ANIMAL BREEDING NOTES

CHAPTER 14M

BREEDING VALUE OF AN ANIMAL FOR A SINGLE TRAIT IN A MULTIBREED POPULATION

Assumptions

(a) n_b unrelated breeds,

(b) k_b loci per breed, k_b arbitrary,

(c) r_{bk} alleles per locus in breed b, r_{bk} arbitrary, and

(d) no linkage.

Multibreed Breeding Value

Genetic value of an individual based on the mean genotypic value of its progeny. Because parents pass on their genes (not their genotypes) to their progeny, the mean genotypic value of their progeny is determined by the average effects of the parent's alleles (Falconer, 1981, p 106). Thus, the breeding value of animal i for trait X in a multibreed population would be determined by the sum of the average effects of alleles all the breeds present in each of its parents:

$$u_{iX} = a_{sX} + a_{dX} \qquad [1]$$

where

u_{iX} = random variable representing the breeding value of individual i for trait X

a_{sX} = random variable representing the sum of the average effects of all alleles from all breeds affecting trait X coming from the gamete of the sire of animal i, i.e.,

$$a_{sX} = \left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} \delta_{bk} a_{bk} \right)_{sire}$$

, where a_{bk} belongs to the sire of animal i, k_b = number of loci in breed b, and δ_{bk} is a Kronecker delta, i.e., $\delta_{bk} = 0$ or 1; δ_{bk} will be zero $(2k_b/2)$ times and one $(2k_b/2)$ times, because a random sample of only ½ of the male alleles from each breed is expected to be passed on to individual i.

a_{dX} = random variable representing the sum of the average effects of all alleles affecting a trait, coming from the gamete of the dam of animal i, i.e.,

$$a_{dX} = \left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} \delta_{bk} a_{bk} \right)_{dam}$$, where a_{bk} belongs to the dam of animal i.

Average genetic effects are defined as deviations from the average gene at each locus within each breed; thus, the **expected value of u_{iX}** is:

$$\begin{aligned}
E[u_{iX}] &= E[a_{sX} + a_{dX}] \\
&= E[a_{sX}] + E[a_{dX}] \\
&= E\left[\left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} \delta_{bk} a_{bk} \right)_{sire} \right] + E\left[\left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} \delta_{bk} a_{bk} \right)_{dam} \right] \\
&= \left[\left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} E[\delta_{bk} a_{bk} \mid k, b] \right)_{sire} \right] + \left[\left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} E[\delta_{bk} a_{bk} \mid k, b] \right)_{dam} \right] \\
&= 0 + 0 \\
&= 0
\end{aligned}$$

The **variance of u_{iX}** is:

$$\begin{aligned}
var(u_{iX}) &= var(a_{sX} + a_{dX}) \\
&= var(a_{sX}) + var(a_{dX}) + 2\, cov(a_{sX}, a_{dX})
\end{aligned}$$

The covariance between the breeding values of animal i for traits X and Z is:

$$cov(u_{iX}, u_{iZ}) = cov(a_{sX} + a_{dX}, a_{sZ} + a_{dZ})$$

$$= \mathrm{cov}(a_{sX}, a_{sZ}) + \mathrm{cov}(a_{sX}, a_{dZ}) + \mathrm{cov}(a_{dX}, a_{sZ}) + \mathrm{cov}(a_{dX}, a_{dZ})$$

By **conditioning on the breeding values of the sire (u_{sX}) and the dam (u_{dX}) of animal i**, var(u_{iX}) becomes:

$$\begin{aligned}
\mathrm{var}(u_{iX}) = \; & \mathrm{var}(E[a_{sX} \mid u_{sX}]) + E[\mathrm{var}(a_{sX} \mid u_{sX})] \\
+ \; & \mathrm{var}(E(a_{dX} \mid u_{dX}]) + E[\mathrm{var}(a_{dX} \mid u_{dX})] \\
+ \; & 2\,\mathrm{cov}(E[a_{sX} \mid u_{sX}], E[a_{dX} \mid u_{dX}]) \\
+ \; & 2\,E[\mathrm{cov}(a_{sX} \mid u_{sX}, a_{dX} \mid u_{dX})]
\end{aligned}$$

But **a_{sX} is the average effect of ½ of the alleles affecting the trait in the sire**, i.e., $a_{sX} = \tfrac{1}{2} u_{sX}$.

Thus,

$$\begin{aligned}
E[a_{sX} \mid u_{sX}] &= E[\tfrac{1}{2} u_{sX} \mid u_{sX}] \\
&= \tfrac{1}{2} u_{sX}
\end{aligned}$$

Applying a similar argument to a_{dX} yields:

$$E[a_{dX} \mid u_{dX}] = \tfrac{1}{2} u_{dX}$$

Thus, for the **sire of animal i**,

$$\begin{aligned}
\mathrm{var}(E[a_{sX} \mid u_{sX}]) &= \mathrm{var}(\tfrac{1}{2} u_{sX}) \\
&= \tfrac{1}{4}\,\mathrm{var}(u_{sX}) \\
&= \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_s + \sum_{c=1}^{n_{cs}} \left(\tfrac{1}{2} a_{ss,ds}\right)_{cs} \left(\sigma^2_{AX}\right)_{cs} \right) \\
&= \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_s + \sum_{c=1}^{n_{cs}} F_{cs} \left(\sigma^2_{AX}\right)_{cs} \right)
\end{aligned}$$

(= $\tfrac{1}{4} (\sigma_A^2 + F_s \sigma_A^2)$ in a single breed population)

where

n_{cs} = number of common ancestors for sire s

$\left(a_{ss,ds}\right)_{cs}$ = additive relationship between the sire and the dam of sire s through common ancestor c

$\left(\sigma^2_{AX}\right)_s$ = multibreed additive genetic variance for trait X of the sire

$\left(\sigma^2_{AX}\right)_{cs}$ = multibreed additive genetic variance for trait X of the common ancestor of the sire and dam of sire s and, for the dam of animal i,

Similarly, for the **dam of animal i**,

$$\mathrm{var}(E[a_{dX} \mid u_{dX}]) = \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{cd}} \left(\tfrac{1}{2} a_{sd,dd}\right)_{cd} \left(\sigma^2_{AX}\right)_{cd}\right)$$

$$= \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{cd}} F_{cd} \left(\sigma^2_{AX}\right)_{cd}\right)$$

$(= \tfrac{1}{4}(\sigma_A^2 + F_d\,\sigma_A^2)$ in a single breed population)

What is $E[\mathrm{var}(a_{sX} \mid u_{sX})]$?

$$E[\mathrm{var}(a_{sX} \mid u_{sX})] = \mathrm{var}(a_{sX}) - \mathrm{var}(E[a_{sX} \mid u_{sX}])$$

Here,

$$\mathrm{var}(a_{sX}) = \mathrm{var}\left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} \delta_{bk} a_{bk}\right)_{sire}$$

Also, by definition,

$$\mathrm{var}\left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} a_{bk}\right) = \sigma_{AX}^2 = \text{multibreed additive genetic variance for trait X.}$$

An expression for the multibreed additive genetic variance for trait X can be obtained by conditioning on the breed of origin of alleles as follows:

$$\sigma_{AX}^2 = E[\mathrm{var}(X \mid b] + \mathrm{var}(E[X \mid b])$$

$$\sigma_{AX}^2 = \sum_{b=1}^{n_b} p_b^i \left(\sigma_{aX}^2\right)_b + \sum_{b=1}^{n_b-1} \sum_{b'>b}^{n_b} (p_b^s p_{b'}^s + p_b^d p_{b'}^d)\left(\sigma_{aX}^2\right)_{bb'}$$

where superscripts i, s and d correspond to animal, sire and dam, subscripts b and b′ represent two breeds, and

n_b = number of breeds

p_b^x = expected fraction of breed b in animal x, x = i, s, d

$(\sigma_{aX}^2)_b$ = additive intrabreed variance of trait X for breed b

$(\sigma_{aX}^2)_{bb'}$ = additive interbreed variance of trait X for the pair of breeds b and b′

