Estimation of quantitative genetic parameters

Robin Thompson
A.R.C. UNIT OF STATISTICS
THE KING'S BUILDINGS, EDINBURGH EH9 3JZ, SCOTLAND

1. INTRODUCTION

The genetic parameters that will be considered are genetic variances and covariances. Standard references (for instance Cockerham [1963], Falconer [1960] and Kempthorne [1957]) have discussed the common methods and problems associated with estimating these parameters. Typically some system of mating is used to generate sets of relatives raised in one or more environments. Often an analysis of variance (for collateral relatives) or covariance (for non-collateral relatives) based on the mating and environmental design can be easily constructed. The resulting variance and covariance components are usually easily interpreted in terms of covariances between relatives. These covariances between relatives can be also interpreted in terms of genetic and environmental components and hence estimates of genetic variance can be derived. The key role of the analysis of variance is not

Proceedings of the International Conference on QUANTITATIVE GENETICS

surprising since it is a way of partitioning variance, and genetic variances arose out of partitioning phenotypic variance (Fisher [1918]).

In many cases (in fact most of those discussed in standard texts) this partition of variance is enough to make estimation simple and efficient. However there are cases when this is not so and in this paper we discuss maximum likelihood (ML) methods in some of these cases. There are two main cases. One case is the balanced designs where a partition of variance is possible but there are more covariances between relatives than parameters to estimate and hence for some parameters more than one estimate can be derived. In Section 2 ML estimation is discussed and in particular a simple estimation procedure very similar to weighted least squares is given.

The other case considered is the unbalanced designs, which can occur, for example, when it is impossible to raise families of equal size. In this situation many estimation procedures for variance components have been suggested (Searle [1971]). These are usually based on analogies with the analysis of variance for balanced data. In the past, ML estimation has been rarely attempted primarily because of the computational difficulties. It is argued in Section 3 that, in some cases at least, the computational difficulties are no worse than in many ad hoc schemes and that the terms in the ML estimating equations are often useful in animal breeding studies.

In Section 4 we discuss the modifications needed when animals used as parents are selected on their phenotypic performance. In Section 5 we discuss the case when data are available from more than 2 discrete generations.

2. ESTIMATION FROM BALANCED DESIGNS

The case when a balanced design generates more covariances between relatives than parameters is now discussed. For example consider a hierarchical structure in which there are s sires,

d dams mated to each sire and n offspring raised from each dam, and data available on offspring and parents (Hill and Nicholas [1974], Thompson [1976]). This design generates five variances and covariances between relatives, namely, covariances between full sibs, σ_{fs}, between half sibs, σ_{hs}, between father and offspring, σ_{fo}, between mother and offspring, σ_{mo}, and phenotypic variance σ_p^2. In heritability estimation these structural parameters are often interpreted in terms of three environmental and genetic parameters: phenotypic variance, additive genetic variance, σ_A^2, and σ_K^2, the part of the full sib covariance not due to additive variance, which therefore contains dominance and common environmental terms. The relationship between the two sets of parameters is given in Table I.

TABLE I

Covariances among Relatives in terms of σ_A^2, σ_K^2 and σ_p^2.

Covariance	σ_A^2	σ_K^2	σ_p^2
σ_{fs}	1/2	1	0
σ_{hs}	1/4	0	0
σ_{fo}	1/2	0	0
σ_{mo}	1/2	0	0
σ_p^2	0	0	1

We will use this model for illustration even though there are several assumptions about genetic relationships, for instance no epistasis and no maternal effects, that might not be appropriate. We see there are 5 covariances between relatives and 3 parameters to estimate. Hill and Nicholas [1974] have shown how the parent-offspring and half sib estimates can be combined by evaluating the variances and covariances of the estimates. This is tedious and difficult to generalize. We now know how the ML estimates can be conveniently calculated.

Suppose X_i, Y_{ij} and Z_{ijm} represent observations on sires, dams and offspring and these are normally distributed about means μ_X, μ_Y and μ_Z, with covariances between observations given in

Table I. A convenient way of summarizing the data is to calculate 3 sum of squares and products matrices representing variation within dams ($\underset{\sim}{S}_1$) between dams within sires ($\underset{\sim}{S}_2$) and between sires ($\underset{\sim}{S}_3$). We let $\bar{x}_{\cdot}$, $\bar{y}_{i\cdot}$ etc. denote means taken over the subscript replaced by a dot and

$$z_{ijm} = z_{ijm} - \bar{z}_{ij\cdot}, \; z_{ij\cdot} = \bar{z}_{ij\cdot} - \bar{z}_{i\cdot\cdot}, \; z_{i\cdot\cdot} = \bar{z}_{i\cdot\cdot} - \bar{z}_{\cdot\cdot\cdot},$$

$$y_{ij} = y_{ij} - \bar{y}_{i\cdot}, \; y_{i\cdot} = \bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}, \; x_i = x_i - \bar{x}_{\cdot}.$$

