ANIMAL BREEDING NOTES

CHAPTER 13M

MULTIBREED VARIANCES AND COVARIANCES

Definitions

Multibreed population: a population composed of straightbred and crossbred animals that interbreed (Elzo, 1983; 1990a; 1994).

Breed group: a group of animals whose genetic composition falls within a range of fractions of breeds. For example, if five breed groups are constructed to group animals in a two-breed multibreed population (A = breed 1, B = breed 2), the group ranges could be: group 1 = (1.0 to .81)A (.0 to .19)B, group 2 = (.80 to .61)A (.20 to .39)B, group 3 = (.60 to .41)A (.40 to .59)B, group 4 = (.40 to .21)A (.60 to .79)B, and group 5 = (.0 to .19)A (1.0 to .81)B.

Regression model: a model that defines multibreed nonadditive effects in terms of intra- and interbreed interactions between alleles at K loci, K = arbitrary.

Additive intrabreed genetic covariance: a covariance due to additive genetic effects within a breed.

Additive interbreed genetic covariance: a covariance arising from differences between intrabreed means of additive genetic effects; it is equal to twice the segregation covariance (Lo et al., 1994).

Multibreed additive genetic covariance: an additive genetic covariance for animals in a multibreed population; equal to either an additive intrabreed genetic covariance (straightbred animals) or a weighted sum of additive intrabreed and interbreed genetic covariances (progeny of at least one crossbred parent).

Nonadditive configuration: a representation of K loci using the breed of origin of the alleles.

For example: if animals are composed of two breeds, A and B, there are four configurations at one locus: A/A, A/B, B/A and B/B; three configurations result if A/B and B/A are defined as one configuration. A possible set of configurations for one and two loci is shown in Elzo (1990b).

Nonadditive intraconfiguration genetic covariance: a covariance due to nonadditive genetic effects caused by the interaction between alleles from one or more breeds within a nonadditive configuration. Nonadditive configurations are to nonadditive genetic covariances as breeds are to additive genetic covariances.

Nonadditive interconfiguration genetic covariance: a covariance arising from differences between intraconfiguration means of nonadditive genetic effects.

Multibreed nonadditive genetic covariance: a nonadditive genetic covariance for animals in a multibreed population; equal to a weighted sum of nonadditive intraconfiguration and interconfiguration genetic covariances.

Environmental intrabreed covariance: a covariance due to environmental effects within a breed.

Environmental interbreed covariance: a covariance arising from differences between intrabreed means of environmental effects.

Environmental multibreed covariance: an environmental covariance equal to either an intrabreed environmental covariance (straightbred animals) or a weighted sum of intrabreed and interbreed environmental covariances (progeny of one or two crossbred parents).

Residual intrabreed, interbreed and multibreed covariances: weighted sums of additive and environmental intrabreed, interbreed and multibreed covariances.

Assumptions

1) Traits are determined by the combined effects of alleles at a large number of unlinked loci,

2) Random segregation and assortment of alleles occur during meiosis, and

3) Covariances remain constant over time.

Multibreed additive genetic covariances

Multibreed additive genetic covariances could potentially be different for each breed group combination. **Multibreed covariances can be computed as the sum of two terms**. The first term is equal to the weighted sum of the intrabreed covariances between traits X and Z, where the weights are the expected frequencies of each breed in the gth breed group combination (Elzo, 1983,1990a; Lo et al., 1994). The second term is equal to the weighted sum of the interbreed covariances between traits X and Z, where the weights are the sums of the products of the expected breed frequencies in the parental breed groups (Lo et al., 1994). Thus, the **multibreed additive genetic covariance** between traits X and Z for an animal in a noninbred multibreed population computed by conditioning on the breed of origin of alleles is:

$$\text{cov}_a(X,Z) = E[\text{cov}(X,Z \mid b] + \text{cov}(E[X \mid b], E[Z \mid b])$$

$$\text{cov}_a(X,Z) = \sum_{b=1}^{nb} p_b^i (\sigma_{aXZ})_b + \sum_{b=1}^{nb-1} \sum_{b'>b}^{nb} (p_b^s p_{b'}^s + p_b^d p_{b'}^d)(\sigma_{aXZ})_{bb'} \qquad [1]$$

Proof:

Let

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

$$u_{iZ} = a_{sZ} + a_{dZ}$$

where

u_{iX} , u_{iZ} = random variables representing the breeding values of individual i for traits X and Z

a_{sX} , a_{sZ} = random variables representing the sum of the average effects of all alleles from all breeds affecting traits X and Z coming from the gamete of the sire of animal i, e.g.,

$$a_{sX} = \left(\sum_{b=1}^{n_b} \sum_{k=1}^{2k_b} \delta_{bk} a_{bkX} \right)_{sire}$$, where a_{bkX} 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} , a_{dZ} = random variables representing the sum of the average effects of all alleles affecting

traits X and Z, coming from the gamete of the dam of animal i, e.g.,

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

To simplify notation, let

$$X = \sum_{b=1}^{n_b} \sum_{b'>b}^{2k_b} \delta_{bk} \, a_{bkX}$$

and

$$Z = \sum_{b=1}^{n_b} \sum_{b'>b}^{2k_b} \delta_{bk} \, a_{bkZ}$$

Thus, for the sire of animal i,

$$cov_a(X,Z)^s = E[cov(X,Z \mid b] + cov(E[X \mid b], E[Z \mid b])$$

$$E[cov(X,Z \mid b] = \sum_b p_b^s (\frac{1}{2} \sigma_{aXZ})_b$$

$$cov(E[X \mid b], E[Z \mid b])] = E[E[X \mid b][E[Z \mid b]] - (E[E[X \mid b]])(E[E[Z \mid b]])$$

$$E[E[X \mid b][E[Z \mid b]] = \sum_b p_b^s (E_{aX} E_{aZ})_{bb}$$

$$E[E[X \mid b][E[Z \mid b]] = E[E[XZ \mid b]]$$

$$E[E[XZ \mid b]] = \sum_b p_b^s (E_{aXZ})_{bb}$$

$$(E[E[X \mid b]])(E[E[Z \mid b]]) = (\sum_b p_b^s (E_{aX})_b)(\sum_b p_b^s (E_{aZ})_b)$$

$$(E[E[X \mid b]])(E[E[Z \mid b]]) = \sum_b (p_b^s)^2 (E_{aX} E_{aZ})_{bb} + \sum_b \sum_{b' \neq b} p_b^s p_{b'}^s (E_{aX} E_{aZ})_{bb'}$$

$$(E[E[X \mid b]])(E[E[Z \mid b]]) = \sum_b (p_b^s)^2 (E_{aXZ})_{bb} + \sum_b \sum_{b' \neq b} p_b^s p_{b'}^s (E_{aXZ})_{bb'}$$

Thus,

$$cov(E[X \mid b], E[Z \mid b])]$$

$$= \sum_b p_b^s (E_{aXZ})_{bb} - \sum_b (p_b^s)^2 (E_{aXZ})_{bb} - \sum_b \sum_{b' > b} p_b^s p_{b'}^s (E_{aXZ})_{bb'}$$

Premultiplying the first term by

$$\left(\sum_b p_b^s \right) = 1$$

yields

$$\left(\sum_b p_b^s \right) \left(\sum_b p_b^s (E_{aXZ})_{bb} \right) = \sum_b (p_b^s)^2 (E_{aXZ})_{bb} + \sum_b \sum_{b' > b} p_b^s p_{b'}^s (E_{aXZ})_{bb'}$$

Inserting the above expression in $cov(E[X \mid b], E[Z \mid b])$ yields

$$cov(E[X \mid b], E[Z \mid b])]$$

$$= \sum_b (p_b^s)^2 (E_{aX}E_{aZ})_b + \sum_b \sum_{b'>b} p_b^s p_{b'}^s (E_{aX}E_{aZ})_b (E_{aX}E_{aZ})_{b'}$$

$$- \sum_b (p_b^s)^2 (E_{aX}E_{aZ})_b - \sum_b \sum_{b'>b} p_b^s p_{b'}^s (E_{aX}E_{aZ})_b (E_{aX}E_{aZ})_{b'}$$

$$cov(E[X \mid b], E[Z \mid b])]$$

$$= \sum_b \sum_{b' \neq b} p_b^s p_{b'}^s (E_{aX}E_{aZ})_b - \sum_b \sum_{b'>b} p_b^s p_{b'}^s (E_{aX}E_{aZ})_b (E_{aX}E_{aZ})_{b'}$$

$$cov(E[X \mid b], E[Z \mid b])]$$

$$= \sum_b \sum_{b'>b} p_b^s p_{b'}^s (E_{aX}E_{aZ})_b - \sum_b \sum_{b'>b} p_b^s p_{b'}^s (E_{aX}E_{aZ})_b (E_{aX}E_{aZ})_{b'}$$

$$cov(E[X \mid b], E[Z \mid b])] = \sum_b \sum_{b'>b} p_b^s p_{b'}^s [(E_{aX}E_{aZ})_b - (E_{aX}E_{aZ})_{b'}]^2$$

$$cov(E[X \mid b], E[Z \mid b])] = \sum_b \sum_{b'>b} p_b^s p_{b'}^s (\sigma_{aXZ})_{bb'}$$

where $(\sigma_{aXZ})_{bb'}$ is the segregation covariance between traits X and Z due to differences in allele frequencies in breeds b and b'.