Thus, for the sire of animal i,

$$\mathrm{var}(a_{sX}) = \mathrm{var}\left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} \delta_{bk} a_{bk}\right)_{\text{sire}}$$

$$= \tfrac{1}{2} (\sigma_{AX}^2)_s$$

and

$$E[\mathrm{var}(a_{sX} \mid u_{sX})] = \tfrac{1}{2}\left(\sigma_{AX}^2\right)_s - \tfrac{1}{4}\left(\left(\sigma_{AX}^2\right)_s + \sum_{c=1}^{n_{cs}} \left(\tfrac{1}{2} a_{ss,ds}\right)_{cs} \left(\sigma_{AX}^2\right)_{cs}\right)$$

$$= \tfrac{1}{4}\left(\left(\sigma_{AX}^2\right)_s - \sum_{c=1}^{n_{cs}} \left(\tfrac{1}{2} a_{ss,ds}\right)_{cs} \left(\sigma_{AX}^2\right)_{cs}\right)$$

$$= \tfrac{1}{4}\left(\left(\sigma_{AX}^2\right)_s - \sum_{c=1}^{n_{cs}} F_{cs} \left(\sigma_{AX}^2\right)_{cs}\right)$$

(= $\tfrac{1}{4} (\sigma_A^2 - F_s \sigma_A^2)$ in a single breed population)

Similarly, for the dam of animal i,

$$\text{var}(a_{dX}) \quad = \quad \text{var}\left(\sum_{b=1}^{n_b}\sum_{k=1}^{2k_b}\delta_{bk}a_{bk}\right)_{\text{dam}} \quad = \quad ½\,(\sigma_{AX}^2)_d$$

and

$$\text{E}[\text{var}(a_{dX}\,|\,u_{dX})] \quad = \quad ½\left(\sigma^2_{AX}\right)_d - ¼\left(\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{cd}}\left(½\,a_{sd,dd}\right)_{cd}\left(\sigma^2_{AX}\right)_{cd}\right)$$

$$= \quad ¼\left(\left(\sigma^2_{AX}\right)_d - \sum_{c=1}^{n_{cd}}\left(½\,a_{sd,dd}\right)_{cd}\left(\sigma^2_{AX}\right)_{cd}\right)$$

$$= \quad ¼\left(\left(\sigma^2_{AX}\right)_d - \sum_{c=1}^{n_{cd}}F_{cd}\left(\sigma^2_{AX}\right)_{cd}\right)$$

(= ¼ (σ_A^2 – F_d σ_A^2) in a single breed population)

The covariance terms of var(u_{iX}) are:

$$2\,\text{cov}(\text{E}[a_{sX}\,|\,u_{sX}],\,\text{E}[a_{dX}\,|\,u_{dX}]) \quad = \quad 2\,\text{cov}(½\,u_{sX},\,½\,u_{dX})$$

$$= \quad ½\,\text{cov}(u_{sX},\,u_{dX})$$

$$= \quad ½\left(\sum_{c=1}^{n_{ci}}\left(a_{s,d}\right)_{ci}\left(\sigma^2_{AX}\right)_{ci}\right)$$

$$= \quad \sum_{c=1}^{n_{ci}}\left(½\,a_{s,d}\right)_{ci}\left(\sigma^2_{AX}\right)_{ci}$$

$$= \quad \sum_{c=1}^{n_{ci}}F_{ci}\left(\sigma^2_{AX}\right)_{ci}$$

(= F_i σ_A^2 in a single breed population)

The value of 2 E[cov(a_{sX} | u_{sX}, a_{dX} | u_{dX})] = 0 because the sire and the dam random samples of ½ of their respective set of alleles are taken independently of each other, i.e., there is no connection between the formation of sire gametes and dam gametes (biological fact).

Thus,

$$\mathrm{var}(u_{iX}) = \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_s + \sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma^2_{AX}\right)_{cs}\right) + \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_s - \sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma^2_{AX}\right)_{cs}\right)$$

$$+ \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma^2_{AX}\right)_{cd}\right) + \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_d - \sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma^2_{AX}\right)_{cd}\right)$$

$$+ \sum_{c=1}^{n_{cs}} F_{ci}\left(\sigma^2_{AX}\right)_{ci} + 0$$

$$\mathrm{var}(u_{iX}) = \tfrac{1}{2}\left(\sigma^2_{AX}\right)_s + \tfrac{1}{2}\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{ci}} F_{ci}\left(\sigma^2_{AX}\right)_{ci}$$

$$= \left(\sigma^2_{AX}\right)_i + \sum_{c=1}^{n_{ci}} F_{ci}\left(\sigma^2_{AX}\right)_{ci}$$

Equivalent model for the breeding value of an animal for a single trait in a multibreed population

An equivalent model (same first and second moments; Henderson, 1985) to the single trait multibreed breeding value model,

$$u_{iX} = a_{sX} + a_{dX}$$

$$E[u_x] = 0$$

$$\mathrm{var}(u_{iX}) = \left(\sigma^2_{AX}\right)_i + \sum_{c=1}^{n_{ci}} F_{ci}\left(\sigma^2_{AX}\right)_{ci}$$

is:

$$u_x = \tfrac{1}{2}\, u_s + \tfrac{1}{2}\, u_d + \tfrac{1}{2}\, \varepsilon_s + \tfrac{1}{2}\, \varepsilon_d \qquad [2]$$

where

u_{iX} = breeding value of animal i for trait X (random),

u_{sX} = breeding value of sire s for trait X (random),

u_{dX} = breeding value of dam d for trait X (random),

ε_{sX} = Mendelian sampling occurring during gametogenesis in sire s for trait X (random),

ε_{dX} = Mendelian sampling in dam d for trait X (random).

The Mendelian sampling terms ε_{sX} and ε_{dX} are independent of each other and independent from breeding values. All random variables have expected values equal to zero, i.e.,

$$E[u_{iX}] = 0$$

and

$$\mathrm{var}(u_{iX}) = \mathrm{var}(½\, u_{sX} + ½\, u_{dX}) + \mathrm{var}(½\, \varepsilon_{sX}) + \mathrm{var}(½\, \varepsilon_{dX})$$

where

$$\begin{aligned}\mathrm{var}(½\, u_{sX} + ½\, u_{dX}) &= \mathrm{var}(½\, u_{sX}) + \mathrm{var}(½\, u_{dX}) + 2\,\mathrm{cov}(½\, u_{sX}, ½\, u_{dX}) \\ &= ¼\left(\left(\sigma^2_{AX}\right)_s + \sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma^2_{AX}\right)_{cs}\right) + ¼\left(\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma^2_{AX}\right)_{cd}\right) \\ &\quad + \sum_{c=1}^{n_{cs}} F_{ci}\left(\sigma^2_{AX}\right)_{ci}\end{aligned}$$

(= $¼\ (1 + F_s)\ \sigma_A^2 + ¼\ (1 + F_d)\ \sigma_A^2 + F_x\ \sigma_A^2$ in a single breed population)

Because the sampling process during gamete formation (Mendelian sampling) in the male and female gametes is completely independent of each other, any loss of variation during this process would be due to (i) inbreeding in the male present in the male gamete, and (ii) inbreeding in the female present in the female gamete. Thus,

$$\mathrm{var}(½\, \varepsilon_{sX}) = ¼\left(\sigma^2_{AX}\right)_s - ¼\left(\sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma^2_{AX}\right)_{cs}\right)$$

$$= \quad \frac{1}{4}\left(\left(\sigma^2_{AX}\right)_s - \sum_{c=1}^{n_{cs}} F_{cs} \left(\sigma^2_{AX}\right)_{cs} \right)$$

(= ¼ (1 – F_s) σ_A^2 in a single breed population)

and

$$\text{var}(\tfrac{1}{2}\,\varepsilon_{dX}) \quad = \quad \frac{1}{4}\left(\sigma^2_{AX}\right)_d - \frac{1}{4}\left(\sum_{c=1}^{n_{cd}} F_{cd} \left(\sigma^2_{AX}\right)_{cd} \right)$$

$$= \quad \frac{1}{4}\left(\left(\sigma^2_{AX}\right)_d - \sum_{c=1}^{n_{cd}} F_{cd} \left(\sigma^2_{AX}\right)_{cd} \right)$$

(= ¼ (1 – F_d) σ_A^2 in a single breed population)

and

$$\text{var}(u_{iX}) \quad = \quad \left(\sigma^2_{AX}\right)_i + \sum_{c=1}^{n_{ci}} F_{ci} \left(\sigma^2_{AX}\right)_{ci}$$

(= (1 + F_x) σ_A^2 in a single breed population)

Remarks: var(½ ε_{sX}) + var(½ ε_{dX}) = var(φ_X) represents the variability that exists among gametes in the sire and in the dam. Thus, it is affected negatively by the level of inbreeding in the sire (F_{sX}) and dam (F_{dX}) and it is independent of the level of inbreeding of the individual (F_{iX}).

