METHODS TO IMPLEMENT DOMINANCE GENETIC EFFECTS INTO PREDICTION EQUATIONS

Asko Mäki-Tanila[1] and Stephen P. Smith[2]
[1]Department of Animal Breeding, Agricultural Research Centre
31600 Jokioinen, FINLAND
and
[2]AGBU, University of New England
Armidale NSW 2351, AUSTRALIA

SUMMARY

Predictions in quantitative genetics have mostly been restricted to additive genetic effects. The flexibility of Henderson's mixed model methodology allows for straightforward consideration of dominance effects if there is no inbreeding. The proper treatment of inbreeding and dominance involves three parameters beside additive and dominance genetic variance. Tabular methods can be used for calculating genotypic covariances. The algorithm involves writing out animals in terms of pairs of gametes and, if necessary, considering also new combinations of gametes. The matrix is called an extended genomic table. The inversion of this matrix requires inverting several block diagonal matrices corresponding to a particular gamete.

INTRODUCTION

With the advance in computer technology and statistical theory, all the breeding value predictions are now dealt within the framework of Henderson's (1973) mixed model equations. It provides animal breeders with logical machinery and flexible tools for extending the scope of prediction beyond conventional additive genetic models to include non-additive intra- and interlocus effects (Henderson, 1985) or even extranuclear genetic effects (Kennedy and Schaeffer, 1989).

Most of the heterosis seeked for can be explained by dominance genetic effects. The most common and possibly the most practical way to handle the gains due to dominance is via minimizing the level of inbreeding, i.e. covering the effects of unfavourable recessive alleles by more desirable dominant ones. In within-line breeding work this is done by avoiding inbreeding. Lines highly performing in some traits but suffering from inbreeding depression in other traits, can be utilized by crossing them with other valuable lines. When breeds or lines show only small amounts of heterosis or when most emphasis is in within-breed selection, more sophisticated prediction methodology could utilize the accrued knowledge and technology on dominance genetic variation. The present paper will outline some of the developments in this area with most emphasis on techniques related to mixed model methodology.

NONINBRED POPULATIONS

In situations when dominance is regarded as a nuisance effect, it can be handled like any other factor causing detectable co-variation between observations. When there is no inbreeding, the removal of effects due to dominance can be easily built into the statistical model and into the mixed model equations (Henderson, 1985). However, there is no easy algorithm to

form the inverse of dominance relationship matrix (D). It can here only be treated as a special case of the more general instructions for D^{-1} in populations which may also be inbred (Smith and Mäki-Tanila, 1990). Another route is to first form D by simple rules and use a matrix inversion routine on it (Mäki-Tanila and Kennedy, 1986).

INBRED POPULATIONS

When we contemplate serious utilization of dominance variation, we do it almost by definition in cases involving considerable inbreeding. With the dominance model the simple linear relationship between performance and level of inbreeding springs into mind. This has been used also in the prediction context (e.g. Hudson and Van Vleck, 1984) and would apparently gain more popularity along the availability of faster routines to compute inbreeding levels (e.g. Tier, 1990). Linear predictions are the rule with additive models but selection on dominance effects relies on choosing between matings rather than individuals (c.f. Jansen and Wilton, 1985).

For more accurate predictions the second moments are required. If we stay with the conventional partition of genotypic values in random mating population into the average effect (a) of an allele (gamete) and into the dominance deviation (d) of a genotype (a pair of gametes), five genetic parameters are necessary to fully describe the co-variation between individuals (Gillois, 1964; Harris, 1964). These parameters define the first and second moments - including additive (σ^2_a) and dominance variance (σ^2_d). All these parameters are functions of allele frequencies in a way familiar from the heritability concept.

The accommodation of inbreeding requires parameters for the complete inbreeding depression (u_δ), the dominance variance amongst homozygotes (σ^2_δ), and the covariance between additive and dominance effects amongst homozygotes ($\sigma_{a\delta}$). In some cases it is convenient to work with parameter $\delta^2 = \sigma^2_\delta + u^2_\delta$. Corresponding identity coefficients can be devised and used in evaluating covariances amongst inbred relatives. This can, however, be avoided by resorting to mendelism and tabular methods at the first and second moment level (Smith and Mäki-Tanila, 1990).

For each genotypic effect $g_{ij} = a_i + a_j + d_{ij}$ (or g_{km}) where i (k) refers to paternal and j (m) to maternal gamete, the first and second moments can be formed. We use a notation where $i \leq j$ indicates that gamete j may be a descendant of gamete i, and arrange so that $i \leq k$, $k \leq m$. The parental gametes of j (m) are x and y (z and w).

One approach is to deal with additive and dominance effects associated with gametes, that is to construct an extended genomic table, notated by E (Smith and Allaire, 1985; Smith and Mäki-Tanila, 1990). The same principles may be applied on genotypic values, the resulting matrix is then a compression of E. The first moments can be obtained following the recursion rule

$$E(g_{ij}) = \begin{cases} \tfrac{1}{2} E(d_{ix}) + \tfrac{1}{2} E(d_{iy}) & i \neq j \\ \tfrac{1}{2} E(d_{ix}) + \tfrac{1}{2} E(d_{iy}) & i = j \end{cases}$$

We follow only through needed routes and use values already computed (c.f. Tier, 1990). Calculations are also terminated when base population gametes

are met, in which case

$$E(g_{ij}) = \begin{cases} 0 & i \neq j \\ u_\delta & i = j \end{cases}$$

For the covariances we need also the second moments for which the most general recursion rule is

$$E(g_{ij}g_{km}) = \tfrac{1}{2} E(g_{ij}g_{kz}) + \tfrac{1}{2} E(g_{ij}g_{kw}) \qquad k \neq m, j \neq m$$

The analogous counterpart for dominance effects was fully elaborated for dominance effects by Smith (1984). To complete the second moment rules for genotypic effects we write for the combinations of base gametes

$$E(g_{ij}g_{km}) = \begin{cases} 0 & \text{none of } i,j,k,l \text{ are equal} \\ u^2{}_\delta & i=j \neq k=m \\ \sigma^2{}_a & i \neq j; \ k \neq m; \ i=k \text{ or } j=k \text{ or } j=m \\ 2\sigma^2{}_a + \sigma_{a\delta} & \text{three of } i,j,k,m \text{ are equal} \\ 4\sigma^2{}_a + 4\,\sigma_{a\delta} + \delta^2 & i=j=k=m \end{cases}$$

In constructing these moments, some shortcuts may be used. For example, if we know about the gametic identities that $Prob(i \equiv j) = Prob(k \equiv m) = 0$, we have

$$E(d_{ij}d_{km}) = \sigma^2{}_d \ \sigma^{-4}{}_a \ [E(a_ia_k)E(a_ja_m) + E(a_ia_m)E(a_ja_k)]$$

implying that in some cases the prediction of dominance effects may be derived from gametic relationships (Mäki-Tanila and Kennedy, 1986). This algorithm is useful only if inbreeding is of moderate amount and would become meaningless with heavy inbreeding.

INVERSE OF (CO)VARIANCE MATRIX

If the proper (co)variance structures could be built into the mixed model equations, we would have a powerful technique to predict future progeny merits under a wide range of factors including selection and drift (e.g. Gianola and Fernando, 1986). For small data sets or simulation studies, a direct inversion may be possible but in any useful genetic improvement program the number of animals would soon make the computing impossible. Smith and Mäki-Tanila (1990) have proposed a method to model dominance and inbreeding via the mixed model and presented the formulae for inverting E.

Matrix E contains second moments and not (co)variances as required by the mixed model equations. However, deleting the rows and columns corresponding to first moments gives an inverse matrix of (co)variances. E is characterized by diagonal blocks associated with a particular gamete. Assuming that A_k contains the first $k+1$ blocks and given its inverse, we can evaluate the inverse of A_{k+1} where

$$A_{k+1} = \begin{bmatrix} A_k & A_kL_k \\ L_k'A_k & B_k \end{bmatrix}$$

B_k is block k+2 and L_k a simple matrix defined by column indices. The general inversion formula is

$$A_{k+1}^{-1} = \begin{bmatrix} A_k^{-1}+L_k(B_k-L_k'A_kL_k)^{-1}L_k' & -L_k(B_k-L_k'A_kL_k)^{-1} \\ -(B_k-L_k'A_kL_k)^{-1}L_k' & (B_k-L_k'A_kL_k)^{-1} \end{bmatrix}$$

To evaluate E^{-1} we apply this rule recursively starting witk k=0. It is hoped that $B_k-L_k'A_kL_k$ will be sufficiently small or sparse so that its inversion is feasible (e.g. Tier and Smith, 1990). In inverting the additive relationship matrix this element is always a scalar (Henderson, 1975).

APPLICATION

The methodology was tried on a selection experiment data involving egg-laying hens (2 386 females and 867 males) (Smith and Mäki-Tanila, 1989). Although the order of the E submatrix was 72 327, the matrix was very sparse there being only 10 775 off-diagonal non-zero elements. A_0 was of the order 20 179 and the distribution of size for the other diagonal blocks was such that most of them were of order 2 or 3 and none higher than 275.

The example study showed that E^{-1} can be evaluated for a real pedigree. Although it is difficult and misleading to generalize, it would seem that many practical data sets can be handled. The approach is so involved that there must be obvious sections to be improved. New research is needed in evaluating the method and important questions on the estimation of the five genetic parameters need answers.

REFERENCES

GILLOIS, M. 1964. La relation d'identité en génétique. Thése Fac. Sci. Paris. 295 pp.
GIANOLA, D. and FERNANDO, R.L. 1986. J. Anim. Sci. 63: 217-244.
HARRIS, D.L. 1964. Genetics 50: 1319-1348.
HENDERSON, C.R. 1973. Proc. Symp. in Honor of Dr. Jay L. Lush, pp. 10-41.
HENDERSON, C.R. 1975. J. Dairy Sci. 58: 1727-1730.
HENDERSON, C.R. 1985. J. Anim. Sci. 60: 111-117.
HUDSON, G.F.S. and VAN VLECK, L.D. 1984. J. Dairy Sci. 67: 171-179.
JANSEN, G.B. and WILTON, J.W. 1985. J. Dairy Sci. 68: 1302-1305.
KENNEDY, B.W.K. and SCHAEFFER, L.R. 1990. Advances in Statistical Methods for Genetic Improvement of Livestock (eds. D. Gianola and K. Hammond) (in press)
MÄKI-TANILA, A. and KENNEDY, B.W.K. 1986. Proc. 3rd World Cong. Genet. Appl. Livestock Prod. 12: 443-448.
SMITH, S.P. 1984. Dominance relationship matrix and inverse for an inbred population. Unpublished mimeo.
SMITH, S.P. and MÄKI-TANILA, A. 1989. Inverting the extended genomic table: a case study. Unpublished mimeo.
SMITH, S.P. and MÄKI-TANILA, A. 1990. Genet. Sel. Evol. 22 (1) (in press)
TIER, B. 1990. Computing inbreeding coefficients quickly. Genet. Sel. Evol. (in press)
TIER, B. and SMITH, S.P. 1990. Genet. Sel. Evol. 21: 457-466.