Thus,

$$cov_a(X, Z)^s = E[cov(X, Z \mid b] + cov(E[X \mid b], E[Z \mid b])$$

$$cov_a(X, Z)^s = \sum_b p_b^s (\frac{1}{2}\sigma_{aXZ})_b + \sum_b \sum_{b'>b} p_b^s p_{b'}^s (\sigma_{aXZ})_{bb'}$$

Similarly, for the dam of animal i,

$$cov_a(X,Z)^d = \sum_b p_b^d (\frac{1}{2}\sigma_{aXZ})_b \; + \; \sum_b \sum_{b'>b} p_b^d p_{b'}^d (\sigma_{aXZ})_{bb'}$$

Thus, the covariance between traits X and Z for animal i is:

$$cov_a(X,Z)^d = cov_a(X,Z)^s \; + \; cov_a(X,Z)^d$$

$$cov_a(X,Z)^d = \sum_b (p_b^s + p_b^d) \, (\frac{1}{2}\sigma_{aXZ})_b \; + \; \sum_b \sum_{b'>b} (p_b^s p_{b'}^s + p_b^d p_{b'}^d) \, (\sigma_{aXZ})_{bb'}$$

$$cov_a(X,Z)^d = \sum_b \frac{1}{2}(p_b^s + p_b^d) \, 2(\frac{1}{2}\sigma_{aXZ})_b \; + \; \sum_b \sum_{b'>b} (p_b^s p_{b'}^s + p_b^d p_{b'}^d) \, (\sigma_{aXZ})_{bb'}$$

$$cov_a(X,Z)^d = \sum_b p_b^i \, (\sigma_{aXZ})_b \; + \; \sum_b \sum_{b'>b} (p_b^s p_{b'}^s + p_b^d p_{b'}^d) \, (\sigma_{aXZ})_{bb'}$$

where the superscripts i, s and d correspond to an individual animal, its sire and its dam, the

subscripts b and b′ represent two breeds, and

nb	=	number of breeds,
p_b^x	=	expected fraction of breed b in animal x, x = i, s, d,
$(\sigma_{aXZ})_b$	=	additive intrabreed covariance for breed b, and
$(\sigma_{aXZ})_{bb'}$	=	additive interbreed covariance for the pair of breeds b and b′.

Multibreed nonadditive genetic covariances

Nonadditive genetic effects can be modeled in terms of subclass or regression procedures.

Regression multibreed nonadditive genetic covariances. These covariances involve

interactions between alleles from one or more breeds at k loci. Only n_k intraconfiguration

nonadditive genetic covariances between traits Y and Z at k loci, i.e., $(\sigma_{nYZ})_k$, k = 1, ... , n_k, need

to be estimated. No interconfiguration nonadditive genetic covariances are needed. For the case

of two breeds and assuming three configurations at 1 locus, there would be three

intraconfiguration variances: two intrabreed: var(A/A) and var(B/B), and one interbreed:

var(A/B).

Subclass multibreed nonadditive genetic covariances. These covariances are assumed to

differ for each breed-group-of-sire × breed-group-of-dam combination. They are computed by

conditioning them on the intraconfiguration populations of nonadditive genetic effects (as

multibreed additive covariances are conditioned on base breeds). Thus, the **subclass multibreed**

nonadditive genetic covariances between traits X and Z at k loci, are computed using the

following expression:

$$\text{cov}_n(X, Z) = E[\text{cov}(X, Z \mid b] + \text{cov}(E[X \mid b], E[Z \mid b])$$

$$\text{cov}_n(X, Z)_T = \left(\sum_{i=1}^{n_i} t_{ik} (\sigma_{nXZ})_{ik} + \sum_{i=1}^{n_{ik}-1} \sum_{i'>i}^{nk} t_{ik} t_{ik'} (\sigma_{nXZ})_{ik,i'k'} \right)_T \qquad [2]$$

where t_{ik} is the probability of intraconfiguration i at k loci in breed-group-of-sire × breed-group-

of-dam combination T, $(\sigma_{nXZ})_{ik}$ = intraconfiguration covariance i between traits X and Z,

$(\sigma_{nXZ})_{ik,i'k'}$ = interconfiguration covariance ik,i'k' at k loci between traits X and Z, and n_{ik} =

number of intraconfigurations at k loci. For example, assuming 2 breeds and 1 locus there would

be 2 intrabreed (A/A and B/B) and 1 interbreed (A/B) configurations. The probabilities of the 2

intrabreed configurations would be $t_{i1} = [p_A{}^s p_A{}^d]$ and $t_{i2} = [p_B{}^s p_B{}^d]$, and the probability of the

interbreed configuration would be $t_{i3} = [p_A{}^s p_B{}^d + p_B{}^s p_A{}^d]$. Notice that the terms

intraconfiguration and interconfiguration when referring to multibreed nonadditive covariances are equivalent to intrabreed and interbreed when referring to multibreed additive covariances.