Derivation of general rules to compute the inverse of the multibreed additive covariance matrix directly

The multibreed additive genetic model can be generalized to include not only the case when both parents are identified, but also the cases when only the sire, only the dam or neither parent is known.

Unidentified parents are assumed to be **non-inbred and unrelated** among themselves and to all other identified animals in the population, just like base animals.

Base animals are those that all other animals in the population descend from. They are assumed to be unrelated and non-inbred.

The **extended version of the multibreed additive genetic model** with animals ordered such that parents precede progeny (e.g., by birth date), in matrix notation, is:

$$u_X = \tfrac{1}{2} P u_X + \varphi_X \qquad [3]$$

where

u_X = vector multibreed breeding values of animals for trait X, where parents precede progeny;

P = lower triangular matrix relating parents to progeny. A row of P contains ones in the columns corresponding to the known parents and zeroes elsewhere. Thus, a row of P contains:

- (i) two 1's if both parents are known,
- (ii) one 1 if either the sire or the dam of an animal is known,
- (iii) zeroes if both parents are unknown;

φ_X = vector of independent random variables representing:

(i) Mendelian sampling in the sire and in the dam, if both parents of animal i are known, i.e.,

$$\varphi_{iX} = \tfrac{1}{2} \varepsilon_{sX} + \tfrac{1}{2} \varepsilon_{dX}$$

(ii) Mendelian sampling in the sire and the dam plus:

(a) the breeding value of the dam, if only the sire of animal i is known, i.e.,

$$\varphi_{iX} = \tfrac{1}{2} u_{dX} + \tfrac{1}{2} \varepsilon_{sX} + \tfrac{1}{2} \varepsilon_{dX}$$

(b) the breeding value of the sire, if only the dam animal i is known, i.e.,

$$\varphi_{iX} = ½\, u_{sX} + ½\, \varepsilon_{sX} + ½\, \varepsilon_{dX}$$

(iii) the breeding value of animal i if neither its sire nor its dam are known, i.e.,

$$\varphi_{iX} = u_{iX}$$

From equation [3] we can see that:

$$\varphi_X = (I - ½\, P)\, u_X$$

$$\Rightarrow \quad u_X = (I - ½\, P)^{-1}\, \varphi_X$$

$$\Rightarrow \quad E[u_X] = (I - ½\, P)^{-1}\, E[\varphi_X]$$

$$= 0$$

$$\Rightarrow \quad \mathrm{var}(u_X) = (I - ½\, P)^{-1}\, \mathrm{var}(\varphi_X)\, (I - ½\, P')^{-1}$$

$$\mathrm{var}(u_X) = (I - ½\, P)^{-1}\, D\, (I - ½\, P')^{-1}$$

where

$$D = \mathrm{diag}\{\mathrm{var}(\varphi_{iX})\}$$

$$= \mathrm{diag}\,\{d_{ii}\}$$

The $\mathrm{var}(\varphi_{iX}) = d_{ii}$ are:

(i) $$\mathrm{var}\left(\frac{1}{2}\varepsilon_{s_X} + \frac{1}{2}\varepsilon_{d_X}\right) = \mathrm{var}\,(u_{iX}) - \mathrm{var}\left(\frac{1}{2}u_{sX} + \frac{1}{2}u_{dX}\right)$$

$$= \left(\sigma^2_{AX}\right)_i + \sum_{c=1}^{n_{ci}} F_{ci}\left(\sigma^2_{AX}\right)_{ci} - ¼\left(\left(\sigma^2_{AX}\right)_s + \sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma^2_{AX}\right)_{cs}\right)$$

$$- ¼\left(\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma^2_{AX}\right)_{cd}\right) - \sum_{c=1}^{n_{ci}} F_{ci}\left(\sigma^2_{AX}\right)_{ci}$$

$$= \left(\sigma^2_{AX}\right)_i - ¼\left(\sigma^2_{AX}\right)_s - ¼\left(\sigma^2_{AX}\right)_d - ¼\sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma^2_{AX}\right)_{cs}$$

$$- ¼\sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma^2_{AX}\right)_{cd}$$

if s and d are known

$$\text{(ii)}\ \ \mathrm{var}\left(\frac{1}{2}u_{dx} + \frac{1}{2}\varepsilon_{sx} + \frac{1}{2}\varepsilon_{dx}\right) = \mathrm{var}\ (u_{iX}) - \mathrm{var}\left(\frac{1}{2}u_{sx}\right)$$

$$= \left(\sigma^2_{AX}\right)_i - \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_s + \sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma^2_{AX}\right)_{cs}\right)$$

$$= \left(\sigma^2_{AX}\right)_i - \tfrac{1}{4}\left(\sigma^2_{AX}\right)_s - \tfrac{1}{4}\sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma^2_{AX}\right)_{cs}$$

if s is known only

$$\text{(iii)}\ \ \mathrm{var}\left(\frac{1}{2}u_{sx} + \frac{1}{2}\varepsilon_{sx} + \frac{1}{2}\varepsilon_{dx}\right) = \mathrm{var}\ (u_{iX}) - \mathrm{var}\left(\frac{1}{2}u_{dx}\right)$$

$$= \left(\sigma^2_{AX}\right)_i - \tfrac{1}{4}\left(\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma^2_{AX}\right)_{cd}\right)$$

$$= \left(\sigma^2_{AX}\right)_i - \tfrac{1}{4}\left(\sigma^2_{AX}\right)_d - \tfrac{1}{4}\sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma^2_{AX}\right)_{cd}$$

if d is known only

$$\text{(iv)}\ \ \mathrm{var}\left(\frac{1}{2}u_{sx} + \frac{1}{2}u_{dx} + \frac{1}{2}\varepsilon_{sx} + \frac{1}{2}\varepsilon_{dx}\right) = \mathrm{var}(u_{iX})$$

$$= \left(\sigma^2_{AX}\right)_i$$

if neither s nor d are known

But

$$\mathrm{var}(u_X) = \left(I - \frac{1}{2}P\right)^{-1} D \left(I - \frac{1}{2}P'\right)^{-1}$$

$$= G_{AX}$$

$$\Rightarrow \qquad G_{AX}^{-1} \;=\; \left(I - \frac{1}{2}P'\right) D^{-1} \left(I - \frac{1}{2}P\right)$$

$$\begin{array}{ccccccccccc} \Rightarrow & G_{AX}^{-1} & = & D^{-1} & - & \tfrac{1}{2} D^{-1} P & - & \tfrac{1}{2} P' D^{-1} & + & \tfrac{1}{4} P' D^{-1} P \\ & & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ & & & \text{diagonals} & & & \text{parent–progeny} & & & \text{parent–parent} \end{array}$$

Let $P = \begin{bmatrix} p_1' \\ p_2' \\ \vdots \\ p_n' \end{bmatrix}$ where p_i' is a vector with at most two 1's and the rest zeroes.

