mathbf{T}^{-1})'(\mathbf{D}^{-1})^2\mathbf{T}^{-1}.$$

Henderson (1976) demonstrated that each row of $\mathbf{T}^{-1}$ has a maximum of two nonzero elements which have the value $-.5$ in the columns corresponding to the parents of the individuals represented by the row. Exploiting this structure, $\mathbf{A}^{-1}$ can be easily constructed from a list of individual ($i$), sire ($s$), and dam ($d$) identifications, ordered by the age of the individuals. For completeness those rules are simply to add:

$$b_i \quad \text{to } a^{ii}$$

$$-.5b_i \quad \text{to } a^{is}, a^{si}, a^{id}, a^{di} \qquad (7)$$

$$.25b_i \quad \text{to } a^{ss}, a^{sd}, a^{ds}, a^{dd}$$

where $a^{ij}$ is the $(i, j)$ element of $\mathbf{A}^{-1}$ and $b_i = l_{ii}^{-2} = d_i^{-2}$. The only requirement, therefore, is that the diagonal elements are known and these are $\text{var}(\phi_i)/\sigma_a^2$ (Thompson, 1977), where $\text{var}(\phi_i)$ was defined in equation (5). Each $d_i^2$ is simply $(\frac{1}{2} - f_s/4 - f_d/4)$, $\frac{3}{4} - f_s/4$ (or $\frac{3}{4} - f_d/4$), or 1, depending on whether 2, 1, or 0 parents, respectively, are identified, and these items are easily found if $\mathbf{A}$ is known. Alternatively, Henderson (1976) found the diagonal elements by construction and storage of $\mathbf{L}$, which could require substantial computer memory when the number of individuals is large. Quaas (1976) presented an

algorithm that does not require **L** or **A** to be stored and, therefore, can be used on very large data sets. These algorithms can be easily updated to account for mutations.

### 4.1 *Adapting the Method of Henderson* (1976) *to Account for Mutation Effects*

**M** is a symmetric matrix and can be written as a product of a lower triangular matrix **L** and its transpose, **L**′. Similarly, each of the $\mathbf{A}_k$ ($k \geq 0$) matrices in equation (2) can be written as a product of lower triangular matrices $\mathbf{L}_k$ and their transpose $\mathbf{L}'_k$:

$$\mathbf{M}\sigma_a^2 = \mathbf{LL}'\sigma_a^2 = \mathbf{L}_0\mathbf{L}'_0\sigma_a^2 + \sum_{k=0}^{t} \mathbf{L}_k\mathbf{L}'_k\sigma_m^2.$$

$\mathbf{L}_k$ for $k > 0$ has null columns corresponding to individuals in generations 0 to $k - 1$; the remaining columns are identical to $\mathbf{L}_0$. In addition, $\mathbf{L}_k = \mathbf{T}_0\mathbf{D}_k$ ($k \geq 0$), where, analogous to the presentation of Henderson (1976) above, $\mathbf{D}_k$ is a diagonal matrix with diagonal elements identical to those of $\mathbf{L}_k$ and $\mathbf{T}_0$ is a lower triangular matrix with diagonal elements 1. Let $\mathbf{E}_0 = \mathbf{D}_0$ and let $\mathbf{E}_k = \mathbf{D}_k\sqrt{\theta}$. Then

$$\mathbf{M}\sigma_a^2 = \mathbf{T}_0\mathbf{D}_0^2\mathbf{T}_0\sigma_a^2 + \sum_{k=1}^{t} \mathbf{T}_k\mathbf{D}_k^2\mathbf{T}'_k\sigma_a^2 = \mathbf{T}_0\left(\sum_{k=0}^{t} \mathbf{E}_k^2\right)\mathbf{T}'_0\sigma_a^2$$

and

$$\mathbf{M}^{-1}\sigma_a^{-2} = (\mathbf{T}_0^{-1})'\left(\sum_{k=0}^{t} \mathbf{E}_k^2\right)^{-1}\mathbf{T}_0^{-1}\sigma_a^{-2}.$$

Therefore, exactly the same rules may be used to construct $\mathbf{M}^{-1}$ as for $\mathbf{A}^{-1}$ above [equations (7)], with the $b_i$ being the $i$th diagonal element of $(\sum_{k=0}^{t} \mathbf{E}_k^2)$, which is identical to $l_{ii}^{-2}$, where $l_{ii}$ is the $i$th diagonal element of **L** from $\mathbf{M} = \mathbf{LL}'$. If an alternative but equivalent derivation, analogous to that of Thompson (1977), is followed, then the diagonal elements are easily observed to be $[\text{var}(\phi_i) + \sigma_m^2]/\sigma_a^2$.