The sum of squares and products matrices are:

$$\underset{\sim}{S}_1 = \left(\sum_i \sum_j \sum_m z_{ijm}^2\right), \quad \underset{\sim}{S}_2 = \begin{pmatrix} n\sum_i\sum_j z_{ij\cdot}^2 & \sum_i\sum_j z_{ij\cdot}y_{ij} \\ \sum_i\sum_j z_{ij\cdot}y_{ij} & \sum_i\sum_j y_{ij}^2 \end{pmatrix},$$

$$\underset{\sim}{S}_3 = \begin{pmatrix} nd\sum_i z_{i\cdot\cdot}^2 & d\sum_i z_{i\cdot\cdot}y_{i\cdot} & \sum_i z_{i\cdot\cdot}x_i \\ d\sum_i z_{i\cdot\cdot}y_{i\cdot} & d\sum_i y_{i\cdot}^2 & \sum_i y_{i\cdot}x_i \\ \sum_i z_{i\cdot\cdot}x_i & \sum_i y_{i\cdot}x_i & \sum_i x_i^2 \end{pmatrix}. \tag{1}$$

We note that the z, y and x squared terms represent terms in the analysis of variance of offspring, dam and sire measurements respectively and the cross product terms represent terms in the analysis of covariance. The degrees of freedom associated with $\underset{\sim}{S}_1$, $\underset{\sim}{S}_2$ and $\underset{\sim}{S}_3$ are $\nu_1 = sd(n-1)$, $\nu_2 = s(d-1)$ and $\nu_3 = s-1$. The expected value of $\underset{\sim}{S}_h$ denoted by $\nu_h \underset{\sim}{V}_h$ can be written in terms of the covariances between relatives. We find

$$\underset{\sim}{V}_1 = (\sigma_p^2 - \sigma_{fs}), \quad \underset{\sim}{V}_2 = \begin{pmatrix} \sigma_p^2 - \sigma_{fs} + n(\sigma_{fs} - \sigma_{hs}) & \sigma_{mo} \\ \sigma_{mo} & \sigma_p \end{pmatrix},$$

$$\underset{\sim}{V}_3 = \begin{pmatrix} (\sigma_p^2 - \sigma_{fs} + n(\sigma_{fs} - \sigma_{hs}) + nd\sigma_{hs} & \sigma_{mo} & \sigma_{fo} \\ \sigma_{mo} & \sigma_p^2 & 0 \\ \sigma_{fo} & 0 & \sigma_p^2 \end{pmatrix}. \tag{2}$$

The likelihood of all the data can be partitioned into two parts, one due to the fixed effects and one due to error contrasts i.e. contrasts with expectation independent of the fixed effects.

We use this latter log-likelihood, $\mathcal{L}$, to estimate the variance parameters arguing that in the absence of knowledge about the fixed effects the former provide no information about the variance parameters (Patterson and Thompson [1971]). In this example the $\mathcal{L}$ is equivalent to the log-likelihood of $\underset{\sim}{S}_1$, $\underset{\sim}{S}_2$ and $\underset{\sim}{S}_3$ and can be written as

$$\mathcal{L} = \text{const} - \frac{1}{2} \sum_{h=1}^{3} \nu_h \left[\log|\underset{\sim}{V}_h| + \text{tr}(\underset{\sim}{M}_h \underset{\sim}{V}_h^{-1})\right] \qquad (3)$$

where $\underset{\sim}{M}_h = \underset{\sim}{S}_h/\nu_h$. In order to differentiate (3) with respect to the parameters we express the $\underset{\sim}{V}$'s as a linear function of the parameters, i.e.

$$\underset{\sim}{V}_h = \underset{\sim}{X}_{hA}\,\sigma_A^2 + \underset{\sim}{X}_{hK}\sigma_K^2 + \underset{\sim}{X}_{hp}\sigma_p^2 \qquad (4)$$

where the $\underset{\sim}{X}$'s are known matrices. The matrices $\underset{\sim}{X}_{hi}$ (h=1,2,3; i=A,K,p) can be derived by replacing the covariances in $\underset{\sim}{V}_h$ by the corresponding coefficients for σ_i^2 in Table I, for instance $\underset{\sim}{X}_{1A} = (\frac{1}{2})$, $\underset{\sim}{X}_{1K} = (-1)$ and $\underset{\sim}{X}_{1p} = 1$. The values of σ_i^2 that maximize (3) satisfy

$$\frac{\partial \mathcal{L}}{\partial \sigma_i^2} = \sum_{h=1}^{3} \nu_k \,\text{tr}(\underset{\sim}{V}_h^{-1}\underset{\sim}{M}_h\underset{\sim}{V}_h^{-1}\underset{\sim}{X}_{hi}) - \sum_{h=1}^{3} \nu_K \,\text{tr}(\underset{\sim}{V}_h^{-1}\underset{\sim}{X}_{hi}) = 0 \quad . \qquad (5)$$

Usually (5) cannot be solved explicitly and an iterative solution is needed. One based on using the expected values of the second differentials that is very similar to weighted least squares is suggested by Anderson [1973]. In this scheme $\hat{\sigma}_i^2$ is estimated from

$$\sum_{j=1}^{3} \tilde{A}_{ij}\,\hat{\sigma}_j^2 = \tilde{B}_i \qquad (6)$$

where $\tilde{A}_{ij} = \sum_{h=1}^{3} \nu_h \,\text{tr}(\tilde{\underset{\sim}{V}}_h^{-1}\underset{\sim}{X}_{hi}\tilde{\underset{\sim}{V}}_h^{-1}\underset{\sim}{X}_{hj})$, $\tilde{B}_i = \sum_{h=1}^{3} \nu_h \,\text{tr}(\tilde{\underset{\sim}{V}}_h^{-1}\underset{\sim}{X}_{hi}\tilde{\underset{\sim}{V}}_h^{-1}\underset{\sim}{M}_h)$ (7)

and $\tilde{\underset{\sim}{V}}_h$ is an initial estimate of $\underset{\sim}{V}_h$. The procedure is repeated using $\hat{\sigma}_i^2$ to give $\tilde{\underset{\sim}{V}}_h$ (from (4)) until the estimates converge. The relationship with weighted least squares becomes apparent if we consider the linear model