Thus,

$$-\tfrac{1}{2} D^{-1} P \;=\; \begin{bmatrix} -\frac{1}{2} d_{11}^{-1} p_1' \\ \vdots \\ -\frac{1}{2} d_{nn}^{-1} p' \end{bmatrix}$$

$$= \left\{ \begin{array}{cccccc} & & [\text{sire}] & & [\text{dam}] & \\ & & \vdots & & \vdots & \\ \cdots & -\frac{1}{2} d_{ii}^{-1} & \cdots & -\frac{1}{2} d_{ii}^{-1} & \cdots & [\text{animal}] \\ & & \vdots & & \vdots & \end{array} \right\}$$

$$-\tfrac{1}{2} P' D^{-1} \;=\; \begin{bmatrix} -\frac{1}{2} d_{11}^{-1} p_1 & \cdots & -\frac{1}{2} d_{nn}^{-1} p_n \end{bmatrix}$$

$$= \left\{\begin{array}{cccc} & & [\text{animal}] & \\ & & \vdots & \\ \cdots & -\frac{1}{2}\,d_{ii}^{-1} & \cdots & [\text{sire}] \\ & & \vdots & \\ \cdots & -\frac{1}{2}\,d_{ii}^{-1} & \cdots & [\text{dam}] \\ & & \vdots & \end{array}\right\}$$

$$\tfrac{1}{4}\,P'\,D^{-1}\,P = \sum_{i=1}^{n} d_{ii}^{-1}\,p_i\,p_i'$$

$$= \left\{\begin{array}{cccccc} & [\text{sire}] & & [\text{dam}] & & \\ & \vdots & & \vdots & & \\ \cdots & \frac{1}{4}\,d_{ii}^{-1} & \cdots & \frac{1}{4}\,d_{ii}^{-1} & \cdots & [\text{sire}] \\ & \vdots & & \vdots & & \\ \cdots & \frac{1}{4}\,d_{ii}^{-1} & \cdots & \frac{1}{4}\,d_{ii}^{-1} & \cdots & [\text{dam}] \\ & \vdots & & \vdots & & \end{array}\right\}$$

Based on the pattern of contributions of the four matrices contributing to G_{AX}^{-1}, the following **rules to compute** G_{AX}^{-1} **directly** can be inferred:

(1) if both the sire (s) and the dam (d) of animal i are identified, add:

d_{ii}^{-1} to $i \times i$

$-\frac{1}{2}\,d_{ii}^{-1}$ to $i \times s$, $i \times d$, $s \times i$, $d \times i$

$\frac{1}{4}\,d_{ii}^{-1}$ to $s \times s$, $s \times d$, $d \times s$, $d \times d$

where

$$d_{ii}^{-1} = \left[\left(\sigma_{AX}^2\right)_i - \tfrac{1}{4}\left(\left(\sigma_{AX}^2\right)_s + \sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma_{AX}^2\right)_{cs} \right) - \tfrac{1}{4}\left(\left(\sigma_{AX}^2\right)_d + \sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma_{AX}^2\right)_{cd} \right) \right]^{-1}$$

(2) if only the sire (s) is known, add:

d_{ii}^{-1} to i × i

$-\frac{1}{2} d_{ii}^{-1}$ to i × s, s × i

$\frac{1}{4} d_{ii}^{-1}$ to s × s

where

$$d_{ii}^{-1} = \left[\left(\sigma_{AX}^2\right)_i - \tfrac{1}{4}\left(\left(\sigma_{AX}^2\right)_s + \sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma_{AX}^2\right)_{cs} \right) \right]^{-1}$$

(3) if only the dam (d) is known, add:

d_{ii}^{-1} to i × i

$-\frac{1}{2} d_{ii}^{-1}$ to i × d, d × i

$\frac{1}{4} d_{ii}^{-1}$ to d × d

where

$$d_{ii}^{-1} = \left[\left(\sigma_{AX}^2\right)_i - \tfrac{1}{4}\left(\left(\sigma_{AX}^2\right)_d + \sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma_{AX}^2\right)_{cd} \right) \right]^{-1}$$

(4) if neither s nor d are known, add:

d_{ii}^{-1} to i × i

where

$$d_{ii}^{-1} = \left[\left(\sigma_{AX}^2\right)_i\right]^{-1}$$

These computational rules correspond to the **multibreed version** of Henderson's rules.

In order to apply these rules we **must know the d_{ii}**. If there is no inbreeding the d_{ii} can be easily computed using only the multibreed additive genetic covariances of the genetic groups of the animal, its sire, and its dam. But, if there is inbreeding, the computation of the d_{ii} requires knowledge of the inbreeding coefficients of each common ancestor of an animal and its corresponding multibreed variance for trait X. Because we are computing G_{AX}^{-1} without computing G_{AX} first, it is easier to compute the d_{ii} directly using a recursive procedure based on computing C, where $CC' = G_{AX}$. This approach will be used here.

Rules to compute the d_{ii} in a non-inbred multibreed population

If all animals in a multibreed population are non-inbred then:

(1) if both the sire (s) and the dam (d) of animal i are identified, add:

d_{ii}^{-1} to $i \times i$

$-\frac{1}{2} d_{ii}^{-1}$ to $i \times s, i \times d, s \times i, d \times i$

$\frac{1}{4} d_{ii}^{-1}$ to $s \times s, s \times d, d \times s, d \times d$

where

$$d_{ii}^{-1} = \left[\left(\sigma_{AX}^2\right)_i - {}^1\!/_4\left(\sigma_{AX}^2\right)_s - {}^1\!/_4\left(\sigma_{AX}^2\right)_d\right]^{-1}$$

(2) if only the sire (s) is known, add:

d_{ii}^{-1} to $i \times i$

$-\frac{1}{2}d_{ii}^{-1}$ to i × s, s × i

$\frac{1}{4}d_{ii}^{-1}$ to s × s

where

$$d_{ii}^{-1} = \left[\left(\sigma_{AX}^2\right)_i - {}^1\!/\!_4\left(\sigma_{AX}^2\right)_s\right]^{-1}$$

(3) if only the dam (d) is known, add:

d_{ii}^{-1} to i × i

$-\frac{1}{2}d_{ii}^{-1}$ to i × d, d × i

$\frac{1}{4}d_{ii}^{-1}$ to d × d

where

$$d_{ii}^{-1} = \left[\left(\sigma_{AX}^2\right)_i - {}^1\!/\!_4\left(\sigma_{AX}^2\right)_d\right]^{-1},$$

(4) if neither s nor d are known, add:

d_{ii}^{-1} to i × i

where

$$d_{ii}^{-1} = \left[\left(\sigma_{AX}^2\right)_i\right]^{-1}$$

Example of G_{AX}^{-1} in a non-inbred multibreed population

	Animal		Sire		Dam	
	Number	Breed Group	Number	Breed Group	Number	Breed Group
Parents	1	A	0	A	0	A
	2	B	0	B	0	B
	3	½ A ½ B	0	A	2	B
Non-parents	4	½ A ½ B	1	A	0	B
	5	½ A ½ B	1	A	2	B
	6	¾ A ¼ B	1	A	3	½ A ½ B

Intrabreed additive genetic variance of trait X for breed A = σ_{AAX}^2 = 36

Intrabreed additive genetic variance of trait X for breed B = σ_{BBX}^2 = 9

Interbreed additive genetic variance of trait X for combination of breeds AB = σ_{ABX}^2 = 4

Multibreed additive genetic variances for the 6 animals are computed using the expression:

$$\sigma_{AX}^2 = \sum_{b=1}^{n_b} p_b^i \left(\sigma_{aX}^2\right)_b + \sum_{b=1}^{n_b-1} \sum_{b'>b}^{n_b} (p_b^s p_{b'}^s + p_b^d p_{b'}^d)\left(\sigma_{aX}^2\right)_{bb'}$$

	Animal		
	Number	**Breed Group**	**Multibreed Additive Genetic Variance**
Parents	1	A	36
	2	B	9
	3	½ A ½ B	½ (36) + ½ (9) = 22.5
Non-parents	4	½ A ½ B	½ (36) + ½ (9) = 22.5
	5	½ A ½ B	½ (36) + ½ (9) = 22.5
	6	¾ A ¼ B	¾ (36) + ¼ (9) + [(1)(0) + (½)(½)](4) = 30.25

