UNIVARIATE REML ANALYSES FOR MULTIVARIATE DATA WITH THE ANIMAL MODEL

R. Thompson\(^1\) and W.C. Hill\(^2\)

\(^1\)AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS, U.K.

\(^2\)Department of Genetics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN, U.K.

SUMMARY

In order to reduce computations involved in multivariate estimation of variance components, a sequence of univariate analyses is suggested based on analysing transformed traits and sums of pairs of transformed traits. The efficiency of the method is investigated.

INTRODUCTION

Use of the animal model (AM) in analysis of quantitative genetic data by Restricted Maximum Likelihood (REML) (Patterson and Thompson, 1971) gives both flexibility and efficiency in estimation of variance and covariance components. In order to handle complicated pedigree structures with unbalanced data it is convenient to use a derivative free approach to maximise the likelihood directly (Graser \textit{et al.}, 1987; Meyer, 1989), and package programs are available (Meyer, 1988). Whilst it is feasible in theory to handle any number \((P)\) of traits with any number \((C)\) of variance components per trait, or covariance components per pair of traits, in practice it is likely to be difficult to achieve convergence unless \(P\) is small. There are \(CQ\) components of (co)variance to estimate, where \(Q = P(P+1)/2\), and thus a \(CQ\) dimension hill to climb in the likelihood search, e.g. 12 dimensions for the simple case of 3 traits and 2 components.

Calculation of multivariate likelihoods in general take of the order of \(P^2\) as much space and \(P^3\) as much time as univariate calculations. Transformations to make some of the variance matrices diagonal can help to reduce calculations. It is natural, however, to consider undertaking \(Q\) univariate analyses each estimating \(C\) parameters and then deducing the required parameters from these. We indicate how to find the variance for such a procedure, find that use of canonical transformations can improve its efficiency, and give examples of its use.

METHOD

Consider multivariate data with

\[
y_a = X_b + \sum_{c=1}^{C} Z_c u_{ac} + a - 1, \ldots, P (1)
\]

484
for trait a, with \( u_{ac} \) a vector of random effects with \( \text{cov}(u_{ac}, u_{bd}) = I \psi_{cab} \) and \( \text{cov}(u_{ac}, u_{bd}) = 0 \) if \( c \neq d \). Typically \( Z_c \) will be the identity matrix and \( u_{ac} \) represents an environmental effect for the ath trait. Let \( \sigma_{ab} \) be the vector of \( C \) covariance components between traits a and b with cth element \( \psi_{cab} \) and \( \psi_c \) be the symmetric variance matrices with elements \( \psi_{cab} \). The procedure is to set up \( Q \) new traits as linear combinations of the \( P \) traits. Let \( z_a = \Sigma h_{ab} y_b \) or \( z = Hy \). Let the variances and covariances of the transformed traits be defined similarly to \( y \) but in terms of \( \psi^* \) and \( \sigma^* \). The variance \( \psi_{cua}^* \) is given by

\[
\psi_{cua}^* = \Sigma h_{ab} h_{ad} \psi_{cbd} \\
\psi^* = C \psi
\]

where \( \psi^* \) is a \( CQ \) vector of \( \psi_{cua}^* (c = 1, \ldots, C; a = 1, \ldots, Q) \), \( \psi \) is a \( CQ \) vector of \( \psi_{cab} (c = 1, \ldots, C; a = 1, \ldots, P, b < a) \), and \( G \) defines the relation between the elements of \( \psi \) and \( \psi^* \) as in (2).

From univariate ML analyses of the \( Q \) traits, estimates of \( \psi^* \) and \( \psi = C^{-1} \psi^* \) can be found. The simplest case, with transformation matrix, \( H_0 \), is to use \( P \) traits, \( y_a \), and \( Q-P \) sums of pairs of traits, \( y_a + y_b \).

If the data structure is such that an orthogonal analysis of variance can be constructed, then the variance estimates, \( \psi \), can be found for any known transformation matrix. In this analysis of variance for each trait there are \( S \) strata mean squares \( M_{saa} \) with \( df_s \) degrees of freedom and expectation \( E(M_{saa}) = \Sigma K_{sc} \psi_{cua}^* = V_{saa} \), where \( K_{sc} \) is the coefficient of the cth variance component in \( M_{saa} \), \( s = 1, \ldots, S \). The components \( \sigma^*_{aa} \) satisfy equations of the form \( I_{aa} \sigma^*_{aa} = E_a \) with elements of \( I_{aa} \) and \( E_a \) defined by

\[
I_{cd} = \Sigma (K_{sc} V_{saa}^{-1} K_{sd} V_{saa}^{-1}) \frac{df_s}{2} \quad c, d = 1, \ldots, C \quad (3)
\]

\[
E_c = \Sigma (K_{sc} V_{saa}^{-1} M_{saa} V_{saa}^{-1}) \frac{df_s}{2} \quad c = 1, \ldots, C \quad (4)
\]

with asymptotic variance \( I_{aa}^{-1} \) (Thompson, 1976). Similarly covariances between estimates of \( \sigma^*_{aa} \) and \( \sigma^*_{bb} \) can be found from \( I_{aa}^{-1} F_{ab} I_{bb}^{-1} \) with the elements of \( F_{ab} \) given by

\[
F_{cd} = \Sigma (K_{sc} V_{saa}^{-2} V_{sab} V_{sbb} V_{saa}^{-2}) \frac{df_s}{2} \quad c, d = 1, \ldots, C \quad (5)
\]

where \( V_{sab} = \Sigma k_{sc} \psi_{cab} \) using \( \text{cov}(M_{saa}, M_{sbb}) = 2V_{sab}/df_s \); (5) reduces to (3) if \( a = b \). Using (2) to (5) the variances of \( \psi^* \) and \( \psi \) can be found.

Alternatively \( \psi \) can be found from maximising the likelihood of all the variates together. When an orthogonal analysis of variance can be constructed then estimation equations are matrix analogues of (3) and (4) (Thompson, 1976).

These two alternative procedures have different variances because in the separate analysis of \( Q \) traits the covariances between \( M_{saa} \) and \( M_{sbb} \) are ignored. These covariances can be reduced by using independent traits. One suggestion is to form an initial estimate \( \psi^{(1)} \) using \( H_0 y \) and using \( \psi^{(1)} \) to find a transformation, \( T \), to make the component matrices \( \psi \) more diagonal and then use as traits \( z = H_0 Ty \) to give an estimate \( \psi^{(2)} \). If \( C = 2 \) a canonical transformation can be found to diagonalise both variance component matrices. An obvious extension is to iterate this procedure,
using \( \psi^{(i)} \) to form \( T \) in the \( i \)th iterate and deriving estimates \( \psi^{(i+1)} \), continuing until \( \psi^{(i)} \) converge.

The variance of \( \psi^{(2)} \) has two parts: one, as before, based on thinking of \( H \) as known and a second part due to \( H \) depending on \( \psi^{(1)} \), the initial estimate. By expressing the estimate \( \psi^{(2)} \) as a function of \( \psi^{(1)} \) and the mean square matrices and using a Taylor-series expansion involving first
differentials of \( G, I_a \) and \( E_a \) with respect to \( \psi^{(1)} \), it was found that the first order approximation to the extra part of the variance is zero.

EXAMPLES

Analysis is illustrated in Table 1 for a limited set of data, on two
traits of 141 dairy heifers (see Persaud et al., 1990, for more details).
The procedure was (i) to scale the traits by approximate phenotypic
standard deviation; (ii) to undertake univariate analyses on the original
traits, \( y_1 \), \( y_2 \) and \( y_1+y_2 \) (cycle 1); (iii) from these compute the canonical
transformation \( T \) and new variates; (iv) undertake univariate analyses on
these transformed variates and their sum (cycle 2); (v) back transform to
reestimate the variances of the original \( y_1 \) (cycle 2); (vi) repeat (iii),
(iv), (v) (cycle 3). Following cycle 1 on the untransformed and cycle 2 on
the transformed variates, the retransformed variates analysed in cycle 3
were nearly uncorrelated and iteration was ceased.

Details of an analysis undertaken on 251 cows, with a total of 475
lactations, fitting three covariances, additive genetic \( (\psi_1) \), environment
common to all lactations \( (\psi_2) \), and to specific lactations \( (\psi_3) \),
transformation diagonalizing \( \psi_1 \) and \( \psi_2 + \psi_3 \) was used. Covariances of
transformed variates were for cycle 2: 0.03453, 0.00577, -0.01771 and
-0.01194 and for cycle 3: -0.00986, 0.01343, -0.00787 and 0.00556, for \( \psi_1 \),
\( \psi_2 \), \( \psi_3 \) and \( \psi_2+\psi_3 \), respectively. It was found that convergence was slower
than with two components and that \( \psi_2 \) and \( \psi_3 \) were individually not diagonal,
although the correlations in \( \psi_2 \) and \( \psi_3 \) were typically in the range \( \pm 0.1 \).
Calculations suggest that the estimates of \( \psi_1 \) are at least 97% efficient.

DISCUSSION

Numerical results with a sire model with two traits and two components
suggest that there is a small loss of efficiency by using the suggested
analysis. For example, in a sire model with a residual mean square with 250
degrees of freedom estimating \( \psi_{2aa} \) and with 100 degrees of freedom between
sires, 93 with expectation \( \psi_{2aa} + 40\psi_{1aa} \) and 7 with expectation \( \psi_{2aa} +
12\psi_{1aa} \), and with uncorrelated traits of heritability 0.15 and 0.6, then the
estimates of genetic variances and covariances based on \( H_0 \) are 100% and 98%
efficient respectively. The covariance estimate becomes more efficient as
the family sizes and the heritabilities become more equal.

For three components we have shown in one specific example that a
canonical transformation based on \( \psi_1 \) and \( \psi_2 + \psi_3 \) has advantages, but it is
not clear if it is optimal to diagonalise these two linear combinations of
component matrices. Other possibilities can now easily be investigated.

ML methods are advocated partly because they take account of selection,
provided records on all traits and on culled animals are included.
Univariate analyses, albeit on uncorrelated traits, may produce biased
estimates; but perhaps these may be reduced by analysis of more linear
combinations, e.g. differences as well as sums.
The dimensionality of finding ML estimates can be reduced by partitioning the parameters into two sets, a $P \times P$ transformation matrix and a set of $CQ-P^2$ other parameters. For a given value of these other parameters it is relatively simple to find the optimal transformation matrix (Meyer, 1990; Juga and Thompson, 1990). A synthesis of this method with the one suggested here could be useful.

ACKNOWLEDGEMENTS

We are grateful to P. Persaud for data and to the editors for allowing a late submission.

REFERENCES


Table 1. Example of estimates of variances and covariances by univariate analysis (Meyer, 1988) of traits and sums of traits from successive iterates using original data (cycle 1) and transformed data (cycles 2 and 3). Data are on heifers at Langhill Farm. Trait 1 is yield of fat plus protein and trait 2 is live weight at calving, both scaled by approximate phenotypic standard deviation, and trait 3 the sum of 1 and 2.

<table>
<thead>
<tr>
<th>Variates</th>
<th>Original</th>
<th>Transformed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Re components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{211}$</td>
<td>0.50503</td>
<td>0.25331</td>
</tr>
<tr>
<td>$\psi_{222}$</td>
<td>0.20405</td>
<td>0.20761</td>
</tr>
<tr>
<td>$\psi_{212}$</td>
<td>0.15343$^+$</td>
<td>0.09776</td>
</tr>
<tr>
<td>$\psi_{233}$</td>
<td>1.01594</td>
<td></td>
</tr>
<tr>
<td>Residual components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{111}$</td>
<td>0.33889</td>
<td>0.17827</td>
</tr>
<tr>
<td>$\psi_{122}$</td>
<td>0.88423</td>
<td>0.88015</td>
</tr>
<tr>
<td>$\psi_{112}$</td>
<td>-0.07646$^+$</td>
<td>-0.03986</td>
</tr>
<tr>
<td>$\psi_{133}$</td>
<td>1.07020</td>
<td></td>
</tr>
</tbody>
</table>

+ Estimated from $(\psi_{c33} - \psi_{c11} - \psi_{c22})/2$. 

487