Each row of **L** can be computed in order by the following method from the ordered pedigree list. Vector **h** contains the $h_i$ values.

(1) If both parents of individual $i$ are known, say $p$ and $q$ where $p < q$,

$$l_{ij} = .5(l_{pj} + l_{qj}) \quad \text{for } j = 1, \ldots, p$$

$$l_{ij} = .5l_{qj} \qquad\qquad \text{for } j = p + 1, \ldots, q$$

$$l_{ij} = 0 \qquad\qquad\quad \text{for } j = q + 1, \ldots, i - 1$$

$$l_{ii} = \sqrt{.25\left(\sum_{j=1}^{p} l_j^2 + \sum_{j=1}^{q} l_j^2\right) - .5(h_p + h_q) + \theta}$$

$$h_i = .5\sum_{j=1}^{p} l_{pj}l_{qj}$$

(2) If only one parent of individual $i$ is known, say $p$,

$$l_{ij} = .5l_{pj} \quad \text{for } j = 1, \ldots, p$$

$$l_{ij} = 0 \qquad \text{for } j = p + 1, \ldots, i - 1$$

$$l_{ii} = \sqrt{.5 + .25\left(\sum_{j=1}^{p} l_{pj}^2\right) - .5h_p + \theta}$$

$$h_i = 0$$

(3) If neither parent of individual $i$ is known,

$$l_{ij} = 0 \quad \text{for } j = 1, \ldots, i - 1$$

$$l_{ii} = 1$$

$$h_i = 0.$$

$\mathbf{M}^{-1}$ is constructed by the rules (7) above using $b_i = l_{ii}^{-2}$. If $\mathbf{M}$ is known, the $l_{ii}^2$ can be found directly without constructing $\mathbf{L}$ as

(1) If both parents are known, say $p$ and $q$,

$$l_{ii}^2 = .25(m_{pp} + m_{qq}) - .5(h_s + h_d) + \theta$$

(2) If only one parent is known, say $p$,

$$l_{ii}^2 = .5 + .25m_{pp} - .5h_s + \theta$$

(3) If neither parent is known,

$$l_{ii}^2 = 1$$

## 4.2 *Adapting the Method of Quaas* (1976) *to Account for Mutation Effects*

The method of Quaas (1976) is adapted to find the diagonal elements of $\mathbf{L}$, where $\mathbf{M} = \mathbf{LL}'$, without storing $\mathbf{L}$. Again the procedure requires the ordered pedigree list. For $n$ individuals, $n$ rounds of the procedure are performed. Three vectors, $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{h}$, are defined with $i$th elements $u_i$, $v_i$, and $h_i$, respectively. The sum of squares of the entries of each row of $\mathbf{L}$ are stored in $\mathbf{u}$. After the $k$th round of the procedure the first $k$ elements of $\mathbf{v}$ contain the diagonal elements of $\mathbf{L}$. Vector $\mathbf{h}$ stores the $h_i$ coefficients.

For the $k$th round:

(a)

$$v_k = l_{kk} = \begin{cases} \sqrt{.25(u_p + u_q) - .5(h_p + h_q) + \theta} & \text{if } 0 < p < q \\ \sqrt{.5 + .25u_q - .5h_q + \theta} & \text{if } p = 0 < q \\ 1 & \text{if } p = q = 0 \end{cases}$$

where $p$ and $q$ are the parents of the $k$th individual and an unknown parent is denoted 0.

(b) For $j = k + 1, \ldots, n$,

$$v_j = l_{jk} = \begin{cases} .5(v_{p_j} + v_{q_j}) & \text{if } k \leq p_j < q_j \\ .5v_{q_j} & \text{if } p_j < k \leq q_j \\ 0 & \text{if } p_j \leq q_j < k \end{cases}$$

where $p_j$ and $q_j$ are the parents of individual $j$, $p_j < q_j$.

(c) For $j = k + 1, \ldots, n$,

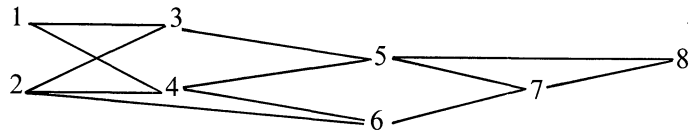$$h_j = \begin{cases} h_j + .5v_{p_j}v_{q_j} & \text{if } k \leq p_j < q_j \\ h_j & \text{if } p_j < k \end{cases}$$

(d) For $j = k, n$,

$$u_j = u_j + v_j^2$$

To construct $\mathbf{M}^{-1}$ by equations (7), use $b_i = l_{ii}^{-2} = v_i^2$.