The $(\mathbf{d_{ii}})^{-1}$ values for the 6 animals are:

	Animal		
	Number	**Breed Group**	$(\mathbf{d_{ii}})^{-1}$
Parents	1	A	$(36)^{-1}$
	2	B	$(9)^{-1}$
	3	½ A ½ B	$22.5 – ¼\ (9) = (20.25)^{-1}$
Non-parents	4	½ A ½ B	$22.5 – ¼\ (36) = (13.5)^{-1}$
	5	½ A ½ B	$22.5 – ¼\ (36) – ¼\ (9) = (11.25)^{-1}$
	6	¾ A ¼ B	$30.25 – ¼\ (36) – ¼\ (22.5) = (15.625)^{-1}$

The inverse of the G_{AX} matrix is: $G_{AX}^{-1} = D^{-1} - ½\ D^{-1}\ P - ½\ P'\ D^{-1} + ¼\ P'\ D^{-1}\ P$, where :

$$P = \left[\begin{array}{ccc|ccc} 0 & & & & & \\ 0 & 0 & & & & \\ 0 & 1 & 0 & & & \\ \hline 1 & 0 & 0 & 0 & & \\ 1 & 1 & 0 & 0 & 0 & \\ 1 & 0 & 1 & 0 & 0 & 0 \end{array}\right],$$

$$D^{-1} = \left[\begin{array}{ccc|ccc} (36)^{-1} & & & & & \\ & (9)^{-1} & & & & \\ & & (20.25)^{-1} & & & \\ \hline & & & (13.5)^{-1} & & \\ & & & & (11.25)^{-1} & \\ & & & & & (15.625)^{-1} \end{array}\right], \text{ thus}$$

$$G_{AX}^{-1} = \begin{array}{cc} \left[\begin{array}{ccc|ccc} g^{11} & g^{12} & g^{13} & g^{14} & g^{15} & g^{16} \\ & g^{22} & g^{23} & 0 & g^{25} & 0 \\ & & g^{33} & 0 & 0 & g^{36} \\ - & - & - & - & - & - \\ & & & g^{44} & 0 & 0 \\ & & & & g^{55} & 0 \\ & & & & & g^{66} \end{array}\right] & \begin{array}{l} \\ \text{parents} \\ \\ \\ \\ \text{nonparents} \\ \\ \end{array} \\ \text{parents} \qquad \text{non-parents} & \end{array}$$

where each g^{ij} term is computed using the recursive multibreed rules above, i.e.,

$$g^{11} = (36)^{-1} + ¼\,(13.5)^{-1} + ¼\,(11.25)^{-1} + ¼\,(15.625)^{-1}$$

$$g^{12} = ¼\,(11.25)^{-1}$$

$$g^{13} = ¼\,(15.625)^{-1}$$

$$g^{14} = -½\,(13.5)^{-1}$$

$$g^{15} = -½\,(11.25)^{-1}$$

$$g^{16} = -½\,(15.625)^{-1}$$

$$g^{22} = (9)^{-1} + ¼\,(20.25)^{-1} + ¼\,(11.25)^{-1}$$

$$g^{23} = -½\,(20.25)^{-1}$$

$$g^{25} = -½\,(11.25)^{-1}$$

$$g^{33} = (20.25)^{-1} + ¼\,(11.25)^{-1}$$

$$g^{36} = -½\,(15.625)^{-1}$$

$$g^{44} = (13.5)^{-1}$$

$$g^{55} = (11.25)^{-1}$$

$$g^{66} = (15.625)^{-1}$$

Multibreed Inbred Population

The multibreed additive genetic covariance of an animal with itself for trait X is:

$$g_{ii} = \mathrm{var}(u_{iX}) = \left(\sigma^2_{AX}\right)_i + \sum_{c=1}^{n_{ci}} \left(\tfrac{1}{2} a_{sd}\right)_{ci} \left(\sigma^2_{AX}\right)_{ci}$$

or

$$g_{ii} = \mathrm{var}(u_{iX}) = \left(\sigma^2_{AX}\right)_i + \sum_{c=1}^{n_{ci}} F_{ci} \left(\sigma^2_{AX}\right)_{ci}$$

where

$\left(a_{sd}\right)_{ci}$ = additive genetic relationship between the sire (s) and the dam (d) of animal i through common ancestor ci

F_{ci} = coefficient of inbreeding of animal i due to common ancestor ci

= $\left(\tfrac{1}{2} a_{sd}\right)_{ci}$

In an inbred multibreed population the diagonal elements of the matrix D, i.e., the d_{ii}, will depend on the coefficient of inbreeding of the parents of the animals included in the relationship matrix A. Thus, the d_{ii} cannot be computed based solely on knowledge of $(\sigma_{AX}^2)_{bx}$, b = 1 to n_{bx}, where n_{bx} = number of breed group combinations. We also need to know the coefficients of inbreeding of the parents thorough each one of their common ancestors weighted by the multibreed additive genetic covariances of their breed group. However, we do **not** need to compute the complete matrix G_{AX}, only need the diagonal elements, because

$$d_{ii} = \left(\sigma^2_{AX}\right)_i - \frac{\delta_s}{4}\left(\left(\sigma^2_{AX}\right)_s + \sum_{c=1}^{n_{cs}} F_{cs}\left(\sigma^2_{AX}\right)_{cs}\right) - \frac{\delta_d}{4}\left(\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{cd}} F_{cd}\left(\sigma^2_{AX}\right)_{cd}\right)$$

where

$$\delta_s(\delta_d) = \begin{cases} 1 & \text{if } s(d) \text{ is known} \\ 0 & \text{otherwise} \end{cases}$$

Thus, what we need is an efficient method to compute the diagonal elements of G_{AX}, i.e., the g_{ii}.

This is accomplished by a recursive algorithm (Elzo, 1990), based on an algorithm developed by Quaas (1976) to compute the diagonals of the additive relationship matrix.

Recursive method to compute the diagonals of the multibreed additive genetic covariance matrix G_{AX}

The multibreed additive genetic covariance matrix G_{AX} is equal to:

$$G_{AX} = \left(I - \frac{1}{2}P\right)^{-1} D \left(I - \frac{1}{2}P'\right)^{-1}$$

Claim:

$$\left(I - \frac{1}{2}P\right)^{-1} = I + \frac{1}{2}P + \left(\frac{1}{2}P\right)^2 + \left(\frac{1}{2}P\right)^3 + \ldots + \left(\frac{1}{2}P\right)^m$$

where

m = maximum number of generations separating two individuals in G_{AX}, $m \leq n$, where

n = order of matrix G_{AX}

= number of animals in the pedigree

Proof (Quaas, 1986): P is a lower triangular matrix with zeroes on and above the diagonal. Thus, $P^{m+1} = 0$ for $m \leq n$.