Multibreed Environmental and Residual Covariances

Multibreed environmental covariances could be assumed 1) to be equal for all breeds and crossbred groups, 2) to be different for each breed and crossbred group (i.e., given an environment each genotype reacts differently), and 3) something in between alternatives 1 and 2. If **multibreed environmental covariances** were assumed to be different across breed groups, and to behave in an additive fashion, then their computation would be similar to the procedure used to compute additive genetic covariances. Thus, the multibreed environmental covariance between traits Y and Z, would be:

$$\text{cov}_e(Y,Z) = \sum_{b=1}^{nb} p_b^i (\sigma_{eYZ})_b + \sum_{b=1}^{nb-1} \sum_{b'>b}^{nb} (p_b^s p_{b'}^s + p_b^d p_{b'}^d)\ (\sigma_{eYZ})_{bb'} \qquad [3]$$

where the superscript i represents an individual animal, the subscripts b and b′ represent two breeds and

$(\sigma_{eYZ})_b$ =   environmental intrabreed covariance for breed b, and

$(\sigma_{eYZ})_{bb'}$ =   environmental interbreed covariance for pair of breeds b and b′.

The structure of **multibreed residual covariances** will depend on 1) the additive model used (animal, reduced animal, sire-dam, bull, sire model), 2) the ancestors identified on an animal with records, and 3) the assumptions made with respect to multibreed environmental covariances. For example, the expression of the multibreed residual covariance between traits Y and Z for a **sire-maternal grandsire model** is the following:

$$\text{cov}_v(Y,Z) = \text{cov}_a(Y,Z)^i - \delta^s\ .25\,\text{cov}_a(Y,Z)^s - \delta^{mgs}\ .0625\,\text{cov}_a(Y,Z)^{mgs}$$

$$+ (1 - \delta^{mgs}) \text{cov}_n(Y,Z)^{mgs} + \text{cov}_e(Y,Z)^i \qquad [3]$$

where superscripts i, s, and mgs refer to an animal, its sire, and its maternal grandsire, the

subscripts v, a, and e represent residual, additive genetic, and environmental, and

δ^x = indicator equal to 1 if animal x is not identified and to 0 if animal x is

identified, where x = s, mgs,

$\text{cov}_a(Y,Z)^i$ = $\text{cov}_a(Y_D,Z_D)^i$, where the subscript D = direct genetic effects.

$\text{cov}_a(Y,Z)^s$ = $\text{cov}_a(Y_D,Z_D)^s$, and

$\text{cov}_a(Y,Z)^{mgs}$ = $\text{cov}_a(Y_D,Z_D)^{mgs}$.

Multibreed additive genetic covariances in Equation [3] are computed using Equation [1], and

multibreed environmental covariances using Equation [2]. If a model includes sires and dams,

then additive dam covariances (and multiplying factors δ^d and .25) will be substituted for those

of the maternal grandsire. For an animal model with all relatives known, multibreed residual

covariances will contain environmental covariance components and covariances due to

nonadditive multibreed effects not accounted for sire × breed group of dam interactions.

References

Elzo, M. A. 1983. Multibreed sire evaluation within and across countries. PhD Dissertation,

University of California, Davis.

Elzo, M. A. 1990a. Recursive procedures to compute the inverse of the multiple trait additive

genetic covariance matrix in inbred and noninbred multibreed populations. J. Anim. Sci.

68:1215-1228.

Elzo, M. A. 1990b. Covariances among sire × breed group of dam interaction effects in

multibreed sire evaluation procedures. J. Anim. Sci. 68:4079-4099.

Mauricio A. Elzo, University of Florida, 2005, 2006, 2007, 2010, 2014.   [13M-11]

Elzo, M. A. 1994. Restricted maximum likelihood estimation of additive and nonadditive genetic

    variances and covariances in multibreed populations. J. Anim. Sci. 72:3055-3065.

Lo, L. L., R. L. Fernando, and M. Grossman. 1994. Genotypic covariance between relatives in

    multibreed populations. Theor. Appl. Genet. 87:423-430.