## 5. Example

Consider the following pedigree:



For all methods the pedigree data entered are

$$
\begin{array}{ccccc}
 & & & & \text{Time unit} \\
1 & 0 & 0 & & 0 \\
2 & 0 & 0 & & 0 \\
3 & 1 & 2 & & 1 \\
4 & 1 & 2 & & 1 \\
5 & 4 & 3 & & 2 \\
6 & 4 & 2 & & 2 \\
7 & 6 & 5 & & 3 \\
8 & 7 & 5 & & 4 \\
\end{array}
$$

$$\mathbf{M} = \mathbf{A}_0 + \theta(\mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3 + \mathbf{A}_4);$$

$$
\mathbf{A}_0 =
\begin{bmatrix}
1 & 0 & 1/2 & 1/2 & 1/2 & 1/4 & 3/8 & 7/16 \\
0 & 1 & 1/2 & 1/2 & 1/2 & 3/4 & 5/8 & 9/16 \\
1/2 & 1/2 & 1 & 1/2 & 3/4 & 1/2 & 5/8 & 11/16 \\
1/2 & 1/2 & 1/2 & 1 & 3/4 & 3/4 & 3/4 & 3/4 \\
1/2 & 1/2 & 3/4 & 3/4 & 5/4 & 5/8 & 15/16 & 35/32 \\
1/4 & 3/4 & 1/2 & 3/4 & 5/8 & 5/4 & 15/16 & 25/32 \\
3/8 & 5/8 & 5/8 & 3/4 & 15/16 & 15/16 & 21/16 & 9/8 \\
7/16 & 9/16 & 11/16 & 3/4 & 35/32 & 25/32 & 9/8 & 47/32 \\
\end{bmatrix}
$$

$$
\mathbf{A}_1 =
\begin{bmatrix}
\mathbf{0} & & & & \mathbf{0} & & & \\
 & & 1 & 0 & 1/2 & 0 & 1/4 & 3/8 \\
 & & 0 & 1 & 1/2 & 1/2 & 1/2 & 1/2 \\
\mathbf{0} & & 1/2 & 1/2 & 1 & 1/4 & 5/8 & 7/16 \\
 & & 0 & 1/2 & 1/4 & 1 & 5/8 & 13/16 \\
 & & 1/4 & 1/2 & 5/8 & 5/8 & 9/8 & 7/8 \\
 & & 3/8 & 1/2 & 7/16 & 13/16 & 7/8 & 21/16 \\
\end{bmatrix}
$$

$$
\mathbf{A}_2 =
\begin{bmatrix}
\mathbf{0} & & \mathbf{0} & & & \\
 & & 1 & 0 & 1/2 & 3/4 \\
\mathbf{0} & & 0 & 1 & 1/2 & 1/4 \\
 & & 1/2 & 1/2 & 1 & 3/4 \\
 & & 3/4 & 1/4 & 3/4 & 5/4 \\
\end{bmatrix}
$$

$$
\mathbf{A}_3 =
\begin{bmatrix}
\mathbf{0} & & \mathbf{0} & \\
 & & 1 & 1/2 \\
\mathbf{0} & & 1/2 & 1 \\
\end{bmatrix}
$$

$\mathbf{A}_4$ is null except for the last element, which is 1.

$$\mathbf{h}' = [0 \quad 0 \quad 0 \quad 0 \quad 1/4 \quad 1/4 \quad 5/16 + \theta/8 \quad 15/32 + 9\theta/16]$$

Using $\theta = 10$ to demonstrate clearly what is happening,

$$
M = \begin{bmatrix}
1.000 & .000 & .500 & .500 & .500 & .250 & .375 & .438 \\
.000 & 1.000 & .500 & .500 & .500 & .750 & .625 & .563 \\
.500 & .500 & 11.000 & .500 & 5.750 & .500 & 3.125 & 4.438 \\
.500 & .500 & .500 & 11.000 & 5.750 & 5.750 & 5.750 & 5.750 \\
.500 & .500 & 5.750 & 5.750 & 21.250 & 3.125 & 12.188 & 16.719 \\
.250 & .750 & .500 & 5.750 & 3.125 & 16.250 & 9.688 & 6.406 \\
.375 & .625 & 3.125 & 5.750 & 12.188 & 9.688 & 30.063 & 21.125 \\
.438 & .563 & 4.438 & 5.750 & 16.719 & 6.406 & 21.125 & 40.844
\end{bmatrix}
$$