Consider,

$$T = I + \frac{1}{2}P + \left(\frac{1}{2}P\right)^2 + \ldots + \left(\frac{1}{2}P\right)^m$$

and

$$\left(\frac{1}{2}P\right) T = \frac{1}{2}P + \left(\frac{1}{2}P\right)^2 + \left(\frac{1}{2}P\right)^3 + \ldots + \left(\frac{1}{2}P\right)^m + \left(\frac{1}{2}P\right)^{m+1}$$

Subtract the second term from the first term:

$$T - \left(\frac{1}{2}P\right) T = I - \left(\frac{1}{2}P\right)^{m+1}$$

But

$$\left(\frac{1}{2}P\right)^{m+1} = 0$$

Thus,

$$\Rightarrow \qquad T\left(I - \frac{1}{2}P\right) = I$$

$$\Rightarrow \qquad T = \left(I - \frac{1}{2}P\right)^{-1}$$

$$\Rightarrow \qquad \left(I - \frac{1}{2}P\right)^{-1} = I + \frac{1}{2}P + \left(\frac{1}{2}P\right)^2 + \ldots + \left(\frac{1}{2}P\right)^m$$

Remarks:

The rows of the matrices P, P^2, P^3, ..., identify parents, grandparents, great-grandparents, ..., i.e.,

Rows of	identify
P	parents
P^2	grandparents
P^3	great-grandparents
$\vdots$	$\vdots$
P^m	ancestors m generations back from the current generation

Example of G_{AX}^{-1} in an inbred multibreed population

	Animal		Sire		Dam	
	Number	**Breed Group**	**Number**	**Breed Group**	**Number**	**Breed Group**
Parents	1	A	0	A	0	A
	2	B	1	B	0	B
	3	½ A ½ B	1	A	2	B
	4	¼ A ¾ B	3	½ A ½ B	2	B
Non-parents	5	⅜ A ⅝ B	3	½ A ½ B	4	¼ A ¾ B
	6	⅜ A ⅝ B	3	½ A ½ B	4	¼ A ¾ B

Intrabreed additive genetic variance of trait X for breed A = $\sigma_{AAX}^2 = 36$

Intrabreed additive genetic variance of trait X for breed B = $\sigma_{BBX}^2 = 9$

Interbreed additive genetic variance of trait X for combination of breeds AB = $\sigma_{ABX}^2 = 4$

	Animal		
	Number	**Breed Group**	**Multibreed Additive Genetic Variance**
Parents	1	A	36
	2	B	9
	3	½ A ½ B	½ (36) + ½ (9) = 22.5
	4	¼ A ¾ B	¼ (36) + ¾ (9) + (½)(½)(4) = 16.75
Non-parents	5	⅜ A ⅝ B	⅜ (36) + ⅝ (9) + [(½)(½) + (¼)(¾)](4) = 20.875
	6	⅜ A ⅝ B	⅜ (36) + ⅝ (9) + [(½)(½) + (¼)(¾)](4) = 20.875

$$
P = \left[\begin{array}{cccc|cc}
0 & & & & & \\
1 & 0 & & & & \\
1 & 1 & 0 & & & \\
0 & 1 & 1 & 0 & & \\
\hline
0 & 0 & 1 & 1 & 0 & \\
0 & 0 & 1 & 1 & 0 & 0
\end{array}\right]
$$

$$
P^2 = P\,P = \underset{\substack{\downarrow \\ \text{parents}}}{\left[\begin{array}{cccc|cc}
0 & & & & & \\
1 & 0 & & & & \\
1 & 1 & 0 & & & \\
0 & 1 & 1 & 0 & & \\
\hline
0 & 0 & 1 & 1 & 0 & \\
0 & 0 & 1 & 1 & 0 & 0
\end{array}\right]}
\left[\begin{array}{cccc|cc}
0 & & & & & \\
1 & 0 & & & & \\
1 & 1 & 0 & & & \\
0 & 1 & 1 & 0 & & \\
\hline
0 & 0 & 1 & 1 & 0 & \\
0 & 0 & 1 & 1 & 0 & 0
\end{array}\right]
= \underset{\substack{\downarrow \\ \text{grandparents}}}{\left[\begin{array}{cccc|cc}
0 & & & & & \\
0 & 0 & & & & \\
1 & 0 & 0 & & & \\
2 & 1 & 0 & 0 & & \\
\hline
1 & 2 & 1 & 0 & 0 & \\
1 & 2 & 1 & 0 & 0 & 0
\end{array}\right]}
$$

$$
P^3 = P^2\,P = \left[\begin{array}{cccc|cc}
0 & & & & & \\
0 & 0 & & & & \\
1 & 0 & 0 & & & \\
2 & 1 & 0 & 0 & & \\
\hline
1 & 2 & 1 & 0 & 0 & \\
1 & 2 & 1 & 0 & 0 & 0
\end{array}\right]
\left[\begin{array}{cccc|cc}
0 & & & & & \\
1 & 0 & & & & \\
1 & 1 & 0 & & & \\
0 & 1 & 1 & 0 & & \\
\hline
0 & 0 & 1 & 1 & 0 & \\
0 & 0 & 1 & 1 & 0 & 0
\end{array}\right]
= \underset{\substack{\downarrow \\ \text{great-grandparents}}}{\left[\begin{array}{cccc|cc}
0 & & & & & \\
0 & 0 & & & & \\
0 & 0 & 0 & & & \\
1 & 0 & 0 & 0 & & \\
\hline
3 & 1 & 0 & 0 & 0 & \\
3 & 1 & 0 & 0 & 0 & 0
\end{array}\right]}
$$

$$P_4 \;=\; P^3\,P \;=\; \left[\begin{array}{cccc|cc} 0 & & & & & \\ 0 & 0 & & & & \\ 0 & 0 & 0 & & & \\ 1 & 0 & 0 & 0 & & \\ \hline 3 & 1 & 0 & 0 & 0 & \\ 3 & 1 & 0 & 0 & 0 & 0 \end{array}\right] \left[\begin{array}{cccc|cc} 0 & & & & & \\ 1 & 0 & & & & \\ 1 & 1 & 0 & & & \\ 0 & 1 & 1 & 0 & & \\ \hline 0 & 0 & 1 & 1 & 0 & \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right] = \left[\begin{array}{cccc|cc} 0 & & & & & \\ 0 & 0 & & & & \\ 0 & 0 & 0 & & & \\ 0 & 0 & 0 & 0 & & \\ \hline 1 & 0 & 0 & 0 & 0 & \\ 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

↓

great-great-grandparents

$$P^5 \;=\; P^4\,P \;=\; \left[\begin{array}{cccc|cc} 0 & & & & & \\ 0 & 0 & & & & \\ 0 & 0 & 0 & & & \\ 0 & 0 & 0 & 0 & & \\ \hline 1 & 0 & 0 & 0 & 0 & \\ 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \left[\begin{array}{cccc|cc} 0 & & & & & \\ 1 & 0 & & & & \\ 1 & 1 & 0 & & & \\ 0 & 1 & 1 & 0 & & \\ \hline 0 & 0 & 1 & 1 & 0 & \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right] \;=\; \{0\}$$

⇒ animals in the current generation have no known ancestors beyond their 4th ancestral generation.

$$\left(I - \frac{1}{2}P\right)^{-1} \;=\; \begin{bmatrix} 1 & & & & & \\ -\frac{1}{2} & 1 & & & & \\ -\frac{1}{2} & -\frac{1}{2} & 1 & & & \\ 0 & -\frac{1}{2} & -\frac{1}{2} & 1 & & \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & 1 & \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & 1 \end{bmatrix}^{-1}$$

$$\left(I-\frac{1}{2}P\right)^{-1} = \begin{bmatrix} 1.0 & & & & & \\ 0.5 & 1.0 & & & & \\ 0.75 & 0.5 & 1.0 & & & \\ 0.625 & 0.75 & 0.5 & 1.0 & & \\ 0.6875 & 0.625 & 0.75 & 0.5 & 1.0 & \\ 0.6875 & 0.625 & 0.75 & 0.5 & 0 & 1.0 \end{bmatrix}$$

Alternatively,

$$\left(I-\frac{1}{2}P\right)^{-1} = I+\frac{1}{2}P+\left(\frac{1}{2}P\right)^{2}+\left(\frac{1}{2}P\right)^{3}+\left(\frac{1}{2}P\right)^{4}$$

$$= \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix} + \begin{bmatrix} 0 & & & & & \\ 0.5 & 0 & & & & \\ 0.5 & 0.5 & 0 & & & \\ 0 & 0.5 & 0.5 & 0 & & \\ 0 & 0 & 0.5 & 0.5 & 0 & \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & & & & & \\ 0 & 0 & & & & \\ 0.25 & 0 & 0 & & & \\ 0.50 & 0.25 & 0 & 0 & & \\ 0.25 & 0.50 & 0.25 & 0 & 1 & \\ 0.25 & 0.50 & 0.25 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & & & & & \\ 0 & 0 & & & & \\ 0 & 0 & 0 & & & \\ 0.125 & 0 & 0 & 0 & & \\ 0.375 & 0.125 & 0 & 0 & 0 & \\ 0.375 & 0.125 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & & & & & \\ 0 & 0 & & & & \\ 0 & 0 & 0 & & & \\ 0 & 0 & 0 & 0 & & \\ 0.0625 & 0 & 0 & 0 & 0 & \\ 0.0625 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left(I - \frac{1}{2}P\right)^{-1} = \left[\begin{array}{cccccc} 1 & & & & & \\ 0.5 & 1 & & & & \\ 0.75 & 0.5 & 1 & & & \\ 0.625 & 0.75 & 0.5 & 1 & & \\ 0.6875 & 0.625 & 0.75 & 0.5 & 1 & \\ 0.6875 & 0.625 & 0.75 & 0.5 & 0 & 1 \end{array}\right]$$