$m_h = \sum_{i=1}^{q} x_{hi}\,\theta_i + e_h$ (h = 1, ..., H) with the e's uncorrelated,

with variances w_h. The weighted least squares estimates of θ_i satisfy

$$\sum_{j=1}^{q} A_{ij} \hat{\theta}_j = B_i \qquad (8)$$

where $A_{ij} = \sum_{h=1}^{H} w_h^{-1} x_{hi} x_{hj}$ and $B_i = \sum_{h=1}^{H} w_h^{-1} x_{hi} m_h$. (9)

Obviously the weight given in (6) to $\underset{\sim}{M}_h$ depends on ν_h and $\underset{\sim}{V}_h$. If the $\underset{\sim}{M}_h$ are scalars then the weights are inversely proportional to V_h^2/ν_h which is not surprising since then M_h has a χ^2 distribution with mean V_h and degrees of freedom ν_h. This procedure has been introduced using the hierarchical example but can be used whenever the data can be split into independent sum of squares and product matrices and their expectation is a linear function of variance parameters. Other analogies with least squares carry over; if $\underset{\sim}{A}$ is singular not all the parameters can be estimated, $2\hat{\underset{\sim}{A}}^{-1}$ gives the asymptotic variance-covariance matrix of the estimates which makes it relatively easy to compare alternative designs and the efficiency of the ML versus other estimation procedures.

Equations similar to (6) and (7) arise in other estimation procedures. For example they occur in minimum norm quadratic unbiased estimation (MINQUE) (Rao [1973]) if $\tilde{\underset{\sim}{V}}_h$ is chosen to correspond with the norm being minimized. Other methods (for instance Horn, Horn and Duncan [1975]) follow from replacing $\tilde{\underset{\sim}{V}}_h$ by $\underset{\sim}{V}_h$ in part of (7) and manipulating (6). Another possibility is to use weighted least squares on the elements of $\underset{\sim}{M}_h$ or the covariances between relatives (Hayman [1960]). This leads to the same estimates as in the ML procedure but needs the derivation of variances of and covariances between the elements of $\underset{\sim}{M}_h$ or the covariances between relatives. Although, in theory, these can be found, in practice the calculations can be intractable. Eisen [1967] suggested a design for estimating maternal genetic variances that generated 13 covariances between relatives and so $13 \times 14/2 = 91$ variances and covariances would be needed to implement the weighted least squares procedure.

3. ESTIMATION IN UNBALANCED DESIGNS

We discuss in this section ML estimation in unbalanced designs. In unbalanced designs sensible partitions of the data are not as obvious as in balanced designs. Often linear models for the observations (as opposed to linear models for the variance parameters as in Section 2) are introduced to generate appropriate partitions. A simple two factor model will be used to illustrate the main points. Extensions to more general models follow naturally but need matrix algebra to express the results compactly.

We assume a linear model of the form

$$y_{ki\ell} = \alpha_k + b_i + e_{ki\ell} \quad . \tag{10}$$

In sire evaluation this model is often used and then $y_{ki\ell}$ is the yield of the ℓ-th daughter of sire i in herd-season k, α_k is the effect of herd-season k and $e_{ki\ell}$ is a random variable normally distributed with mean zero and variance σ^2. If no other assumptions are made about α_k and b_i then α_k and b_i are called fixed effects and (10) is a fixed effects model. Alternatively if we assume the b_i are normally distributed with mean zero and variance σ_b^2 then the b_i are called random effects and (10) is then a mixed effects model. This mixed model implies that $\text{var}(y_{ki\ell}) = \sigma^2 + \sigma_b^2$ and $\text{cov}(y_{ki\ell}, y_{k'i'\ell'}) = \sigma_b^2$ if $i = i'$ and $= 0$ if $i \neq i'$. So the model can be written

$$y_{ki\ell} = \alpha_k + e'_{ki\ell} \tag{11}$$

where $e'_{ki\ell}$ has variance $\sigma^2 + \sigma_b^2$ and $\text{cov}(e'_{ki\ell}, e'_{k'i'\ell'}) = \sigma_b^2$ if $i = i'$ and 0 if $i \neq i'$. Often (11) is a covenient way of thinking about genetic models and helps in formulating the linear model (10). In the sire evaluation case if the covariances between daughters of a bull are assumed to be $\sigma_A^2/4$ and the variance of an observation is σ_p^2 this is consistent with a mixed model with $\sigma_b^2 = \sigma_A^2/4$ and $\sigma^2 = \sigma_p^2 - \sigma_A^2/4 = \sigma_p^2(1-h^2/4)$, where $h^2 = \sigma_A^2/\sigma_p^2$.

Searle [1971] has reviewed methods for estimating σ_b^2 and σ^2. Most follow the simple recipe of equating two sums of

squares to their expectations. One of the commonest (called the method of fitting constants or Henderson's method 3) is now outlined because the development is useful in understanding ML estimation. If α_k and b_i were fixed effects and were estimated by least squares they would satisfy

$$n_{kO}\,\hat{\alpha}_k + \sum_i n_{ki}\,\hat{b}_i = y_{kO} \quad , \tag{12}$$

$$\sum_k n_{ki}\,\hat{\alpha}_k + n_{Oi}\,\hat{b}_i = y_{Oi} \quad , \tag{13}$$

where n_{ki} is the number of daughters of sire i in herd-season k and O indicates summation over a suffix. An analysis of variance can be constructed:

Source	Sum of squares	
Herd-seasons	$\sum y_{kO}^2/n_{kO}$	
Sires (adjusted for Herd-seasons)	$\sum \hat{b}_i\, y_{Oi}$	(14)
Residual	$\sum y_{ki\ell}^2 - \sum \hat{\alpha}_k\, y_{kO} - \sum \hat{b}_i\, y_{Oi}$	(15)

The sires sum of squares is the difference between fitting a model with α_k and b_i and with α_k. The residual sum of squares is the sum of squares of deviations $(y_{kij} - \hat{\alpha}_k - \hat{b}_i)$. In the method of fitting constants (14) and (15) are equated to their expectation, which are functions of σ^2 and σ_b^2, and hence σ^2 and σ_b^2 can be estimated. The efficiency of this procedure is in general unknown and depends on the degree of unbalance and relative magnitude of σ^2/σ_b^2. However using this method with some unbalanced designs more precise estimates can be obtained if some of the data are removed (Swiger, Harvey, Everson and Gregory [1964]). Another misgiving I feel is in the ambivalence in the role of b_i. They are first assumed to be fixed effects to generate the sum of squares (14) and (15) and then assumed to be random effects to calculate their expectations. Further, if the mixed model is interpreted as (11) i.e. as a model for α_k with correlated errors it can be argued that weighted least squares and not least squares should be used to estimate α_k. Weighted least squares would usually require the inversion of a matrix of size the number of observations. However in mixed models of this

type Henderson (in Henderson, Kempthorne, Searle and Von Krosigk [1959]) has shown that this inversion can be eliminated and that the weighted least squares estimate of α_k satisfies

$$n_{kO}\,\tilde{\alpha}_k + \sum_i n_{ki}\,\tilde{\beta}_i = y_{kO} \quad , \tag{16}$$

$$\sum_k n_{ki}\,\tilde{\alpha}_k + (n_{Oi} + \gamma^{-1})\,\tilde{\beta}_i = y_{Oi} \quad , \tag{17}$$

where $\gamma = \sigma_b^2/\sigma^2$. These equations are very similar to (12) and (13) except that the coefficient n_{Oi} in (13) is replaced by $(n_{Oi} + \gamma^{-1})$ in (17). Henderson [1973] has emphasized that the $\tilde{\beta}_i$ can be interpreted as the predicted breeding values of bull i . In the sire evaluation case $\tilde{\beta}_i = [n_{Oi}\,h^2/(4 + (n_{Oi} - 1)h^2)]\,(y_{Oi} - \sum_k n_{ki}\,\tilde{\alpha}_k)/n_{Oi}$ i.e. the mean daughter-yield corrected for herd seasons is regressed back by a factor $[n_{Oi}\,h^2/(4 + (n_{Oi} - 1)h^2)]$.

Further $\tilde{\beta}_i$ plays a key part in the ML estimation of σ^2 and σ_b^2 . Patterson and Thompson [1971] have shown that the ML estimating equations are equivalent to equating $\Sigma y_{kij}^2 - \Sigma\tilde{\alpha}_k\,y_{kO} - \Sigma\tilde{\beta}_i\,y_{Oi}$ and $\Sigma\,\tilde{\beta}_i^2$ to their expected value. The first term is similar to a sum of squares of residuals (cf (15)), and the second term is the sum of squares of bull's predicted values. Once again an iterative scheme is usually needed to estimate σ^2 and σ_b^2 since (17) depends on σ_b^2/σ^2 . In practical cases I have found the iterative scheme outlined by Patterson and Thompson [1971] has converged in two or three iterations. The fitting constant method essentially gives equal weight to each observation, other methods give weights to the family means that are functions of their size. For the one-way classification i.e. only one herd-season $\Sigma\,\tilde{\beta}_i^2$ can be written as $\Sigma\,\gamma^2\,[\gamma + n_{Oi}^{-1}]^{-2}\,[(y_{Oi} - n_{Oi}\,\tilde{\alpha}_1)/n_{Oi}]^2$ and $[\gamma + n_{Oi}^{-1}]^{-2}$ is the weighting suggested by Robertson [1962].

The ML method can be extended to more complicated cases.

Thompson [1977] has considered the unbalanced version of the hierarchical design discussed in Section 2. Kempthorne and Tandon [1953] (for a single family classification) and Ollivier [1974] (for a hierarchical family classification) have suggested regression schemes weighting families according to size in order to make most use of the parent-offspring information. The ML method automatically does this and, where appropriate, uses the extra information from the sib covariances and enables fixed effects to be estimated.

The ML method can also be generalized to deal with q traits. In the two factor case we estimate $q \times q$ matrices $\underset{\sim}{\sigma}^2$ and $\underset{\sim}{\sigma}^2_b$. Useful equations are then (Thompson [1973])

$$n_{kO}\,\tilde{\alpha}_{km} + \sum_i n_{ki}\,\tilde{\beta}_{im} = y_{kOm} \qquad (18)$$

$$\sum_k n_{ki}\,\tilde{\alpha}_{km} + n_{Oi}\,\tilde{\beta}_{im} + \sum_{k=1}^{q} \gamma^{-1}_{mk}\,\tilde{\beta}_{ik} = y_{Oim} \qquad (19)$$

where $\underset{\sim}{\gamma}^{-1} = \underset{\sim}{\sigma}^2(\underset{\sim}{\sigma}^2_b)^{-1}$ and the suffix m represents the m-th trait $(m = 1, \ldots, q)$. Again $\tilde{\beta}_{im}$ can be interpreted as the predicted value for sire i for trait m. It is equivalent to combining the data on all q traits, corrected for herd-season effects, by means of a selection index to give the predicted value for the m-th trait. The ML estimating equations are natural extensions of the univariate equations. For instance, the sum of squares and products of the values $\tilde{\beta}_{im}$ are used. If $\underset{\sim}{\gamma}^{-1}$ is diagonal, then (18) and (19) separate into q parts each like (16) and (17) and there is no connection between the q variates.

If $\underset{\sim}{\gamma}^{-1}$ is not diagonal a canonical transformation of the variates enables the equations to be solved in q parts. The q new derived variates are

$$y^*_{kijr} = \sum_m T_{rm}\, y_{kijm} \qquad (m = 1, \ldots, q)$$

where $\underset{\sim}{T}$, the matrix of coefficients T_{rm} satisfies

$$\underset{\sim}{T}\,\underset{\sim}{\sigma}^2_b\,\underset{\sim}{T}' = \underset{\sim}{I} \quad \text{and} \quad \underset{\sim}{T}\,\underset{\sim}{\sigma}^2\,\underset{\sim}{T}' = \underset{\sim}{D}$$

where $\underset{\sim}{D}$ is a diagonal matrix. These canonical variates sometimes have a genetic interpretation in terms of which linear combination

of traits is most heritable (Rouvier [1969]). They might also be useful in interpreting results on the effects of errors in parameter estimates on the efficiency of selection indices (Harris [1964]).

4. SELECTION OF PARENTS

Sometimes, either through design or accident, the animals that are used as parents are chosen on their phenotypic performance. Then some of the usual methods of estimation are biased, for example heritability if estimated by sib covariances, genetic correlations if estimated by parent-offspring regression. In this section it is argued that these difficulties are removed if ML is used.

Suppose we have observations on $\nu_1 + \nu_2$ parents (y_i) and on ν_2 offspring (z_i). Suppose also y_{1i} and y_{2i} are normally distributed with mean zero and variance V_1 and V_{22} and covariances V_{12} between y_i and z_i and also that parents are chosen at random. Let S_1 be the sum of squares for the parental data and $\underset{\sim}{S}_2$ the sum of squares and cross products matrix for parent and offspring data. Let $\nu_2 \underset{\sim}{V}_2$ be the expected value of $\underset{\sim}{S}_2$, then $\underset{\sim}{V}_2$ and $\underset{\sim}{S}_2$ can be partitioned as

$$\underset{\sim}{V}_2 = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \qquad \underset{\sim}{S}_2 = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

where $V_{11} = V_1$ and $V_{21} = V_{12}$. The log-likelihood can be written as

$$\mathcal{L} = \text{const} - \frac{1}{2}(\nu_1 \log|V_1| + \text{tr}[(S_1 - S_{11})V_1^{-1}] + \nu_2 \log|V_2| + \text{tr}(\underset{\sim}{S}_2 \underset{\sim}{V}_2^{-1})) \quad . \tag{20}$$

$\mathcal{L}$ is of the same form as (3) and hence (6) can be used if V_1 and $\underset{\sim}{V}_2$ are linear functions of the unknown parameters. An alternative instructive form for (20) follows if we partition $\mathcal{L}$ into two independent parts, one part, $\mathcal{L}_1$, the log-likelihood of y_{1i}, the parental data and another part, $\mathcal{L}_2$ the log-likelihood of $z_i - V_{21}V_{11}^{-1}y_i$ which can be thought of as the offspring record given (or conditional on) the parental record. Defining $S_{22.}$ and

$V_{22.}$ as the sum of squares and the variance of $z_i - V_{21}V_{11}^{-1}y_i$, we find $\mathcal{L}_1$ and $\mathcal{L}_2$ can be written as

$$\mathcal{L}_1 = \text{const} - \tfrac{1}{2}((\nu_1+\nu_2)\log|V_1| + \text{tr}(S_1V_1^{-1})) , \qquad (21)$$

$$\mathcal{L}_2 = \text{const} - \tfrac{1}{2}(\nu_2\log|V_{22.}| + \text{tr}(S_{22.}V_{22.}^{-1})) . \qquad (22)$$

We see ML essentially makes use of three pieces of information. The parental data gives information on V_1, regression of z_i on y_i gives information on $V_{21}V_{11}^{-1}$ and $z_i - V_{21}V_{11}^{-1}y_i$ gives information on $V_{22.}$.

Suppose parents are chosen on their parental values, then following Kempthorne and Von Krosigk (in Henderson et al [1959]) and Curnow [1961] we can write the log-likelihood as the log-likelihood of parental values plus the log-likelihood of offspring values given parental values. This log-likelihood is again $\mathcal{L}_1+\mathcal{L}_2 = \mathcal{L}$ (20) and the iterative scheme in Section 2 can be used. One minor modification is needed, A_{ij}, depends on the expected values of the second moments $E(y_iz_i)$ and $E(z_i^2)$. Following Curnow [1961] we express these conditional on the selected parental values. Let $\nu_1M_{11} = S_{11}$, then noting that $z_i - V_{21}V_{11}^{-1}y_i$ has variance $V_{22.}$ and is independent of y_i we find

$$E(y_iz_i) = M_{11}V_{11}^{-1}V_{12} \qquad (23)$$

$$E(z_i^2) = V_{22.} + V_{21}V_{11}^{-1}M_{11}V_{11}^{-1}V_{12} = V_{22} - V_{21}V_{11}^{-1}(V_{11}-M_{11})V_{11}^{-1}V_{12} \qquad (24)$$

$$= V_{22} - (1-K)V_{21}V_{11}^{-1}V_{12} \quad \text{if} \quad M_{11} = KV_{11} .$$

Using (23) and (24) A_{ij} can be found to be

$$A_{ij} = \sum_{h=1}^{2}\nu_h\,\text{tr}(V_h^{-1}X_{hi}V_h^{-1}X_{hj}) - 2\,\text{tr}\,(V_1^{-1}(S_{11}-\nu_1V_1)V_1^{-1}X_{1i}V_1^{-1}X_{1j})$$

$$+ 2\,\text{tr}\left[\begin{pmatrix} V_1^{-1}(S_{11}-\nu_1V_1)V_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} X_{2i}V_2^{-1}X_{2j}\right] . \qquad (25)$$

Terms similar to (25) have been given by Curnow [1961] (for parent-offspring data) and Thompson [1973 and 1976] (for multivariate parent-offspring data and multivariate hierarchical

structures). Using (23) and (24) it can be checked that (5) gives unbiased estimating equations for θ_i. Note that in effect we are estimating the variances in the unselected population. Covariances and variances in the selected population could be evaluated using formulae similar to (23) and (24). Equations (23) and (24) and natural extensions of them have been used to investigate the effect of selection of parents on several common estimation procedures. They can be used to justify parent-offspring regression to estimate heritability (Falconer [1960]), to give measures of biases and to suggest correction factors for sib-covariance estimates of heritability and parent-offspring estimates of genetic correlation (Reeve [1953], Brown and Turner [1968]).

The formulae (20) - (25) have been written so that they hold if y_i and z_i are vectors. Obviously V_1, V_{12}, V_{22} etc. will then be interpreted as matrices of the appropriate size. Sometimes data are only available on the selected parents and the offspring. Then maximizing the conditional likelihood of the offspring given the parents, $\mathcal{L}_2$, seems an obvious suggestion. If $\mathcal{L}_2$ is written as $\mathcal{L} - \mathcal{L}_1$ it is of the form of (3). This way of writing the log-likelihood is similar to writing, in a non-orthogonal analysis of variance, the sum of squares for factor B after adjusting for factor A as the sum of squares for factors A and B minus the sum of squares for A (Searle [1971]). Henderson [1975] has discussed estimating fixed effects and predicting random effects from unbalanced designs using a similar conditional approach.

5. MORE THAN TWO GENERATIONS

In this section we give a convenient form for the covariances between relatives in different generations in terms of the additive genetic variance. We assume for simplicity that generations are discrete so that an individual in generation t is the offspring of parents in generation $(t-1)$ and that there are N_t individuals in the t-th generation.

In each generation we order the individuals by sex with males first. Suppose we start in generation 0 and if we assume the N_0 individuals are unrelated then, the coefficients of parentage of the individuals can be represented by a $N_0 \times N_0$ matrix, $\underset{\sim}{R}_{00}$, equal to $\frac{1}{2}\underset{\sim}{I}$. We now define matrices $\underset{\sim}{Z}_t$ of size $N_{t+1} \times N_t$ relating the individuals of generation (t+1) with those in generation t in order to calculate the other coefficients of parentage. The $\underset{\sim}{Z}_t$ matrices can be written as

$$\begin{pmatrix} \underset{\sim}{Z}_{tmm} & \underset{\sim}{Z}_{tmf} \\ \underset{\sim}{Z}_{tfm} & \underset{\sim}{Z}_{tff} \end{pmatrix} \tag{26}$$

where the elements of $\underset{\sim}{Z}_t$ are either $\frac{1}{2}$ or 0. The (j, k) element of $\underset{\sim}{Z}_{tmm}$ is $\frac{1}{2}$ only if the k-th male of generation t is the father of the j-th male in generation (t + 1). The other blocks of $\underset{\sim}{Z}$ are defined similarly. Hill [1974] used similar matrices and notes that the blocks of $\underset{\sim}{Z}$ represent the alternative pathways of genes

$$\left(\begin{array}{c|c} \text{males from males} & \text{males from females} \\ \hline \text{females from males} & \text{females from females} \end{array}\right).$$

The relationship matrix for the first three generations (which indicates the general form) can now be written, if no individual is inbred, as (Thompson [1977])

$$\underset{\sim}{R} = \frac{1}{2}\begin{pmatrix} \underset{\sim}{I}_0 & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{Z}_0 & \underset{\sim}{I}_1 & \underset{\sim}{0} \\ \underset{\sim}{Z}_1\underset{\sim}{Z}_0 & \underset{\sim}{Z}_1 & \underset{\sim}{I}_2 \end{pmatrix}\begin{pmatrix} \underset{\sim}{I}_0 & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \frac{1}{2}\underset{\sim}{I}_1 & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \frac{1}{2}\underset{\sim}{I}_2 \end{pmatrix}\begin{pmatrix} \underset{\sim}{I}_0 & \underset{\sim}{Z}_0' & \underset{\sim}{Z}_0'\underset{\sim}{Z}_1' \\ \underset{\sim}{0} & \underset{\sim}{I}_1 & \underset{\sim}{Z}_1' \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{I}_2 \end{pmatrix} \tag{27}$$

This is the product of a lower triangular matrix with a diagonal matrix and with an upper triangular matrix. The inverse of $\underset{\sim}{R}$, sometimes used for predicting random effects (Henderson [1976]) has a simple form since the left inverse of the lower triangular matrix is

$$\begin{pmatrix} \underset{\sim}{I}_0 & \underset{\sim}{0} & \underset{\sim}{0} \\ -\underset{\sim}{Z}_0 & \underset{\sim}{I}_1 & \underset{\sim}{0} \\ \underset{\sim}{0} & -\underset{\sim}{Z}_1 & \underset{\sim}{I}_2 \end{pmatrix}.$$

The variance matrix of the observations assuming just an additive genetic component and an environmental component is $\underset{\sim}{V} = 2\underset{\sim}{R}\,\sigma_A^2 + \underset{\sim}{I}\,\sigma_e^2$. Suppose the records in the t-th generation are $\underset{\sim}{y}_t$ and, for simplicity, these are normally distributed about a mean of 0 . ML estimation of σ_A^2 and σ_e^2 depends on calculating $\underset{\sim}{V}^{-1}$ and, except for some special cases (for example observations only available on one sex (Thompson [1977])) in most practical cases this is not feasible. One suggestion is to work with deviations from parental values $\underset{\sim}{y}^+_{t+1} = \underset{\sim}{y}_{t+1} - \underset{\sim}{Z}_t\,\underset{\sim}{y}_t$, since the variance matrix of these deviations is tridiagonal and the covariance between deviations two or more generations apart are zero. When each dam is mated to only one sire the variance structure for $\underset{\sim}{y}^+_t$ corresponds to a hierarchical analysis of variance with sire and dam components σ_e^2 and covariance between full sibs $\sigma_e^2 + \sigma_A^2/2$.

Another possibility is to work with $\underset{\sim}{y}^*_t$, the t-th generation values conditional on, or given, the ancestors' records, which can be interpreted as deviations of actual from predicted values. For example, $\underset{\sim}{y}^*_1 = \underset{\sim}{y}_1 - (\sigma_A^2/(\sigma_e^2+\sigma_A^2))\,\underset{\sim}{Z}_O\underset{\sim}{y}_O$, where $(\sigma_A^2/(\sigma_e^2+\sigma_A^2))\underset{\sim}{Z}_O\underset{\sim}{y}_O$ represents a vector of mid-parent values regressed back by a factor $(\sigma_A^2/(\sigma_e^2+\sigma_A^2))$ and hence are predicted values of $\underset{\sim}{y}_1$. The terms for the next generations are more complicated, but if we approximate the variance of $\underset{\sim}{y}^*_1$ by $v_1\,\underset{\sim}{I}$, where v_1 is the variance of the elements of $\underset{\sim}{y}^*_1$, we can approximate $\underset{\sim}{y}^*_2$ by $\underset{\sim}{y}_2 - \underset{\sim}{Z}_1\underset{\sim}{y}_1 + (\sigma_e^2/v_1)\underset{\sim}{Z}_1\underset{\sim}{y}^*_1$. Using the same type of approximation and a recursive argument similar to that of Bulmer [1971] in succeeding generations we find that $\underset{\sim}{y}^*_{t+1}$ might be approximated by

$$\underset{\sim}{y}^*_{t+1} = \underset{\sim}{y}_{t+1} - \underset{\sim}{Z}_t\underset{\sim}{y}_t + (\sigma_e^2/v_t)\underset{\sim}{Z}_t\underset{\sim}{y}^*_t \qquad (28)$$

where $v_t = \sigma_p^2 + d_t$, $d_{t+1} = \frac{1}{2}\,d_t - \frac{1}{2}H_t v_t^{-1} H_t$, $d_O = 0$, $\sigma_p^2 = \sigma_A^2 + \sigma_e^2$ and $H_i = \sigma_A^2 + d_i$. As in Bulmer's case the d_i quickly converge to a limiting value d^*

$$d^* = (-2\sigma_A^2 - \sigma_p^2 + [\sigma_p^4 + 4\sigma_A^2\ (\sigma_p^2 - \sigma_A^2)]^{\frac{1}{2}})/4 \ .$$

The values of d^*/σ_p^2 for various values of σ_A^2/σ_p^2 are given in Table II and we see approximately $d^*/\sigma_p^2 = -\frac{1}{2}(\sigma_A^2/\sigma_p^2)^2$. $\sigma_p^2 + d^*$ can be thought of as the variance between the actual value and the predicted value using all parental, grand parental etc. information.

Table II

Values of d^*/σ_p^2 for various values of σ_A^2/σ_p^2 .

σ_A^2/σ_p^2	0.0	0.2	0.4	0.6	0.8	1.0
d^*/σ_p^2	-0.000	-0.030	-0.100	-0.200	-0.330	-0.500

Sometimes the regression of response on selection differential is used to estimate heritability (Falconer [1960], Hill [1971, 1972]). This is similar to putting $d_t = 0$ in (28) and working with the mean values of y_t^* rather than the individual values. The variance-covariance matrix of the mean values can be derived by arguments similar to the development of (28) which I find more appealing than Hill's [1971, 1972] intuitive genetic approach.

BIBLIOGRAPHY

ANDERSON, T.W. [1973]. Asymptotically efficient estimation of covariance matrices with linear structure. Ann. Statist. 1, 135-141.

BROWN, G.H. and TURNER, H.N. [1968]. Response to selection in Australian Merino sheep. II Estimates of phenotypic and genetic parameters for some production traits on Merino ewes and an analysis of the possible effects of selection on them. Aust. J. of Agric. Research 19, 303-22 (Corrigendum, 21, 182).

BULMER, M.G. [1971]. The effect of selection on genetic variability. Amer. Natur. 105, 201-211.

COCKERHAM, C.C. [1963]. Estimation of genetic variances. Statistical Genetics and Plant Breeding. Nat. Acad. Sci. Nat. Res. Council Publ. 982, 53-94.

CURNOW, R.N. [1961]. The estimation of repeatability and heritability from records subject to culling. Biometrics 17, 553-66.

EISEN, S.J. [1967]. Mating designs for estimating direct and maternal variances and direct-maternal covariances. Can. J. Genet. Cytol. 9, 13-22.

FALCONER, D.S. [1960]. Introduction to Quantitative Genetics. Ronald Press Co., New York, N.Y.

FISHER, R.A. [1918]. The correlation between relatives on the supposition of Mendelian inheritance. Trans. Royal Soc., Edinburgh 52, 399-433.

HARRIS, D.L. [1964]. Expected and predicted progress from index selection involving estimates of population parameters. Biometrics 20, 46-72.

HAYMAN, B.I. [1960]. Maximum likelihood estimation of genetic components of variation. Biometrics 16, 369-381.

HENDERSON, C.R. [1973]. Sire evaluation and genetic trends. In Proc. Anim. Breed. Genet. Symp. Blacksburg, Virginia pp 10-41. American Society of Animal Sciences, Champaign, Illinois.

HENDERSON, C.R. [1975]. Best linear unbiased estimation and prediction under a selection model. Biometrics 31, 423-447.

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

HENDERSON, C.R., O. KEMPTHORNE, S.R. SEARLE and C.M. von KROSIGK. [1959]. The estimation of environmental and genetic trends from records subject to culling. Biometrics 15, 192-218.

HILL, W.G. [1971]. Design and efficiency of selection experiments for estimating genetic parameters. Biometrics 27, 293-311.

HILL, W.G. [1972]. Estimation of realized heritabilities from selection experiments. I Divergent selection. II Selection in one generation. Biometrics 29, 747-780.

HILL, W.G. [1974]. Prediction and evaluation of response to selection with overlapping generations. Anim. Prod. 18, 117-139.

HILL, W.G. and F.W. NICHOLAS. [1974]. Estimation of heritability by both regression of offspring on parent and intra-class correlation of sibs in one experiment. Biometrics 30, 447-468.

HORN, S.D., R.A. HORN and D.B. DUNCAN [1975]. Estimating variances in linear models. Journal of the American Statistical Association, 70, 380-385.

KEMPTHORNE, O. [1957]. An introduction to Genetic Statistics. John Wiley and Sons, New York, N.Y.

KEMPTHORNE, O. and O.B. TANDON. [1953]. The estimation of heritability by regression of offspring on parent. Biometrics 9, 90-100.

OLLIVIER, L. [1974]. La regression parent-descendant dans le cas de descendances subdivisees en familles de taille inegale. Biometrics 30, 59-66.

PATTERSON, H.D. and R. THOMPSON. [1971]. Recovery of inter-block information when block sizes are unequal. Biometrika 58, 545-554.

RAO, C.R. [1973]. Linear Statistical Inference and its Applications (Second Edition) Wiley, New York, N.Y.

REEVE, E.C.R. [1953]. Studies in quantitative inheritance. III Heritability and genetic correlation in progeny tests using different mating systems. J. Genet. 51, 520-542.

ROBERTSON, A. [1962]. Weighting in estimation of variance components. Biometrics 18, 413-415.

ROUVIER, R. [1969]. Ponderation des valeurs genotypiques dans la selection par indice sur plusieurs caracteres. Biometrics 25, 295-308.

SEARLE, S.R. [1971]. Linear Models. Wiley, New York, N.Y.

SWIGER, L.A., W.R. HARVEY, D.O. EVERSON and K.E. GREGORY [1969]. The variance of intraclass correlation involving groups with one observation. Biometrics, 20, 818-826.

THOMPSON, R. [1973]. The estimation of variance and covariance components with an application when records are subject to culling. Biometrics 29, 527-550.

THOMPSON, R. [1976]. Design of experiments to estimate heritability when observations are available on parents and offspring. Biometrics 32, 283-304.

THOMPSON, R. [1977]. The estimation of heritability with unbalanced data. To appear in Biometrics.