and

$$
M^{-1} = \begin{bmatrix}
1.048 & .048 & -.048 & -.048 & .000 & .000 & .000 & .000 \\
.048 & 1.067 & -.048 & -.028 & .000 & -.039 & .000 & .000 \\
-.048 & -.048 & .111 & .016 & -.032 & .000 & .000 & .000 \\
-.048 & -.028 & .016 & .131 & -.032 & -.039 & .000 & .000 \\
.000 & .000 & -.032 & -.032 & .089 & .013 & -.015 & -.023 \\
.000 & -.039 & .000 & -.039 & .013 & .090 & -.026 & .000 \\
.000 & .000 & .000 & .000 & -.015 & -.026 & .064 & -.023 \\
.000 & .000 & .000 & .000 & -.023 & .000 & -.023 & .046
\end{bmatrix}
$$

Entries of the **u**, **v**, and **h** vectors for the updated Quaas method are presented in Table 1; the first four entries of **h** are 0 in all rounds. Individual 4 is a common ancestor of individual 7; the segregation variance of individual 7 attributable to the new mutations arising in individual 4 is therefore reduced because the new mutations have been subject to inbreeding.

**Table 1**
*Entries for* **u**, **v**, *and* **h** *for each round*

| Entry | Round 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| **Vector u** | | | | | | | | |
| 1 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 2 | .000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 3 | .250 | .500 | 11.000 | 11.000 | 11.000 | 11.000 | 11.000 | 11.000 |
| 4 | .250 | .500 | .500 | 11.000 | 11.000 | 11.000 | 11.000 | 11.000 |
| 5 | .250 | .500 | 3.125 | 5.750 | 21.250 | 21.250 | 21.250 | 21.250 |
| 6 | .063 | .625 | .625 | 3.250 | 3.250 | 16.250 | 16.250 | 16.250 |
| 7 | .141 | .531 | 1.188 | 3.813 | 7.688 | 10.938 | 30.063 | 30.063 |
| 8 | .191 | .508 | 1.984 | 4.609 | 13.328 | 14.141 | 18.922 | 40.844 |
| **Vector v** | | | | | | | | |
| 1 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 2 | .000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 3 | .500 | .500 | 3.240 | 3.240 | 3.240 | 3.240 | 3.240 | 3.240 |
| 4 | .500 | .500 | .000 | 3.240 | 3.240 | 3.240 | 3.240 | 3.240 |
| 5 | .500 | .500 | 1.620 | 1.620 | 3.937 | 3.937 | 3.937 | 3.937 |
| 6 | .250 | .750 | .000 | 1.620 | .000 | 3.606 | 3.606 | 3.606 |
| 7 | .375 | .625 | .810 | 1.620 | 1.969 | 1.803 | 4.372 | 4.372 |
| 8 | .438 | .563 | 1.215 | 1.620 | 2.953 | .901 | 2.187 | 4.682 |
| **Vector h** | | | | | | | | |
| 5 | .125 | .250 | .250 | .250 | .250 | .250 | .250 | .250 |
| 6 | .000 | .250 | .250 | .250 | .250 | .250 | .250 | .250 |
| 7 | .063 | .250 | .250 | 1.563 | 1.563 | 1.563 | 1.563 | 1.563 |
| 8 | .094 | .250 | .906 | 2.219 | 6.094 | 6.094 | 6.094 | 6.094 |

## 6. Discussion

Mutation effects are easily incorporated into the rules for constructing the additive genetic variance–covariance matrix and its inverse, and so $M^{-1}$ could be incorporated into the mixed model equations for predicting breeding values if $\sigma_m^2$ or $\theta$ is known. If they are not known, then $M\sigma_a^2$ could be included in a restricted maximum likelihood (REML) analysis, iterating on different values of $\theta$, resulting in an estimate of $\sigma_m^2$. More explicitly, the likelihood function in the derivative-free REML method of Smith and Graser (1986) depends on a function of observations and $|A|$, the determinant of $A$. In a mutation model this will depend on $|M|$, which in turn is determined by the product of the $d_i$. This method is currently being used to estimate mutation variance in lines of mice selected for increased and decreased body weight (P. Keightley, personal communication). The base population mice are a standard inbred laboratory strain and therefore have a very low base additive genetic variance.