Let us look at the sums $I + \sum_{c=1}^{j} \left(\frac{1}{2}P\right)^{c}$, $j \leq 4$, more closely.

$$I + \frac{1}{2}P = \left[\begin{array}{cccccc} 1.0 & & & & & \\ 0.5 & 1.0 & & & & \\ - & - & - & - & - & - \\ 0.5 & 0.5 & 1.0 & & & \\ 0 & 0.5 & 0.5 & 1.0 & & \\ 0 & 0 & 0.5 & 0.5 & 1.0 & \\ 0 & 0 & 0.5 & 0.5 & 0 & 1.0 \end{array}\right] \begin{array}{l} \text{parental rows } \varepsilon \text{ generation 1} \\ \\ \\ \\ \\ \\ \\ \end{array}$$

$$I + \sum_{c=1}^{2} \left(\frac{1}{2}P\right)^{c} = \left[\begin{array}{cccccc} 1.0 & & & & & \\ 0.5 & 1.0 & & & & \\ ---- & ---- & ---- & --- & --- & --- \\ 0.75 & 0.5 & 1.0 & & & \\ ---- & ---- & ---- & --- & --- & --- \\ 0.50 & 0.75 & 0.5 & 1.0 & & \\ 0.25 & 0.50 & 0.75 & 0.5 & 1.0 & \\ 0.25 & 0.50 & 0.75 & 0.5 & 0 & 1.0 \end{array}\right] \begin{array}{l} \\ \text{parental rows } \varepsilon \text{ generation 1} \\ \\ \text{parental row } \varepsilon \text{ generation 2} \\ \\ \\ \\ \\ \end{array}$$

$$I+\sum_{c=1}^{3}\left(\frac{1}{2}P\right)^{c} = \left[\begin{array}{cccccc} 1.0 &&&&& \\ 0.5 & 1.0 &&&& \\ \text{----} & \text{----} & \text{----} & \text{---} & \text{---} & \text{---} \\ 0.75 & 0.5 & 1.0 &&& \\ \text{----} & \text{----} & \text{----} & \text{---} & \text{---} & \text{---} \\ 0.625 & 0.75 & 0.5 & 1.0 && \\ \text{-----} & \text{-----} & \text{----} & \text{----} & \text{---} & \text{---} \\ 0.625 & 0.625 & 0.75 & 0.5 & 1.0 & \\ 0.625 & 0.625 & 0.75 & 0.5 & 0 & 1.0 \end{array}\right] \begin{array}{l} \\ \text{parental rows } \varepsilon \text{ generation 1} \\ \\ \text{parental row } \varepsilon \text{ generation 2} \\ \\ \text{parental row } \varepsilon \text{ generation 3} \\ \\ \\ \\ \end{array}$$

$$I+\sum_{c=1}^{4}\left(\frac{1}{2}P\right)^{c} = \left[\begin{array}{cccccc} 1.0 &&&&& \\ 0.5 & 1.0 &&&& \\ \text{------} & \text{-----} & \text{----} & \text{---} & \text{---} & \text{---} \\ 0.75 & 0.5 & 1.0 &&& \\ \text{------} & \text{-----} & \text{----} & \text{---} & \text{---} & \text{---} \\ 0.625 & 0.75 & 0.5 & 1.0 && \\ \text{------} & \text{-----} & \text{----} & \text{----} & \text{---} & \text{---} \\ 0.6875 & 0.625 & 0.75 & 0.5 & 1.0 & \\ 0.6875 & 0.625 & 0.75 & 0.5 & 0 & 1.0 \end{array}\right] \begin{array}{l} \\ \text{parental rows } \varepsilon \text{ generation 1} \\ \\ \text{parental row } \varepsilon \text{ generation 2} \\ \\ \text{parental row } \varepsilon \text{ generation 3} \\ \\ \\ \\ \end{array}$$

Generalizing:

(1) The i^{th} row of $T = \left(I-\frac{1}{2}P\right)^{-1}$ is equal to the sum of the i^{th} rows of $\mathbf{I}$, $\frac{1}{2}P$, $\left(\frac{1}{2}P\right)^{2}$, ..., $\left(\frac{1}{2}P\right)^{m}$, i.e., i^{th} row of T is equal to the sum of the i^{th} rows of $\mathbf{I}$, $\frac{1}{2}P$, $\left(\frac{1}{2}P\right)^{2}$, ..., $\left(\frac{1}{2}P\right)^{m}$.

However, the i^{th} row of T is also equal to the sum of the i^{th} rows of **I**, $\frac{1}{2}P$, $\left(\frac{1}{2}P\right)^2$, ..., $\left(\frac{1}{2}P\right)^{m_i}$, where m_i ($m_i \le m$) is the number of generations separating animal i from its oldest known ancestor.

(2) The parental rows of $I - \frac{1}{2}P$ are the same as the corresponding ones of

$$I + \frac{1}{2}P + \left(\frac{1}{2}P\right)^2,\ I + \frac{1}{2}P + \left(\frac{1}{2}P\right)^2 + \left(\frac{1}{2}P\right)^3,\ \ldots,\quad \text{and of}$$

$$I + \frac{1}{2}P + \left(\frac{1}{2}P\right)^2 + \left(\frac{1}{2}P\right)^3 + \ldots + \left(\frac{1}{2}P\right)^m \equiv T.$$

Similarly, the parental rows of $I + \frac{1}{2}P + \left(\frac{1}{2}P\right)^2$ are the same as the corresponding ones of $I + \frac{1}{2}P + \left(\frac{1}{2}P\right)^2 + \left(\frac{1}{2}P\right)^3$, ..., and **T**; etc. The reason for it is that the differences that exist between

$$I + \sum_{c=1}^{j\_1}\left(\frac{1}{2}P\right)^c \qquad \text{and} \qquad I + \sum_{c=1}^{j}\left(\frac{1}{2}P\right)^c$$

are related to accounting for the passage of alleles, from ancestors c generations removed form each animal, to these same individuals. For instance, if c = 3, the difference between $I + \sum_{c=1}^{2}\left(\frac{1}{2}P\right)^c$ and $I + \sum_{c=1}^{3}\left(\frac{1}{2}P\right)^c$, are the elements of $\left(\frac{1}{2}P\right)^3$ which reflect the passage of alleles from great-grandparents to great-grandprogeny. Thus, rows of animals with unknown ancestors from the c^{th} generation backwards remain unchanged, i.e., when the passage of alleles from all known ancestors of an animal has been explained, its row will not change anymore. These fixed rows are called parental rows if these animals have progeny. In particular, notice that:

$$\left\{ \begin{matrix} \text{the parental rows of} \\ \\ I + \sum_{c=1}^{m_{i-1}} \left(\frac{1}{2} P \right)^c \end{matrix} \right\} = \left\{ \begin{matrix} \text{the parental rows of} \\ \\ I + \sum_{c=1}^{m_i} \left(\frac{1}{2} P \right)^c \end{matrix} \right\}$$

because the animals from the last generation have not become parents yet.

(3) The i^{th} row of T is a linear function of the rows of the parents of animals i, i.e.,

$$\begin{aligned} i^{th} \text{ row of } \left\{ I + \sum_{c=1}^{m_i} \left(\frac{1}{2} P \right)^c \right\} &= i^{th} \text{ row of } \left\{ I + \frac{1}{2} P \left[I - \sum_{c=1}^{m_{i-1}} \left(\frac{1}{2} P \right)^c \right] \right\} \\ &= 1 \text{ on diag} + \frac{1}{2} p_i' \left[I + \sum_{c=1}^{m_{i-1}} \left(\frac{1}{2} P \right)^c \right] \end{aligned}$$

where

$p_i' = i^{th}$ row of P, it has at most two non-zero elements: a 1 on the column corresponding to the sire of animal i (s) and another 1 on the column of the dam of animal i (d), i.e.,

$$p_i' = [0 \ldots 0\underset{\substack{\downarrow \\ s}}{1}0 \ldots 0\underset{\substack{\downarrow \\ d}}{1}0 \ldots 0]$$

(4) Because T contains all the parental rows and the row for an animal is a linear function of the rows of its parents, a **recursive procedure to compute T** can be outlined as follows:

(i) Order animals such that parents precede progeny, numbering them from 1 (oldest) to n (youngest).

(ii) Compute the elements of $T = \{t_{ij}\}$, one row or one column at a time, as follows:

$$\text{(a) } t_{ij} = \frac{1}{2} t_{s_i,j} + \frac{1}{2} t_{d_i,j} \text{ for } j < i \qquad \text{if } s_i \text{ and } d_i \text{ are known,}$$

$$= \frac{1}{2} t_{s_i,j} \text{ for } j < i \qquad \text{if } s_i \text{ is known only,}$$

$$= \frac{1}{2} t_{d_i,j} \text{ for } j < i \qquad \text{if } d_i \text{ is known only,}$$

$$= 0 \text{ for } j < i \qquad \text{if neither } s_i \text{ nor } d_i \text{ is known,}$$

$$= 0 \text{ for } j > i$$

$$\text{(b) } t_{ii} = 1$$

The matrix G_{AX}, written in terms of T, is:

$$G_{AX} = T\,D\,T'$$

Because D is diagonal and positive, $D = D^{½}\,D^{½}$. Thus,

$$G_{AX} = T\,D^{½}\,D^{½}\,T'$$

$$G_{AX} = CC'$$

where

$$C = T\,D^{½}$$

$$C = \text{Cholesky decomposition of } G_{AX}$$

The elements of C can be computed recursively, using the procedure to compute T, as follows:

$$\text{(i) } c_{ij} = t_{ij}\, d_{jj}^{1/2}$$

$$= \frac{1}{2}\left(\delta_{s_i}\; t_{s_i,j} + \delta_{d_i}\; t_{d_i,j}\right) d_{jj}^{1/2} \text{ for } j < i$$

where

$$\delta_{s_i}\;\left(\delta_{d_i}\right) = 1 \quad \text{if } s_i\ (d_i) > 0$$

$$= \quad 0 \quad \text{if } s_i\ (d_i) = 0$$

(ii) $c_{ii} \; = \; d_{ii}^{1/2}$

$$c_{ii}^2 = d_{ii}$$

$$c_{ii}^2 = \left(\sigma^2_{AX}\right)_i - \frac{\delta_{s_i}}{4}\left(\left(\sigma^2_{AX}\right)_{s_i} + \sum_{c=1}^{n_{cs_i}} F_{cs_i}\left(\sigma^2_{AX}\right)_{cs_i}\right) - \frac{\delta_{d_i}}{4}\left(\left(\sigma^2_{AX}\right)_d + \sum_{c=1}^{n_{cd_i}} F_{cd_i}\left(\sigma^2_{AX}\right)_{cd_i}\right)$$

$$c_{ii}^2 = \left(\sigma^2_{AX}\right)_i - \frac{\delta_{s\,i}}{4} g_{s_is_i} - \frac{\delta_{d_i}}{4} g_{d_id_i}$$

But

$$G_{AX} = CC'$$

$$\Rightarrow \quad g_{s_is_i} = \sum_{k=1}^{s_i} c^2_{s_i,k}$$

$$g_{d_id_i} = \sum_{k=1}^{d_i} c^2_{d_i,k}$$

$$\Rightarrow \quad c^2_{ii} = \left(\sigma^2_{AX}\right)_i - \frac{\delta_{s_i}}{4}\sum_{k=1}^{s_i} c^2_{s_i,k} - \frac{\delta_{d_i}}{4}\sum_{k=1}^{d_i} c^2_{d_i,k}$$

⇒ to compute the $d_{ii} = c_{ii}^2$ we only need the squares of the diagonal elements of Cholesky matrix C. Also, computations can proceed **one column at a time**. Consequently, the matrix C does **not** need to be stored to compute the g_{ii}.

The **computational procedure to obtain the g_{ii}, proceeding one column at a time (Elzo, 1990)**, is as follows:

[1] Define:

u = vector of sums of squares of the elements of a row of C

v = vector containing the diagonal elements of C and work vector (used to store offdiagonal elements of C temporarily)

[2] Order animals so that parents precede progeny and number them from 1 to n. Set the numbers of unknown parents to zero.

[3] For the i^{th} round (i.e., the i^{th} animal) compute:

(a) $v_i = c_{ii}$

$$= \left[\left(\sigma^2_{AX}\right)_i - \frac{1}{4}\left(u_{s_i} + u_{d_i}\right)\right]^{1/2} \qquad \text{if } s_i, d_i > 0$$

$$= \left[\left(\sigma^2_{AX}\right)_i - \frac{1}{4}u_{s_i}\right]^{1/2} \qquad \text{if } s_i > 0, d_i = 0$$

$$= \left[\left(\sigma^2_{AX}\right)_i - \frac{1}{4}u_{d_i}\right]^{1/2} \qquad \text{if } s_i = 0, d_i > 0$$

$$= \left[\left(\sigma^2_{AX}\right)_i\right]^{1/2} \qquad \text{if } s_i = d_i = 0$$

(b) $v_j = c_{ji}$ for $j = i + 1, \ldots, n$

$$= \frac{1}{2}v_{s_j} + \frac{1}{2}v_{d_j} \qquad \text{if } i \le s_j, d_j$$

$$= \frac{1}{2}v_{s_j} \qquad \text{if } d_j < i \le s_j$$

$$= \frac{1}{2}v_{d_j} \qquad \text{if } s_j < i \le d_j$$

$$= 0 \qquad \text{if } s_j, d_j < i$$

(c) $u_j = u_j + (v_j)^2$ for $j = i, \ldots, n$

(d) $d_{ii}^{-1} = (v_i)^{-2}$

[4] Compute and sum the contributions of the i^{th} animal to G_{AX}^{-1} using the multibreed recursive rules. If the matrix is too big to be kept in core, use a linked-list subroutine to sum and store

only the non-zero elements of G_{AX}^{-1}.

[5] Repeat steps [3] and [4] until the last animal is processed, i.e., do steps [3] and [4] for i = 1, ... , n.

[6] If matrix G_{AX}^{-1} is to be stored on disk or type, copy the non-zero elements accompanied by their row and column numbers.

References

Falconer, D. S. 1981. Introduction to Quantitative Genetics. 2nd Ed., Longman, Inc., New York.

Henderson, C. R. 1976. A simple method for computing the inverse of a large numerator relationship matrix used in prediction of breeding values. Biometrics 32:69-83.

Elzo, M. A. 1990. Recursive procedures to compute the inverse of the multiple trait additive genetic covariance matrix in inbred and noninbred multibreed populations. J. Anim. Sci. 68:1215-1228.

Elzo, M. A. 1996. Animal Breeding Notes. University of Florida, Gainesville, Florida, USA.

Quaas, R. L. 1975. From Mendel's laws to the A inverse. Mimeograph, Cornell University, p 16.

Quaas, R. L. 1976. Computing the diagonal elements and inverse of a large numerator relationship matrix. Biometrics 32:949-953.

Quaas, R. L. 1986. Personal Communication. Animal Science 720. Cornell University, Ithaca, NY.