For simplicity, individuals with unidentified parents are assumed to be sampled from the base population. The procedure could be easily updated to account for the approximate generation of an individual and therefore impose expected accumulated mutation effects. However, it is anticipated that this procedure is most likely to be applicable to experimental populations in which full pedigree records are available.

### RÉSUMÉ

On adapte les procédures de calcul de la matrice de covariance génétique additive et son inverse pour tenir compte des mutations dans le génome. On peut utiliser la matrice inverse dans la méthodologie du modèle mixte pour la meilleure prédiction linéaire non biaisée des valeurs en croisement ou pour l'estimation de composantes de la variance quand un modèle génétique inclut les effets de mutation.

### REFERENCES

Dempfle, L. (1987). Problems in the use of the relationship matrix in animal breeding. In *Advances in Statistical Methods for Genetic Improvement of Livestock*, D. Gianola and K. Hammond (eds). Notes for participants of International Symposium, Armidale, Australia. Chapter 20, 492–511.

Dempfle, L. and Grundl, E. (1988). Identification of superior animals and their use in improvement programmes. In *Advances in Animal Breeding*, Symposium in honour of Professor R. D. Politiek, 56–72, Pudoc, Wageningen.

Dudley, J. W. (1977). 76 generations of selection for oil and protein percentage in maize. In *Proceedings of the International Conference on Quantitative Genetics*, E. Pollak, O. Kempthorne, and T. B. Bailey, Jr (eds), 459–473. Ames, Iowa: Iowa State University Press.

Emik, L. O. and Terrill, C. E. (1949). Systematic procedures for calculating inbreeding coefficients. *Journal of Heredity* **40**, 51–55.

Henderson, C. R. (1973). Sire evaluation and genetic trends. In *Proceedings of Animal Breeding and Genetics in Honor of Dr J. Lush*. ASAS and ASDS, Champaign, Illinois, 10–39.

Henderson, C. R. (1976). A simple method for computing the inverse of a numerator relationship matrix used in the prediction of breeding values. *Biometrics* **32**, 69–83.

Hill, W. G. (1982). Predictions of response to artificial selection from new mutations. *Genetical Research, Cambridge* **40**, 255–278.

Keightley, P. D. and Hill, W. G. (1988). Quantitative genetic variability maintained by mutation stabilising balance in finite populations. *Genetical Research, Cambridge* **52**, 33–43.

Lynch, M. (1988). The rate of polygenic mutation. *Genetical Research, Cambridge* **51**, 137–148.

Quaas, R. L. (1976). Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* **32**, 949–953.

Smith, S. P. and Graser, H. (1986). Estimating variance components in a class of mixed models by restricted maximum likelihood. *Journal of Dairy Science* **69,** 1156–1165.

Sorenson, D. A. and Kennedy, B. W. (1984). Estimation of genetic variances from unselected and selected populations. *Journal of Animal Science* **59,** 1213–1221.

Thompson, R. (1977). The estimation of heritability with unbalanced data. II. Data available on more than two generations. *Biometrics* **33,** 497–504.

Wright, S. (1922). Coefficients of inbreeding and relationship. *American Naturalist* **56,** 330–338.

# LINKED CITATIONS

*- Page 1 of 2 -*

*You have printed the following article:*

**Accounting for Mutation Effects in the Additive Genetic Variance-Covariance Matrix and Its Inverse**
Naomi R. Wray
*Biometrics*, Vol. 46, No. 1. (Mar., 1990), pp. 177-186.
Stable URL:

*This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.*

# References

**A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values**
C. R. Henderson
*Biometrics*, Vol. 32, No. 1. (Mar., 1976), pp. 69-83.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28197603%2932%3A1%3C69%3AASMFCT%3E2.0.CO%3B2-9

**Computing the Diagonal Elements and Inverse of a Large Numerator Relationship Matrix**
R. L. Quaas
*Biometrics*, Vol. 32, No. 4. (Dec., 1976), pp. 949-953.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28197612%2932%3A4%3C949%3ACTDEAI%3E2.0.CO%3B2-9

**The Estimation of Heritability with Unbalanced Data: II. Data Available on More Than Two Generations**
Robin Thompson
*Biometrics*, Vol. 33, No. 3. (Sep., 1977), pp. 497-504.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28197709%2933%3A3%3C497%3ATEOHWU%3E2.0.CO%3B2-2

# LINKED CITATIONS

*- Page 2 of 2 -*

**Coefficients of Inbreeding and Relationship**

Sewall Wright

*The American Naturalist*, Vol. 56, No. 645. (Jul. - Aug., 1922), pp. 330-338.

Stable URL: