Applications of Linear Models in Animal Breeding

Charles R. Henderson

Chapter 1
Models

C. R. Henderson

1984 - Guelph

This book is concerned exclusively with the analysis of data arising from an experiment or sampling scheme for which a linear model is assumed to be a suitable approximation. We should not, however, be so naive as to believe that a linear model is always correct. The important consideration is whether its use permits predictions to be accomplished accurately enough for our purposes. This chapter will deal with a general formulation that encompasses all linear models that have been used in animal breeding and related fields. Some suggestions for choosing a model will also be discussed.

All linear models can, I believe, be written as follows with proper definition of the various elements of the model. Define the observable data vector with n elements as $\mathbf{y}$. In order for the problem to be amenable to a statistical analysis from which we can draw inferences concerning the parameters of the model or can predict future observations it is necessary that the data vector be regarded legitimately as a random sample from some real or conceptual population with some known or assumed distribution. Because we seldom know what the true distribution really is, a commonly used method is to assume as an approximation to the truth that the distribution is multivariate normal. Analyses based on this approximation often have remarkable power. See, for example, Cochran (1937). The multivariate normal distribution is defined completely by its mean and by its central second moments. Consequently we write a linear model for $\mathbf{y}$ with elements in the model that determine these moments. This is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}.$$

$\mathbf{X}$ is a known, fixed, $n \times p$ matrix with rank $= r \leq$ minimum of (n, p).

$\boldsymbol{\beta}$ is a fixed, $p \times 1$ vector generally unknown, although in selection index methodology it is assumed, probably always incorrectly, that it is known.

$\mathbf{Z}$ is a known, fixed, $n \times q$ matrix.

$\mathbf{u}$ is a random, $q \times 1$ vector with null means.

$\mathbf{e}$ is a random, $n \times 1$ vector with null means.

The variance-covariance matrix of $\mathbf{u}$ is $\mathbf{G}$, a $q \times q$ symmetric matrix that is usually non-singular. Hereafter for convenience we shall use the notation $Var(\mathbf{u})$ to mean a variance-covariance matrix of a random vector.

$Var(\mathbf{e}) = \mathbf{R}$ is an $n \times n$, symmetric, usually non-singular matrix. $Cov(\mathbf{u}, \mathbf{e}') = \mathbf{0}$, that is, all elements of the covariance matrix for $\mathbf{u}$ with $\mathbf{e}$ are zero in most but not all applications.

It must be understood that we have hypothesized a population of $\mathbf{u}$ vectors from which a random sample of one has been drawn into the sample associated with the data vector, $\mathbf{y}$, and similarly a population of $\mathbf{e}$ vectors is assumed, and a sample vector has been drawn with the first element of the sample vector being associated with the first element of $\mathbf{y}$, etc.

Generally we do not know the values of the individual elements of $\mathbf{G}$ and $\mathbf{R}$. We usually are willing, however, to make assumptions about the pattern of these values. For example, it is often assumed that all the diagonal elements of $\mathbf{R}$ are equal and that all off-diagonal elements are zero. That is, the elements of $\mathbf{e}$ have equal variances and are mutually uncorrelated. Given some assumed pattern of values of $\mathbf{G}$ and $\mathbf{R}$, it is then possible to estimate these matrices assuming a suitable design (values of $\mathbf{X}$ and $\mathbf{Z}$) and a suitable sampling scheme, that is, guarantee that the data vector arose in accordance with $\mathbf{u}$ and $\mathbf{e}$ being random vectors from their respective populations. With the model just described

$$\begin{aligned} E(\mathbf{y}) &= \text{mean of } \mathbf{y} = \mathbf{X}\boldsymbol{\beta}. \\ Var(\mathbf{y}) &= \mathbf{ZGZ}' + \mathbf{R}. \end{aligned}$$

We shall now present a few examples of well known models and show how these can be formulated by the general model described above. The important advantage to having one model that includes all cases is that we can thereby present in a condensed manner the basic methods for estimation, computing sampling variances, testing hypotheses, and prediction.

1 Simple Regression Model

The simple regression model can be written as follows,

$$y_i = \mu + x_i \, \alpha + e_i.$$

This is a scalar model, y_i being the i^{th} of n observations. The fixed elements of the model are μ and α, the latter representing the regression coefficient. The concomitant variable associated with the i^{th} observation is x_i, regarded as fixed and measured without error.

Note that in conceptual repeated sampling the values of x_i remain constant from one sample to another, but in each sample a new set of e_i is taken, and consequently the values of y_i change. Now relative to our general model,

$$\begin{aligned}
\mathbf{y}' &= (y_1 \quad y_2 \quad ... \quad y_n), \\
\boldsymbol{\beta}' &= (\mu \quad \alpha), \\
\mathbf{X}' &= \begin{bmatrix} 1 & 1 & ... & 1 \\ x_1 & x_2 & ... & x_n \end{bmatrix}, \text{ and} \\
\mathbf{e}' &= (e_1 \quad e_2 \quad ... \quad e_n)
\end{aligned}$$

Zu does not exist in the model. Usually **R** is assumed to be $\mathbf{I}\sigma_e^2$ in regression models.

2 One Way Random Model

Suppose we have a random sample of unrelated sires from some population of sires and that these are mated to a sample of unrelated dams with one progeny per dam. The resulting progeny are reared in a common environment, and one record is observed on each. An appropriate model would seem to be

$$y_{ij} = \mu + s_i + e_{ij},$$

y_{ij} being the observation on the j^{th} progeny of the i^{th} sire.

Suppose that there are 3 sires with progeny numbers 3, 2, l respectively. Then **y** is a vector with 6 elements.

$$\begin{aligned}
\mathbf{y}' &= (y_{11} \quad y_{12} \quad y_{13} \quad y_{21} \quad y_{22} \quad y_{31}), \\
\mathbf{x}' &= (1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1), \\
\mathbf{u}' &= (s_1 \quad s_2 \quad s_3), \text{ and} \\
\mathbf{e}' &= (e_{11} \quad e_{12} \quad e_{13} \quad e_{21} \quad e_{22} \quad e_{23}), \\
Var(\mathbf{u}) &= \mathbf{I}\sigma_s^2, \\
Var(\mathbf{e}) &= \mathbf{I}\sigma_e^2,
\end{aligned}$$

where these two identity matrices are of order 3 and 6, respectively.

$$Cov(\mathbf{u}, \mathbf{e}') = \mathbf{0}.$$

Suppose next that the sires in the sample are related, for example, sires 2 and 3 are half-sib progeny of sire l, and all 3 are non-inbred. Then under an additive genetic model

$$Var(\mathbf{u}) \; = \; \left[\begin{array}{ccc} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/4 \\ 1/2 & 1/4 & 1 \end{array} \right] \; \sigma_s^2.$$

What if the mates are related? Suppose that the numerator relationship matrix, $\mathbf{A}_m$, for the 6 mates is

$$\left(\begin{array}{cccccc} 1 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1 & 1/4 & 0 & 0 \\ 1/2 & 0 & 1/4 & 1 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1 & 1/4 \\ 0 & 1/2 & 0 & 0 & 1/4 & 1 \end{array} \right).$$

Suppose further that we invoke an additive genetic model with $h^2 = 1/4$. Then

$$Var(\mathbf{e}) \; = \; \left(\begin{array}{cccccc} 1 & 0 & 1/30 & 1/30 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1/30 & 1/30 \\ 1/30 & 0 & 1 & 1/60 & 0 & 0 \\ 1/30 & 0 & 1/60 & 1 & 0 & 0 \\ 0 & 1/30 & 0 & 0 & 1 & 1/60 \\ 0 & 1/30 & 0 & 0 & 1/60 & 1 \end{array} \right) \sigma_e^2.$$

This result is based on $\sigma_s^2 = \sigma_y^2/16, \;\; \sigma_e^2 = 15 \; \sigma_y^2/16$, and leads to

$$Var(\mathbf{y}) = (.25 \; \mathbf{A}_p + .75 \; \mathbf{I}) \; \sigma_y^2,$$

where $\mathbf{A}_p$ is the relationship matrix for the 6 progeny.

3 Two Trait Additive Genetic Model

Suppose that we have a random sample of 5 related animals with measurements on 2 correlated traits. We assume an additive genetic model. Let $\mathbf{A}$ be the numerator relationship matrix of the 5 animals. Let

$$\left(\begin{array}{cc} g_{11} & g_{12} \\ g_{12} & g_{22} \end{array} \right)$$

be the genetic variance-covariance matrix and

$$\begin{pmatrix} r_{11} & r_{12} \\ r_{12} & r_{22} \end{pmatrix}$$

be the environmental variance-covariance matrix. Then h^2 for trait 1 is $g_{11}/(g_{11}+r_{11})$, and the genetic correlation between the two traits is $g_{12}/(g_{11} \ g_{22})^{1/2}$. Order the 10 observations, animals within traits. That is, the first 5 elements of $\mathbf{y}$ are the observations on trait 1. Suppose that traits 1 and 2 have common means μ_1, μ_2 respectively. Then

$$\mathbf{X}' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix},$$

and

$$\boldsymbol{\beta}' = (\mu_1 \ \ \mu_2).$$

The first 5 elements of $\mathbf{u}$ are breeding values for trait 1 and the last 5 are breeding values for trait 2. Similarly the errors are partitioned into subvectors with 5 elements each. Then $\mathbf{Z} = \mathbf{I}$ and

$$\mathbf{G} = Var(\mathbf{u}) = \begin{pmatrix} \mathbf{A} \, g_{11} & \mathbf{A} \, g_{12} \\ \mathbf{A} \, g_{12} & \mathbf{A} \, g_{22} \end{pmatrix},$$

$$\mathbf{R} = \text{Var}\ (\mathbf{e}) = \begin{pmatrix} \mathbf{I} \, r_{11} & \mathbf{I} \, r_{12} \\ \mathbf{I} \, r_{12} & \mathbf{I} \, r_{22} \end{pmatrix},$$

where each $\mathbf{I}$ has order, 5.

4 Two Way Mixed Model

Suppose that we have a random sample of 3 unrelated sires and that they are mated to unrelated dams. One progeny of each mating is obtained, and the resulting progeny are assigned at random to two different treatments. The table of subclass numbers is

	Treatments	
Sires	1	2
1	2	1
2	0	2
3	3	0

Ordering the data by treatments within sires,

$$\mathbf{y}' = \begin{pmatrix} y_{111} & y_{112} & y_{121} & y_{221} & y_{222} & y_{311} & y_{312} & y_{313} \end{pmatrix}.$$

Treatments are regarded as fixed, and variances of sires and errors are considered to be unaffected by treatments. Then

$$\mathbf{u}' = \begin{pmatrix} s_1 & s_2 & s_3 & st_{11} & st_{12} & st_{22} & st_{31} \end{pmatrix}.$$

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

$$Var(\mathbf{s}) = \mathbf{I}_3\ \sigma^2_s, \quad Var(\mathbf{st}) = \mathbf{I}_4\ \sigma^2_{st}, \quad Var(\mathbf{e}) = \mathbf{I}_8\ \sigma^2_e.$$

$$Cov(\mathbf{s}, \quad (\mathbf{st}')) = \mathbf{0}.$$

This is certainly not the only linear model that could be invoked for this design. For example, one might want to assume that sire and error variances are related to treatments.

5 Equivalent Models

It was stated above that a linear model must describe the mean and the variance-covariance matrix of $\mathbf{y}$. Given these two, an infinity of models can be written all of which yield the same first and second moments. These models are called linear equivalent models.

Let one model be $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zu} + \mathbf{e}$ with $Var(\mathbf{u}) = \mathbf{G}$, $Var(\mathbf{e}) = \mathbf{R}$. Let a second model be $\mathbf{y} = \mathbf{X}_*\boldsymbol{\beta}_* + \mathbf{Z}_*\mathbf{u}_* + \mathbf{e}_*$, with $Var(\mathbf{u}_*) = \mathbf{G}_*$, $Var(\mathbf{e}_*) = \mathbf{R}_*$. Then the means of $\mathbf{y}$ under these 2 models are $\mathbf{X}\boldsymbol{\beta}$ and $\mathbf{X}_*\boldsymbol{\beta}_*$ respectively. $Var(\mathbf{y})$ under the 2 models is

$$\mathbf{ZGZ}' + \mathbf{R} \text{ and } \mathbf{Z}_*\mathbf{G}_*\mathbf{Z}_*' + \mathbf{R}_*.$$

Consequently we state that these 2 models are linearly equivalent if and only if

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{X}_*\boldsymbol{\beta}_* \text{ and } \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R} = \mathbf{Z}_*\mathbf{G}_*\mathbf{Z}_*' + \mathbf{R}_*.$$

To illustrate, $\mathbf{X}\boldsymbol{\beta} = \mathbf{X}_* \ \boldsymbol{\beta}_*$ suppose we have a treatment design with 3 treatments and 2 observations on each. Suppose we write a model

$$y_{ij} = \mu + t_i + e_{ij},$$

then

$$\mathbf{X}\boldsymbol{\beta} \quad = \quad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ t_1 \\ t_2 \\ t_3 \end{pmatrix}.$$

An alternative model is

$$y_{ij} = \alpha_i + e_{ij},$$

then

$$\mathbf{X}_*\boldsymbol{\beta}_* \quad = \quad \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}.$$

Then if we define $\alpha_i \ = \ \mu \ + \ t_i$, it is seen that $E(\mathbf{y})$ is the same in the two models. To illustrate with two models that give the same $Var(\mathbf{y})$ consider a repeated lactation model. Suppose we have 3 unrelated, random sample cows with $3, \ 2, \ 1$ lactation records, respectively. Invoking a simple repeatability model, that is, the correlation between any pair of records on the same animal is r, one model ignoring the fixed effects is

$$y_{ij} \ = \ c_i \ + \ e_{ij}.$$

$$Var(\mathbf{c}) \ = \ Var \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \ = \ \begin{pmatrix} r & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & r \end{pmatrix} \sigma_y^2.$$

$$Var(\mathbf{e}) \ = \ \mathbf{I}_6 \ (1 - r) \ \sigma_y^2.$$

An alternative for the random part of the model is

$$y_{ij} \ = \ e_{ij},$$

where **Zu** does not exist.

$$Var(\boldsymbol{\epsilon}) \ = \mathbf{R} \ = \ \begin{pmatrix} 1 & r & r & 0 & 0 & 0 \\ r & 1 & r & 0 & 0 & 0 \\ r & r & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & r & 0 \\ 0 & 0 & 0 & r & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \ \sigma_y^2.$$

Relating the 2 models,

$$\begin{aligned} \sigma_\epsilon^2 \ &= \ \sigma_c^2 \ + \ \sigma_e^2. \\ Cov(\epsilon_{ij}, \ \epsilon_{ij},) \ &= \ \sigma_c^2 \ \text{ for } \ j \ \neq j'. \end{aligned}$$

We shall see that some models are much easier computationally than others. Also the parameters of one model can always be written as linear functions of the parameters of any equivalent model. Consequently linear and quadratic estimates under one model can be converted by these same linear functions to estimates for an equivalent model.

6 Subclass Means Model

With some models it is convenient to write them as models for the ”smallest” subclass mean. By ”smallest” we imply a subclass identified by all of the subscripts in the model except for the individual observations. For this model to apply, the variance-covariance matrix of elements of **e** pertaining to observations in the same smallest subclass must have the form

$$\begin{pmatrix} v & & c \\ & \ddots & \\ c & & v \end{pmatrix},$$

no covariates exist, and the covariances between elements of **e** in different subclasses must be zero. Then the model can be written

$$\bar{\mathbf{y}} = \bar{\mathbf{X}}\boldsymbol{\beta} + \bar{\mathbf{Z}}\mathbf{u} + \boldsymbol{\epsilon}.$$

$\bar{\mathbf{y}}$ is the vector of "smallest" subclass means. $\bar{\mathbf{X}}$ and $\bar{\mathbf{Z}}$ relate these means to elements of $\boldsymbol{\beta}$ and $\mathbf{u}$. The error vector, $\boldsymbol{\epsilon}$, is the mean of elements of $\mathbf{e}$ in the same subclass. Its variance-covariance matrix is diagonal with the i^{th} diagonal element being

$$\left(\frac{v}{n_i} + \frac{n_i - 1}{n_i} c \right) \sigma_e^2,$$

where n_i is the number of observations in the i^{th} subclass.

7 Determining Possible Elements In The Model

Henderson(1959) described in detail an algorithm for determining the potential lines of an ANOVA table and correspondingly the elements of a linear model. First, the experiment is described in terms of two types of factors, namely main factors and nested factors. By a main factor is meant a classification, the "levels" of which are identified by a single subscript. By a nesting factor is meant one whose levels are not completely identified except by specifying a main factor or a combination of main factors within which the nesting factor is nested. Identify each of the main factors by a single unique letter, for example, B for breeds and T for treatments. Identify nesting factors by a letter followed by a colon and then the letter or letters describing the main factor or factors within which it is nested. For example, if sires are nested within breeds, this would be described as S:B. On the other hand, if a different set of sires is used for each breed by treatment combination, sires would be identified as S:BT. To determine potential 2 factor interactions combine the letters to the left of the colon (for a main factor a colon is implied with no letters following). Then combine the letters without repetition to the right of the colon. If no letter appears on both the right and left of the colon this is a valid 2 factor interaction. For example, factors are A,B,C:B. Two way combinations are AB, AC:B, BC:B. The third does not qualify since B appears to the left and right of the colon. AC:B means A by C interaction nested within B. Three factor and higher interactions are determined by taking all possible trios and carrying out the above procedure. For example, factors are (A, D, B:D, C:D). Two factor possibilities are (AD, AB:D, AC:D, DB:D, DC:D, BC:D). The 4^{th} and 5^{th} are not valid. Three factor possibilities are (ADB:D, ADC:D, ABC:D, DBC:D). None of these is valid except ABC:D. The four factor possibility is ADBC:D, and this is not valid.

Having written the main factors and interactions one uses each of these as a subvector of either $\boldsymbol{\beta}$ or $\mathbf{u}$. The next question is how to determine which. First consider main factors and nesting factors. If the levels of the factor in the experiment can be regarded as a random sample from some population of levels, the levels would be a subvector of $\mathbf{u}$. With respect to interactions, if one or more letters to the left of the colon represent a factor in $\mathbf{u}$, the interaction levels are subvectors of $\mathbf{u}$. Thus interaction of fixed by random factors is regarded as random, as is the nesting of random within fixed. As a final step we decide the variance-covariance matrix of each subvector of $\mathbf{u}$, the covariance between subvectors of $\mathbf{u}$, and the variance- covariance matrix of $(\mathbf{u}, \mathbf{e})$. These last decisions are based on knowledge of the biology and the sampling scheme that produced the data vector.

It seems to me that modelling is the most important and most difficult aspect of linear models applications. Given the model everything else is essentially computational.

Chapter 2
Linear Unbiased Estimation

C. R. Henderson

1984 - Guelph

We are interested in linear unbiased estimators of $\boldsymbol{\beta}$ or of linear functions of $\boldsymbol{\beta}$, say $\mathbf{k}'\boldsymbol{\beta}$. That is, the estimator has the form, $\mathbf{a}'\mathbf{y}$, and E $(\mathbf{a}'\mathbf{y}) = \mathbf{k}'\boldsymbol{\beta}$, if possible. It is not necessarily the case that $\mathbf{k}'\boldsymbol{\beta}$ can be estimated unbiasedly. If $\mathbf{k}'\boldsymbol{\beta}$ can be estimated unbiasedly, it is called estimable. How do we determine estimability?

1 Verifying Estimability

$$E(\mathbf{a}'\mathbf{y}) = \mathbf{a}'\mathbf{X}\boldsymbol{\beta}.$$

Does this equal $\mathbf{k}'\boldsymbol{\beta}$? It will for any value of $\boldsymbol{\beta}$ if and only if $\mathbf{a}'\mathbf{X} = \mathbf{k}'$.

Consequently, if we can find any $\mathbf{a}$ such that $\mathbf{a}'\mathbf{X} = \mathbf{k}'$, then $\mathbf{k}'\boldsymbol{\beta}$ is estimable. Let us illustrate with

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 1 & 1 & 2 \\ 1 & 3 & 6 \end{pmatrix}.$$

- Is β_1 estimable, that is, $(1 \;\; 0 \;\; 0)$ $\boldsymbol{\beta}$ estimable? Let $\mathbf{a}' = (2 \;\; -1 \;\; 0 \;\; 0)$ then

  $$\mathbf{a}'\mathbf{X} = (1 \;\; 0 \;\; 0) = \mathbf{k}'.$$

  Therefore, $\mathbf{k}'\boldsymbol{\beta}$ is estimable.

- Is $(0 \;\; 1 \;\; 2)$ $\boldsymbol{\beta}$ estimable? Let $\mathbf{a}' = (-1 \;\; 1 \;\; 0 \;\; 0)$ then

  $$\mathbf{a}'\mathbf{X} = (0 \;\; 1 \;\; 2) = \mathbf{k}'.$$

  Therefore, it is estimable.

- Is β_2 estimable? No, because no $\mathbf{a}'$ exists such that $\mathbf{a}'\mathbf{X} = (0 \;\; 1 \;\; 0)$.

Generally it is easier to prove by the above method that an estimable function is indeed estimable than to prove that a non-estimable function is non-estimable. Accordingly, we consider other methods for determining estimability.

1.1 Second Method

Partition $\mathbf{X}$ as follows with possible re-ordering of columns.

$$\mathbf{X} = (\mathbf{X}_1 \quad \mathbf{X}_1\mathbf{L}),$$

where $\mathbf{X}_1$ has r linearly independent columns. Remember that $\mathbf{X}$ is $n \times p$ with rank $= r$. The dimensions of $\mathbf{L}$ are $r \times (p - r)$.

Then $\mathbf{k}'\boldsymbol{\beta}$ is estimable if and only if

$$\mathbf{k}' = (\mathbf{k}_1' \quad \mathbf{k}_1'\mathbf{L}),$$

where $\mathbf{k}_1'$ has r elements, and $\mathbf{k}_1'\mathbf{L}$ has $p - r$ elements. Consider the previous example.

$$\mathbf{X}_1 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \\ 1 & 3 \end{pmatrix}, \text{ and } \mathbf{L} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

- Is $(1 \quad 0 \quad 0)\ \boldsymbol{\beta}$ estimable?

  $$\mathbf{k}_1' = (1 \;\; 0),\ \mathbf{k}_1'\mathbf{L} = (1 \;\; 0) \begin{pmatrix} 0 \\ 2 \end{pmatrix} = 0.$$

  Thus $\mathbf{k}' = (1 \;\; 0 \;\; 0)$, and the function is estimable.

- Is $(0 \;\; 1 \;\; 2)\ \boldsymbol{\beta}$ estimable?

  $$\mathbf{k}_1' = (0 \;\; 1), \text{ and } \mathbf{k}_1'\mathbf{L} = (0 \;\; 1)\begin{pmatrix} 0 \\ 2 \end{pmatrix} = 2.$$

  Thus $\mathbf{k}' = (0 \;\; 1 \;\; 2)$, and the function is estimable.

- Is $(0 \;\; 1 \;\; 0)\ \boldsymbol{\beta}$ estimable?

  $$\mathbf{k}_1' = (0 \;\; 1), \text{ and } \mathbf{k}_1'\mathbf{L} = (0 \;\; 1)\begin{pmatrix} 0 \\ 2 \end{pmatrix} = 2.$$

  Thus $(\mathbf{k}_1' \quad \mathbf{k}_1'\mathbf{L}). = (0 \;\; 1 \;\; 2) \neq (0 \;\; 1 \;\; 0)$. The function is not estimable.

1.2 Third Method

A third method is to find a matrix, $\mathbf{C}$, of order $p \times (p - r)$ and rank, $p - r$, such that

$$\mathbf{XC} = \mathbf{0}.$$

Then $\mathbf{k}'\beta$ is estimable if and only if

$$\mathbf{k}'\mathbf{C} = \mathbf{0}.$$

In the example

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 1 & 1 & 2 \\ 1 & 3 & 6 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

- Therefore $(1 \;\; 0 \;\; 0) \;\; \boldsymbol{\beta}$ is estimable because

$$(1 \;\; 0 \;\; 0) \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = 0.$$

- So is $(0 \;\; 1 \;\; 2) \; \boldsymbol{\beta}$ because

$$(0 \;\; 1 \;\; 2) \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = 0.$$

- But $(0 \;\; 1 \;\; 0) \; \boldsymbol{\beta}$ is not because

$$(0 \;\; 1 \;\; 0) \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = 2 \neq 0.$$

1.3 Fourth Method

A fourth method is to find some g-inverse of $\mathbf{X}'\mathbf{X}$, denoted by $(\mathbf{X}'\mathbf{X})^-$. Then $\mathbf{k}'\boldsymbol{\beta}$ is estimable if and only if

$$\mathbf{k}'(\mathbf{X}'\mathbf{X})^- \; \mathbf{X}'\mathbf{X} \; = \; \mathbf{k}'.$$

A definition of and methods for computing a g-inverse are presented in Chapter 3.

In the example

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 4 & 7 & 14 \\ 7 & 15 & 30 \\ 14 & 30 & 60 \end{pmatrix},$$

and a g-inverse is

$$\frac{1}{11} \begin{pmatrix} 15 & -7 & 0 \\ -7 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

- $(1 \ \ 0 \ \ 0) \ (\mathbf{X}'\mathbf{X})^- \ \mathbf{X}'\mathbf{X} \ = \ (1 \ \ 0 \ \ 0)$. Therefore $(1 \ \ 0 \ \ 0) \ \boldsymbol{\beta}$ is estimable.
- $(0 \ \ 1 \ \ 2) \ (\mathbf{X}'\mathbf{X})^- \ \mathbf{X}'\mathbf{X} \ = \ (0 \ \ 1 \ \ 2)$. Therefore $(0 \ \ 1 \ \ 2) \ \boldsymbol{\beta}$ is estimable.
- $(0 \ \ 1 \ \ 0) \ (\mathbf{X}'\mathbf{X})^- \ \mathbf{X}'\mathbf{X} \ = \ (0 \ \ 1 \ \ 2)$. Therefore $(0 \ \ 1 \ \ 0) \ \boldsymbol{\beta}$ is not estimable.
- Related to this fourth method any linear function of

  $$(\mathbf{X}'\mathbf{X})^- \ \mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

  is estimable.

If rank $(\mathbf{X}) = p =$ the number of columns in $\mathbf{X}$, any linear function of $\boldsymbol{\beta}$ is estimable. In that case the only g-inverse of $\mathbf{X}'\mathbf{X}$ is $(\mathbf{X}'\mathbf{X})^{-1}$, a regular inverse. Then by the fourth method

$$\mathbf{k}' \ (\mathbf{X}'\mathbf{X})^- \ \mathbf{X}'\mathbf{X} \ = \ \mathbf{k}'(\mathbf{X}'\mathbf{X})^{-1} \ \mathbf{X}'\mathbf{X} \ = \ \mathbf{k}' \ \mathbf{I} \ = \ \mathbf{k}'.$$

Therefore, any $\mathbf{k}'\boldsymbol{\beta}$ is estimable.

There is an extensive literature on generalized inverses. See for example, Searle (1971b, 1982), Rao and Mitra (1971) and Harville(1999??).

Chapter 3
Best Linear Unbiased Estimation

C. R. Henderson

1984 - Guelph

In Chapter 2 we discussed linear unbiased estimation of $\mathbf{k}'\boldsymbol{\beta}$, having determined that it is estimable. Let the estimate be $\mathbf{a}'\mathbf{y}$, and if $\mathbf{k}'\boldsymbol{\beta}$ is estimable, some $\mathbf{a}$ exists such that

$$E(\mathbf{a}'\mathbf{y}) = \mathbf{k}'\boldsymbol{\beta}.$$

Assuming that more than one $\mathbf{a}$ gives an unbiased estimator, which one should be chosen? The most common criterion for choice is minimum sampling variance. Such an estimator is called the best linear unbiased estimator (BLUE).

Thus we find $\mathbf{a}'$ such that $E(\mathbf{a}'\mathbf{y}) = \mathbf{k}'\boldsymbol{\beta}$ and, in the class of such estimators, has minimum sampling variance. Now

$$Var(\mathbf{a}'\mathbf{y}) = \mathbf{a}'(Var(\mathbf{y}))\mathbf{a} = \mathbf{a}'\mathbf{V}\mathbf{a},$$

where $Var(\mathbf{y}) = \mathbf{V}$, assumed known, for the moment.

For unbiasedness we require $\mathbf{a}'\mathbf{X} = \mathbf{k}'$. Consequently we find $\mathbf{a}$ that minimizes $\mathbf{a}'\mathbf{V}\mathbf{a}$ subject to $\mathbf{a}'\mathbf{X} = \mathbf{k}'$. Using a Lagrange multiplier, $\boldsymbol{\theta}$, and applying differential calculus we need to solve for $\mathbf{a}$ in equations

$$\begin{pmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{k} \end{pmatrix}.$$

This is a consistent set of equations if and only if $\mathbf{k}'\boldsymbol{\beta}$ is estimable. In that case the unique solution to $\mathbf{a}$ is

$$\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{k}.$$

A solution to $\boldsymbol{\theta}$ is

$$-(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{k},$$

and this is not unique when $\mathbf{X}$ and consequently $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ is not full rank. Nevertheless the solution to $\mathbf{a}$ is invariant to the choice of a g-inverse of $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$. Thus, BLUE of $\mathbf{k}'\boldsymbol{\beta}$ is

$$\mathbf{k}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}.$$

But let

$$\boldsymbol{\beta}^{o} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y},$$

where $\boldsymbol{\beta}^o$ is any solution to

$$(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})\boldsymbol{\beta}^o = \mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$$

known as generalized least squares (GLS) equations, Aitken (1935). Superscript 0 is used to denote some solution, not a unique solution. Therefore BLUE of $\mathbf{k}'\boldsymbol{\beta}$ is $\mathbf{k}'\boldsymbol{\beta}^o$.

Let us illustrate with

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 1 & 1 & 2 \\ 1 & 3 & 6 \end{pmatrix},$$

and $\mathbf{y}' = (5\ 2\ 4\ 3)$. Suppose $Var(\mathbf{y}) = \mathbf{I}\sigma_e^2$. Then the GLS equations are

$$\sigma_e^{-2} \begin{pmatrix} 4 & 7 & 14 \\ 7 & 15 & 30 \\ 14 & 30 & 60 \end{pmatrix} \quad \boldsymbol{\beta}^o = \begin{pmatrix} 14 \\ 22 \\ 44 \end{pmatrix} \sigma_e^{-2}.$$

A solution is

$$(\boldsymbol{\beta}^o)' = (56 \quad -10 \quad 0)/11.$$

Then BLUE of $(0\ \ 1\ \ 2)\boldsymbol{\beta}$, which has been shown to be estimable, is

$$(0\ \ 1\ \ 2)(56 \quad -10 \quad 0)'/11 = -10/11.$$

Another solution to $\boldsymbol{\beta}^o$ is

$$(56 \quad 0 \quad -5)'/11.$$

Then BLUE of $(0\ \ 1\ \ 2)\boldsymbol{\beta}$ is $-10/11$, the same as the other solution to $\boldsymbol{\beta}^o$.

1 Mixed Model Method For BLUE

One frequent difficulty with GLS equations, particularly in the mixed model, is that $\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$ is large and non-diagonal. Consequently $\mathbf{V}^{-1}$ is difficult or impossible to compute by usual methods. It was proved by Henderson *et al*. (1959) that

$$\mathbf{V}^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}.$$

Now if $\mathbf{R}^{-1}$ is easier to compute than $\mathbf{V}^{-1}$, as is often true, if $\mathbf{G}^{-1}$ is easy to compute, and $(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}$ is easy to compute, this way of computing $\mathbf{V}^{-1}$ may have important advantages. Note that this result can be obtained by writing equations, known as Henderson's mixed model equations (1950) as follows,

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}.$$

Note that if we solve for $\hat{\mathbf{u}}$ in the second equation and substitute this in the first we get

$$\begin{aligned} & \mathbf{X}'[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}]\mathbf{X}\boldsymbol{\beta}^o \\ = \; & \mathbf{X}'[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}]\mathbf{y}, \end{aligned}$$

or from the result for $\mathbf{V}^{-1}$

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}^o = \mathbf{X}'\mathbf{V}^{-1}\mathbf{y}.$$

Thus, a solution to $\boldsymbol{\beta}^o$ in the mixed model equations is a GLS solution. An interpretation of $\hat{\mathbf{u}}$ is given in Chapter 5. The mixed model equations are often well suited to an iterative solution. Let us illustrate the mixed model method for BLUE with

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \\ 1 & 3 \end{pmatrix}, \; \mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \; \mathbf{G} = \begin{pmatrix} .1 & 0 \\ 0 & .1 \end{pmatrix},$$

and

$$\mathbf{R} = \mathbf{I}, \;\; \mathbf{y}' = [5 \; 4 \; 3 \; 2].$$

Then the mixed model equations are

$$\begin{pmatrix} 4 & 7 & 3 & 1 \\ 7 & 15 & 4 & 3 \\ 3 & 4 & 13 & 0 \\ 1 & 3 & 0 & 11 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} 14 \\ 22 \\ 12 \\ 2 \end{pmatrix}.$$

The solution is $[286 \;\; -50 \; 2 \;\; -2]'/57$. In this case the solution is unique because $\mathbf{X}$ has full column rank.

Now consider a GLS solution.

$$\mathbf{V} = [\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}] = \begin{pmatrix} 1.1 & .1 & .1 & 0 \\ .1 & 1.1 & .1 & 0 \\ .1 & .1 & 1.1 & 0 \\ 0 & 0 & 0 & 1.1 \end{pmatrix}.$$

$$\mathbf{V}^{-1} = \frac{1}{143} \begin{pmatrix} 132 & -11 & -11 & 0 \\ -11 & 132 & -11 & 0 \\ -11 & -11 & 132 & 0 \\ 0 & 0 & 0 & 130 \end{pmatrix}.$$

Then $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}^o = \mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ becomes

$$\frac{1}{143} \begin{pmatrix} 460 & 830 \\ 830 & 1852 \end{pmatrix} \boldsymbol{\beta}^o = \frac{1}{143} \begin{pmatrix} 1580 \\ 2540 \end{pmatrix}.$$

The solution is $(286 \;\; -50)/57$ as in the mixed model equations.

2 Variance of BLUE

Once having an estimate of $\mathbf{k}'\boldsymbol{\beta}$ we should like to know its sampling variance. Consider a set of estimators, $\mathbf{K}'\boldsymbol{\beta}^o$.

$$\begin{aligned} Var(\mathbf{K}'\boldsymbol{\beta}^o) &= Var[\mathbf{K}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}] \\ &= \mathbf{K}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{V}\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{K} \\ &= \mathbf{K}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{K} \text{ provided } \mathbf{K}'\boldsymbol{\beta} \text{ is estimable.} \end{aligned}$$

The variance is invariant to the choice of a g-inverse provided $\mathbf{K}'\boldsymbol{\beta}$ is estimable. We can also obtain this result from a g-inverse of the coefficient matrix of the mixed model equations. Let a g-inverse of this matrix be

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix}.$$

Then

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) = \mathbf{K}'\mathbf{C}_{11}\mathbf{K}.$$

This result can be proved by noting that

$$\begin{aligned} \mathbf{C}_{11} &= (\mathbf{X}'[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}]\mathbf{X})^{-} \\ &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}. \end{aligned}$$

Using the mixed model example, let

$$\mathbf{K}' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A g-inverse (regular inverse) of the coefficient matrix is

$$\frac{1}{570}\begin{pmatrix} 926 & -415 & -86 & 29 \\ -415 & 230 & 25 & -25 \\ -86 & 25 & 56 & 1 \\ 29 & -25 & 1 & 56 \end{pmatrix}.$$

Then

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) = \frac{1}{570}\left(\begin{array}{rr} 926 & -415 \\ -415 & 230 \end{array} \right].$$

The same result can be obtained from the inverse of the GLS coefficient matrix because

$$\left(143^{-1}\begin{pmatrix} 460 & 830 \\ 830 & 1852 \end{pmatrix} \right)^{-1} = \frac{1}{570}\begin{pmatrix} 926 & -415 \\ -415 & 230 \end{pmatrix}.$$

3 Generalized Inverses and Mixed Model Equations

Earlier in this chapter we found that BLUE of $\mathbf{K}'\boldsymbol{\beta}$, estimable, is $\mathbf{K}'\boldsymbol{\beta}^o$, where $\boldsymbol{\beta}^o$ is any solution to either GLS or mixed model equations. Also the sampling variance requires a g-inverse of the coefficient matrix of either of these sets of equations. We define $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^-$ as a g-inverse of $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$. There are various types of generalized inverses, but the one we shall use is defined as follows.

$\mathbf{A}^-$ is a g-inverse of $\mathbf{A}$ provided that

$$\mathbf{A}\ \mathbf{A}^-\mathbf{A} = \mathbf{A}.$$

Then if we have a set of consistent equations,

$$\mathbf{A}\ \mathbf{p} = \mathbf{z},$$

a solution to $\mathbf{p}$ is

$$\mathbf{A}^-\mathbf{z}.$$

We shall be concerned, in this chapter, only with g-inverses of singular, symmetric matrices characteristic of GLS and mixed model equations.

3.1 First type of g-inverse

Let $\mathbf{W}$ be a symmetric matrix with order, s, and rank, $t < s$. Partition $\mathbf{W}$ with possible re-ordering of rows (and the same re-ordering of columns) as

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{12}' & \mathbf{W}_{22} \end{pmatrix},$$

where $\mathbf{W}_{11}$ is a non-singular matrix with order t. Then $\mathbf{W}^- = \begin{pmatrix} \mathbf{W}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$.

It is of interest that for this type of $\mathbf{W}^-$ it is true that $\mathbf{W}^-\mathbf{W}\ \mathbf{W}^- = \mathbf{W}^-$ as well as $\mathbf{W}\ \mathbf{W}^-\mathbf{W} = \mathbf{W}$. This is called a reflexive g-inverse. To illustrate, suppose $\mathbf{W}$ is a GLS coefficient matrix,

$$\mathbf{W} = \begin{pmatrix} 4 & 7 & 8 & 15 \\ 7 & 15 & 17 & 32 \\ 8 & 17 & 22 & 39 \\ 15 & 32 & 39 & 71 \end{pmatrix}.$$

This matrix has rank 3 and the upper 3×3 is non-singular with inverse

$$30^{-1} \begin{pmatrix} 41 & -18 & -1 \\ -18 & 24 & -12 \\ -1 & -12 & 11 \end{pmatrix}.$$

Therefore a g-inverse is

$$30^{-1} \begin{pmatrix} 41 & -18 & -1 & 0 \\ -18 & 24 & -12 & 0 \\ -1 & -12 & 11 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Another g-inverse of this type is

$$30^{-1} \begin{pmatrix} 41 & -17 & 0 & -1 \\ -17 & 59 & 0 & -23 \\ 0 & 0 & 0 & 0 \\ -1 & -23 & 0 & 11 \end{pmatrix}.$$

This was obtained by inverting the full rank submatrix composed of rows (and columns) 1, 2, 4 of $\mathbf{W}$. This type of g-inverse is described in Searle (1971b).

In the mixed model equations a comparable g-inverse is obtained as follows. Partition $\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}$ with possible re-ordering of rows (and columns) as

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_2 \end{pmatrix}$$

so that $\mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1$ has order r and is full rank. Compute

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{C}_{00} & \mathbf{C}_{02} \\ \mathbf{C}_{02}' & \mathbf{C}_{22} \end{pmatrix}.$$

Then a g-inverse of the coefficient matrix is $\begin{pmatrix} \mathbf{C}_{00} & \mathbf{0} & \mathbf{C}_{02} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{C}_{02}' & \mathbf{0} & \mathbf{C}_{22} \end{pmatrix}$. We illustrate with a mixed model coefficient matrix as follows.

$$\begin{pmatrix} 5 & 8 & -8 & 3 & 2 \\ 8 & 16 & -16 & 4 & 4 \\ -8 & -16 & 16 & -4 & -4 \\ 3 & 4 & -4 & 8 & 0 \\ 2 & 4 & -4 & 0 & 7 \end{pmatrix}$$

where $\mathbf{X}$ has 3 columns and $\mathbf{Z}$ has 2. Therefore $\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}$ is the upper 3 x 3 submatrix. It has rank 2 because the 3rd column is the negative of the second. Consequently find a g-inverse by inverting the matrix with the 3rd row and column deleted. This gives

$$560^{-1} \begin{pmatrix} 656 & -300 & 0 & -96 & -16 \\ -300 & 185 & 0 & 20 & -20 \\ 0 & 0 & 0 & 0 & 0 \\ -96 & 20 & 0 & 96 & 16 \\ -16 & -20 & 0 & 16 & 96 \end{pmatrix}.$$

With this type of g-inverse the solution to $\boldsymbol{\beta}^o$ is $(\boldsymbol{\beta}_1^o \;\; \mathbf{0})'$, where $\boldsymbol{\beta}_1^o$ has r elements. Only the first p rows of the mixed model equations contribute to lack of rank of the mixed model matrix. The matrix has order $p+q$ and rank $r+q$, where r = rank of $\mathbf{X}$, p = columns in $\mathbf{X}$, and q = columns in $\mathbf{Z}$.

3.2 Second type of g-inverse

A second type of g-inverse is one which imposes restrictions on the solution to $\boldsymbol{\beta}^o$. Let $\mathbf{M}'\boldsymbol{\beta}$ be a set of $p-r$ linearly independent, non-estimable functions of $\boldsymbol{\beta}$. Then a g-inverse for the GLS matrix is obtained as follows $\begin{pmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{M} \\ \mathbf{M}' & \mathbf{O} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix}$.

$\mathbf{C}_{11}$ is a reflexive g-inverse of $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$. This type of solution is described in Kempthorne (1952). Let us illustrate GLS equations as follows.

$$\begin{pmatrix} 11 & 5 & 6 & 3 & 8 \\ 5 & 5 & 0 & 2 & 3 \\ 6 & 0 & 6 & 1 & 5 \\ 3 & 2 & 1 & 3 & 0 \\ 8 & 3 & 5 & 0 & 8 \end{pmatrix} \boldsymbol{\beta}^o = \begin{pmatrix} 12 \\ 7 \\ 5 \\ 8 \\ 4 \end{pmatrix}.$$

This matrix has order 5 but rank only 3. Two independent non-estimable functions are needed. Among others the following qualify

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \boldsymbol{\beta}.$$

Therefore we invert

$$\begin{pmatrix} 11 & 5 & 6 & 3 & 8 & 0 & 0 \\ 5 & 5 & 0 & 2 & 3 & 1 & 0 \\ 6 & 0 & 6 & 1 & 5 & 1 & 0 \\ 3 & 2 & 1 & 3 & 0 & 0 & 1 \\ 8 & 3 & 5 & 0 & 8 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix},$$

which is

$$244^{-1} \begin{pmatrix} 28 & -1 & 1 & 13 & -13 & -122 & -122 \\ -1 & 24 & -24 & -7 & 7 & 122 & 0 \\ 1 & -24 & 24 & 7 & -7 & 122 & 0 \\ 13 & -7 & 7 & 30 & -30 & 0 & 122 \\ -13 & 7 & -7 & -30 & 30 & 0 & 122 \\ -122 & 122 & 122 & 0 & 0 & 0 & 0 \\ -122 & 0 & 0 & 122 & 122 & 0 & 0 \end{pmatrix}.$$

The upper 5 x 5 submatrix is a g-inverse. This gives a solution

$$\boldsymbol{\beta}^o \;=\; (386 \;\; 8 \quad -8 \;\; 262 \quad -262)'/244.$$

A corresponding g-inverse for the mixed model is as follows

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{M} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1} & 0 \\ \mathbf{M}' & 0 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{13}' & \mathbf{C}_{23} & \mathbf{C}_{33} \end{pmatrix}.$$

Then

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix}$$

is a g-inverse of the mixed model coefficient matrix. The property of $\boldsymbol{\beta}^o$ coming from this type of g-inverse is

$$\mathbf{M}'\boldsymbol{\beta}^o = \mathbf{0}.$$

3.3 Third type of g-inverse

A third type of g-inverse uses $\mathbf{M}$ of the previous section as follows. $(\mathbf{X}\ \text{'}\mathbf{V}\ ^{-1}\mathbf{X} + \mathbf{M}\mathbf{M}')^{-1} = \mathbf{C}$. Then $\mathbf{C}$ is a g-inverse of $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$. In this case $\mathbf{C}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})\mathbf{C} \neq \mathbf{C}$. This is described in Rao and Mitra (1971).

We illustrate with the same GLS matrix as before and

$$\mathbf{M}' = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

as before.

$$(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} + \mathbf{M}\mathbf{M}') = \begin{pmatrix} 11 & 5 & 6 & 3 & 8 \\ 5 & 6 & 1 & 2 & 3 \\ 6 & 1 & 7 & 1 & 5 \\ 3 & 2 & 1 & 4 & 1 \\ 8 & 3 & 5 & 1 & 9 \end{pmatrix}$$

with inverse

$$244^{-1} \begin{pmatrix} 150 & -62 & -60 & -48 & -74 \\ -62 & 85 & 37 & -7 & 7 \\ -60 & 37 & 85 & 7 & -7 \\ -48 & -7 & 7 & 91 & 31 \\ -74 & 7 & -7 & 31 & 91 \end{pmatrix},$$

which is a g-inverse of the GLS matrix. The resulting solution to $\boldsymbol{\beta}^o$ is the same as the previous section.

The corresponding method for finding a g-inverse of the mixed model matrix is $\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} + \mathbf{M}\mathbf{M}' & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix}^{-1} = \mathbf{C}$. Then $\mathbf{C}$ is a g-inverse. The property of the solution to $\boldsymbol{\beta}^o$ is

$$\mathbf{M}'\boldsymbol{\beta}^o = \mathbf{0}.$$

4 Reparameterization

An entirely different method for dealing with the not full rank $\mathbf{X}$ problem is reparameterization. Let $\mathbf{K}'\boldsymbol{\beta}$ be a set of r linearly independent, estimable functions of $\boldsymbol{\beta}$. Let $\hat{\boldsymbol{\alpha}}$ be BLUE of $\mathbf{K}'\boldsymbol{\beta}$. To find $\hat{\boldsymbol{\alpha}}$ solve $(\mathbf{K}\text{'}\mathbf{K})^{-1}\mathbf{K}'\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\mathbf{K}(\mathbf{K}'\mathbf{K})^{-1}\hat{\boldsymbol{\alpha}} = (\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$. $\hat{\boldsymbol{\alpha}}$ has a unique solution, and the regular inverse of the coefficient matrix is $Var(\hat{\boldsymbol{\alpha}})$. This corresponds to a model

$$E(\mathbf{y}) = \mathbf{X} \;\; \mathbf{K}(\mathbf{K}'\mathbf{K})^{-1}\boldsymbol{\alpha}.$$

This method was suggested to me by Gianola (1980).

From the immediately preceding example we need 3 estimable functions. An independent set is

$$\begin{pmatrix} 1 & 1/2 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

The corresponding GLS equations are

$$\begin{pmatrix} 11 & -.50 & -2.50 \\ -.5 & 2.75 & .75 \\ -2.5 & .75 & 2.75 \end{pmatrix} \hat{\boldsymbol{\alpha}} = \begin{pmatrix} 12 \\ 1 \\ 2 \end{pmatrix}.$$

The solution is

$$\hat{\boldsymbol{\alpha}}' = (193 \;\; 8 \;\; 262)/122.$$

This is identical to

$$\begin{pmatrix} 1 & 1/2 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\beta}^o$$

from the previous solution in which

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \boldsymbol{\beta}^o$$

was forced to equal $\mathbf{0}$.

The corresponding set of equations for mixed models is

$$\begin{pmatrix} (\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\mathbf{K}(\mathbf{K}'\mathbf{K})^{-1} & (\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\mathbf{K}(\mathbf{K}'\mathbf{K})^{-1} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} (\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}.$$

5 Precautions in Solving Equations

Precautions must be observed in the solution to equations, especially if there is some doubt about the rank of the matrix. If a supposed g-inverse is calculated, it may be advisable to check that $\mathbf{A}\mathbf{A}^{-}\mathbf{A} = \mathbf{A}$. Another check is to regenerate the right hand sides as follows. Let the equations be

$$\mathbf{C}\hat{\boldsymbol{\alpha}} = \mathbf{r}.$$

Having computed $\hat{\boldsymbol{\alpha}}$, compute $\mathbf{C}\hat{\boldsymbol{\alpha}}$ and check that it is equal, except for rounding error, to $\mathbf{r}$.

Chapter 4
Test of Hypotheses

C. R. Henderson

1984 - Guelph

Much of the statistical literature for many years dealt primarily with tests of hypotheses (or tests of significance). More recently increased emphasis has been placed, properly I think, on estimation and prediction. Nevertheless, many research workers and certainly most editors of scientific journals insist on tests of significance. Most tests involving linear models can be stated as follows. We wish to test the null hypothesis,

$$\mathbf{H}_0^{'}\boldsymbol{\beta} \;\; = \;\; \mathbf{c}_0,$$

against some alternative hypothesis, most commonly the alternative that $\boldsymbol{\beta}$ can have any value in the parameter space. Another possibility is the general alternative hypothesis,

$$\mathbf{H}_a^{'}\boldsymbol{\beta} \;\; = \;\; \mathbf{c}_a.$$

In both of these hypotheses there may be elements of $\boldsymbol{\beta}$ that are not determined by $\mathbf{H}$. These elements are assumed to have any values in the parameter space. $\mathbf{H}_0^{'}$ and $\mathbf{H}_a^{'}$ are assumed to have full row rank with m and a rows respectively. Also $r \geq m > a$. Under the unrestricted hypothesis $a = 0$.

Two important restrictions are required logically for $\mathbf{H}_0$ and $\mathbf{H}_a$. First, both $\mathbf{H}_0^{'}\boldsymbol{\beta}$ and $\mathbf{H}_a^{'}\boldsymbol{\beta}$ must be estimable. It hardly seems logical that we could test hypotheses about functions of $\boldsymbol{\beta}$ unless we can estimate these functions. Second, the null hypothesis must be contained in the alternative hypothesis. That is, if the null is true, the alternative must be true. For this to be so we require that $\mathbf{H}_a^{'}$ can be written as $\mathbf{M}\mathbf{H}_0^{'}$ and $\mathbf{c}_a$ as $\mathbf{M}\mathbf{c}_0$ for some $\mathbf{M}$.

1 Equivalent Hypotheses

It should be recognized that there are an infinity of hypotheses that are equivalent to $\mathbf{H}_0^{'}\boldsymbol{\beta} = \mathbf{c}$. Let $\mathbf{P}$ be an $m \times m$, non-singular matrix. Then $\mathbf{P}\mathbf{H}_0^{'}\boldsymbol{\beta} = \mathbf{P}\mathbf{c}$ is equivalent to

$\mathbf{H}_0'\boldsymbol{\beta} = \mathbf{c}$. For example, consider a fixed model

$$y_{ij} \ = \ \mu + t_i + e_{ij}, \quad i \ = \ 1,\ 2,\ 3.$$

A null hypothesis often tested is

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \ \mathbf{t} = \mathbf{0}.$$

An equivalent hypothesis is

$$\begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \end{pmatrix} \ \mathbf{t} = \mathbf{0}.$$

To convert the first to the second pre-multiply

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \quad \text{by} \quad \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix}.$$

As an example of use of $\mathbf{H}_a'$ consider a type of analysis sometimes recommended for a two way fixed model without interaction. Let the model be $y_{ijk} \ = \ \mu + a_i + b_j + e_{ijk}$, where $i \ = \ 1,\ 2,\ 3$ and $j \ = \ 1,\ 2,\ 3,\ 4$. The lines of the ANOVA table could be as follows.

Sum of Squares
Rows ignoring columns (column differences regarded as non-existent),
Columns with rows accounted for,
Residual.

The sum of these 3 sums of squares is equal to ($\mathbf{y}'\mathbf{y}-$ correction factor). The first sum of squares is represented as testing the null hypothesis:

$$\begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\beta} = \mathbf{0}.$$

and the alternative hypothesis:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\beta} = \mathbf{0}.$$

The second sum of squares represents testing the null hypothesis:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\beta} = \mathbf{0}.$$

and the alternative hypothesis: entire parameter space.

2 Test Criteria

2.1 Differences between residuals

Now it is assumed for purposes of testing hypotheses that $\mathbf{y}$ has a multivariate normal distribution. Then it can be proved by the likelihood ratio method of testing hypotheses, Neyman and Pearson (1933), that under the null hypothesis the following quantity is distributed as χ^2.

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0)'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0) \quad - \quad (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_a)'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_a). \tag{1}$$

$\boldsymbol{\beta}_0$ is a solution to GLS equations subject to the restriction $\mathbf{H}_0'\boldsymbol{\beta}_0 = \mathbf{c}_0$. $\boldsymbol{\beta}_0$ can be found by solving

$$\begin{pmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{H}_0 \\ \mathbf{H}_0' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_0 \\ \boldsymbol{\theta}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ \mathbf{c}_0 \end{pmatrix}$$

or by solving the comparable mixed model equations

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{H}_0 \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{H}_0' & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_0 \\ \mathbf{u}_0 \\ \boldsymbol{\theta}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{c}_0 \end{pmatrix}.$$

$\boldsymbol{\beta}_a$ is a solution to GLS or mixed model equations with restrictions, $\mathbf{H}_a'\boldsymbol{\beta}_a = \mathbf{c}_a$ rather than $\mathbf{H}_0'\boldsymbol{\beta}_0 = \mathbf{c}_0$.

In case the alternative hypothesis is unrestricted ($\boldsymbol{\beta}$ can have any values), that is, $\boldsymbol{\beta}_a$ is a solution to the unrestricted GLS or mixed model equations. Under the null hypothesis (1) is distributed as χ^2 with $(m - a)$ degrees of freedom, m being the number of rows (independent) in $\mathbf{H}_0'$, and a being the number of rows (independent) in $\mathbf{H}_a'$. If the alternative hypothesis is unrestricted, $a = 0$. Having computed (1) this value is compared with values of χ^2_{m-a} for the chosen level of significance.

Let us illustrate with a model

$$\begin{aligned} y &= \mu + t_i + e_{ij} \\ \mu, t_i \text{ fixed}, \quad i &= 1, 2, 3 \\ \mathbf{R} &= Var(\mathbf{e}) = 5\mathbf{I}. \end{aligned}$$

Suppose that the number of observations on the levels of t_i are 4, 3, 2, and the treatment totals are 25, 15, 9 with individual observations, (6, 7, 8, 4, 4, 5, 6, 5, 4). We wish to test that the levels of t_i are equal, which can be expressed as

$$\begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} (\mu \;\; t_1 \;\; t_2 \;\; t_3)' = (0 \;\; 0)'.$$

We use as the alternative hypothesis the unrestricted hypothesis. The GLS equations under the restriction are

$$.2 \begin{pmatrix} 9 & 4 & 3 & 2 & 0 & 0 \\ 4 & 4 & 0 & 0 & 1 & 0 \\ 3 & 0 & 3 & 0 & 0 & 1 \\ 2 & 0 & 0 & 2 & -1 & -1 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_0 \\ \boldsymbol{\theta}_0 \end{pmatrix} = .2 \begin{pmatrix} 49 \\ 25 \\ 15 \\ 9 \\ 0 \\ 0 \end{pmatrix}.$$

A solution is

$$\boldsymbol{\beta}_o^{'} = (49 \;\; 0 \;\; 0 \;\; 0)/9, \boldsymbol{\theta}_o^{'} = (29 \;\; -12)/9.$$

The GLS equations with no restrictions are

$$.2 \begin{pmatrix} 9 & 4 & 3 & 2 \\ 4 & 4 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 2 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_a \end{pmatrix} = .2 \begin{pmatrix} 49 \\ 25 \\ 15 \\ 9 \end{pmatrix}.$$

A solution is $\boldsymbol{\beta}_a = (0 \;\; 25 \;\; 20 \;\; 18)/4$.

$$\begin{aligned} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_o)' &= (5 \;\; 14 \;\; 23 \;\; -13 \;\; -13 \;\; -4 \;\; 5 \;\; -4 \;\; -13)/9. \\ (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_o)'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_o) &= 146/45. \\ (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_a)' &= [-1, \; 3, \; 7, \; -9, \; -4, \; 0, \; 4, \; 2, \; -2]/4. \\ (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_a)'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_a) &= 9/4. \end{aligned}$$

The difference is $\frac{146}{45} - \frac{9}{4} = \frac{179}{180}$.

2.2 Differences between reductions

Two easier methods of computation that lead to the same result will now be presented. The first, described in Searle (1971b), is

$$\boldsymbol{\beta}_a^{'}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} + \boldsymbol{\theta}_a^{'}\mathbf{c}_a - \boldsymbol{\beta}_o^{'}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} - \boldsymbol{\theta}_o^{'}\mathbf{c}_o. \tag{2}$$

The first 2 terms are called reduction in sums of squares under the alternative hypothesis. The last two terms are the negative of the reduction in sum of squares under the null hypothesis. In our example

$$\begin{aligned} \boldsymbol{\beta}_a^{'}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} + \boldsymbol{\theta}_a^{'}\mathbf{c}_a &= 1087/20. \\ \boldsymbol{\beta}_o^{'}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} + \boldsymbol{\theta}_o^{'}\mathbf{c}_o &= 2401/45. \\ \frac{1087}{20} - \frac{2401}{45} &= \frac{179}{180} \text{ as before.} \end{aligned}$$

If the mixed model equations are used, (2) can be computed as

$$\beta_a^{'}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y} + \mathbf{u}_a^{'}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} + \theta_a^{'}\mathbf{c}_a - \beta_o^{'}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y} - \mathbf{u}_o^{'}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} - \theta_o^{'}\mathbf{c}_o. \quad (3)$$

2.3 Method based on variances of linear functions

A second easier method is

$$\begin{aligned}&(\mathbf{H}_o^{'}\beta^o - \mathbf{c}_o)' \ [\mathbf{H}_o^{'}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{H}_o]^{-1}(\mathbf{H}_o^{'}\beta^o - \mathbf{c}_o) \\ -&(\mathbf{H}_a^{'}\beta^o - \mathbf{c}_a)' \ [\mathbf{H}_a^{'}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{H}_a]^{-1}(\mathbf{H}_a^{'}\beta^o - \mathbf{c}_a). \quad (4)\end{aligned}$$

If $\mathbf{H}_a^{'}\beta$ is unrestricted the second term of (4) is set to 0. Remember that β^o is a solution in the unrestricted GLS equations. In place of $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}$ one can substitute the corresponding submatrix of a g-inverse of the mixed model coefficient matrix.

This is a convenient point to prove that an equivalent hypothesis, $\mathbf{P}(\mathbf{H}'\beta - \mathbf{c}) = \mathbf{0}$ gives the same result as $\mathbf{H}'\beta - \mathbf{c}$, remembering that $\mathbf{P}$ is non-singular. The quantity corresponding to (4) for $\mathbf{P}\ (\mathbf{H}'\beta - \mathbf{c})$ is

$$\begin{aligned}&(\mathbf{H}'\beta^o - \mathbf{c})'\mathbf{P}'[\mathbf{P}\mathbf{H}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{H}\mathbf{P}']^{-1}\mathbf{P}(\mathbf{H}'\beta - \mathbf{c}) \\ &= (\mathbf{H}'\beta^o - \mathbf{c})'\mathbf{P}'(\mathbf{P}')^{-1}[\mathbf{H}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{H}]^{-1}\mathbf{P}^{-1}\mathbf{P}(\mathbf{H}'\beta^o - \mathbf{c}) \\ &= (\mathbf{H}'\beta^o - \mathbf{c})'[\mathbf{H}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})\mathbf{H}]^{-1}(\mathbf{H}'\beta^o - \mathbf{c}),\end{aligned}$$

which proves the equality of the two equivalent hypotheses.

Let us illustrate (3) with our example

$$\begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \beta^o = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} (0 \ \ 25 \ \ 20 \ \ 18)'/4 = \begin{pmatrix} 7 \\ 2 \end{pmatrix}/4.$$

A g-inverse of $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 15 & 0 & 0 \\ 0 & 0 & 20 & 0 \\ 0 & 0 & 0 & 30 \end{pmatrix}/12.$$

$$\mathbf{H}_0^{'}\ (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{H}_0 = \begin{pmatrix} 45 & 30 \\ 30 & 50 \end{pmatrix}/12.$$

The inverse of this is

$$\begin{pmatrix} 20 & -12 \\ -12 & 18 \end{pmatrix} /45.$$

Then

$$\frac{1}{4}(7\ 2) \begin{pmatrix} 20 & -12 \\ -12 & 18 \end{pmatrix} \frac{1}{45} \begin{pmatrix} 7 \\ 2 \end{pmatrix} \frac{1}{4} = \frac{179}{180} \text{ as before.}$$

The d.f. for χ^2 are 2 because $\mathbf{H}_0'$ has 2 rows and the alternative hypothesis is unrestricted.

2.4 Comparison of reductions under reduced models

Another commonly used method is to compare reductions in sums of squares resulting from deletions of different subvectors of $\boldsymbol{\beta}$ from the reduction. The difficulty with this method is the determination of what hypothesis is tested by the difference between a pair of reductions. It is not true in general, as sometimes thought, that $Red(\boldsymbol{\beta}) - Red(\boldsymbol{\beta}_1)$ tests the hypothesis that $\boldsymbol{\beta}_2 = \mathbf{0}$, where $\boldsymbol{\beta}' = (\boldsymbol{\beta}_1' \ \ \boldsymbol{\beta}_2')$. In most designs, $\boldsymbol{\beta}_2$ is not estimable. We need to determine what $\mathbf{H}'\boldsymbol{\beta}$ imposed on a solution will give the same reduction in sum of squares as does $Red(\boldsymbol{\beta}_1)$.

In the latter case we solve

$$(\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_1) \ \ \boldsymbol{\beta}_1^o \ = \ \mathbf{X}_1'\mathbf{V}^{-1}\mathbf{y}$$

and then

$$\text{Reduction} = (\boldsymbol{\beta}_1^o)'\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{y}. \tag{5}$$

Consider a hypothesis, $\mathbf{H}'\boldsymbol{\beta}_2 = \mathbf{0}$. We could solve

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_1 & \mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_2 & \mathbf{0} \\ \mathbf{X}_2'\mathbf{V}^{-1}\mathbf{X}_1 & \mathbf{X}_2'\mathbf{V}^{-1}\mathbf{X}_2 & \mathbf{H} \\ \mathbf{0} & \mathbf{H}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1^o \\ \boldsymbol{\beta}_2^o \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{V}^{-1}\mathbf{y} \\ \mathbf{X}_2'\mathbf{V}^{-1}\mathbf{y} \\ \mathbf{0} \end{pmatrix}. \tag{6}$$

Then

$$\text{Reduction} \ = \ (\boldsymbol{\beta}_1^o)'\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{y} \ + \ (\boldsymbol{\beta}_2^o)'\mathbf{X}_2'\mathbf{V}^{-1}\mathbf{y}. \tag{7}$$

Clearly (7) is equal to (5) if a solution to (6) is $\boldsymbol{\beta}_2^o = \mathbf{0}$, for then

$$\boldsymbol{\beta}_1^o \ = \ (\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_1)^{-}\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{y}.$$

Consequently in order to determine what hypothesis is implied when $\boldsymbol{\beta}_2$ is deleted from the model, we need to find some $\mathbf{H}\,'\boldsymbol{\beta}_2 = \mathbf{0}$ such that a solution to (6) is $\boldsymbol{\beta}_2^o = \mathbf{0}$.

We illustrate with a two way fixed model with interaction. The numbers of observations per subclass are

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 2 & 5 \end{pmatrix}.$$

The subclass totals are

$$\begin{pmatrix} 6 & 2 & 2 \\ 3 & 5 & 9 \end{pmatrix}.$$

An analysis sometimes suggested is

$$Red(\mu, r, c) - Red(\mu, c) \text{ to test rows.}$$

$$Red(full\ model) - Red(\mu, r, c) \text{ to test interaction.}$$

The least squares equations are

$$\begin{pmatrix} 14 & 6 & 8 & 4 & 4 & 6 & 3 & 2 & 1 & 1 & 2 & 5 \\ & 6 & 0 & 3 & 2 & 1 & 3 & 2 & 1 & 0 & 0 & 0 \\ & & 8 & 1 & 2 & 5 & 0 & 0 & 0 & 1 & 2 & 5 \\ & & & 4 & 0 & 0 & 3 & 0 & 0 & 1 & 0 & 0 \\ & & & & 4 & 0 & 0 & 2 & 0 & 0 & 2 & 0 \\ & & & & & 6 & 0 & 0 & 1 & 0 & 0 & 5 \\ & & & & & & 3 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & & 2 & 0 & 0 & 0 & 0 \\ & & & & & & & & 1 & 0 & 0 & 0 \\ & & & & & & & & & 1 & 0 & 0 \\ & & & & & & & & & & 2 & 0 \\ & & & & & & & & & & & 5 \end{pmatrix} \boldsymbol{\beta}^o = \begin{pmatrix} 27 \\ 10 \\ 17 \\ 9 \\ 7 \\ 11 \\ 6 \\ 2 \\ 2 \\ 3 \\ 5 \\ 9 \end{pmatrix}$$

A solution to these equations is

$$[0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 2,\ 1,\ 2,\ 3,\ 2.5,\ 1.8],$$

which gives a reduction of 55.7, the full model reduction. A solution when interaction terms are deleted is

$$[1.9677,\ -.8065,\ 0,\ .8871,\ .1855,\ 0]$$

giving a reduction of 54.3468. This corresponds to an hypothesis,

$$\begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{pmatrix} \ \mathbf{rc} = \mathbf{0}.$$

When this is included as a Lagrange multiplier as in (6), a solution is

$$[1.9677,\ -.8065,\ 0,\ .8871,\ .1855,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ -.1452,\ -.6935].$$

Note that $(\mathbf{rc})^o = \mathbf{0}$, proving that dropping $\mathbf{rc}$ corresponds to the hypothesis stated above. The reduction again is 54.3468.

When $\mathbf{r}$ and $\mathbf{rc}$ are dropped from the equations, a solution is

$$[0,\ 2.25,\ 1.75,\ 1.8333]$$

giving a reduction of 52.6667. This corresponds to an hypothesis

$$\begin{pmatrix} 3 & -3 & 1 & 1 & 1 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ \mathbf{rc} \end{pmatrix} = \mathbf{0}.$$

When this is added as a Lagrange multiplier, a solution is

$$[2.25,\ 0,\ 0,\ 0,\ -.5,\ -.4167, 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ -.6944,\ -.05556,\ -.8056].$$

Note that $\mathbf{r}^o$ and $\mathbf{rc}^o$ are null, verifying the hypothesis. The reduction again is 52.6667. Then the tests are as follows:

Rows assuming **rc** non-existent = 54.3468 - 52.6667.

Interaction = 55.7 - 54.3468.

Chapter 5
Prediction of Random Variables

C. R. Henderson

1984 - Guelph

We have discussed estimation of $\boldsymbol{\beta}$, regarded as fixed. Now we shall consider a rather different problem, prediction of random variables, and especially prediction of $\mathbf{u}$. We can also formulate this problem as estimation of the realized values of random variables. These realized values are fixed, but they are the realization of values from some known population. This knowledge enables better estimates (smaller mean squared errors) to be obtained than if we ignore this information and estimate $\mathbf{u}$ by GLS. In genetics the predictors of $\mathbf{u}$ are used as selection criteria. Some basic results concerning selection are now presented.

Which is the more logical concept, prediction of a random variable or estimation of the realized value of a random variable? If we have an animal already born, it seems reasonable to describe the evaluation of its breeding value as an estimation problem. On the other hand, if we are interested in evaluating the potential breeding value of a mating between two potential parents, this would be a problem in prediction. If we are interested in future records, the problem is clearly one of prediction.

1 Best Prediction

Let $\hat{w} = f(\mathbf{y})$ be a predictor of the random variable w. Find $f(\mathbf{y})$ such that $E(\hat{w} - w)^2$ is minimum. Cochran (1951) proved that

$$f(\mathbf{y}) = E(w \mid \mathbf{y}). \tag{1}$$

This requires knowing the joint distribution of w and $\mathbf{y}$, being able to derive the conditional mean, and knowing the values of parameters appearing in the conditional mean. All of these requirements are seldom possible in practice.

Cochran also proved in his 1951 paper the following important result concerning selection. Let p individuals regarded as a random sample from some population as candidates for selection. The realized values of these individuals are $w_1, \ldots w_p$, not observable. We can observe $\mathbf{y}_i$, a vector of records on each. $(w_i, \mathbf{y}_i)$ are jointly distributed as $f(w, \mathbf{y})$ independent of $(w_j, \mathbf{y}_j)$. Some function, say $f(\mathbf{y}_i)$, is to be used as a selection criterion and the fraction, α, with highest $f(\mathbf{y}_i)$ is to be selected. What f will maximize the expectation of the mean of the associated w_i? Cochran proved that $E(w \mid \mathbf{y})$ accomplishes this goal. This is a very important result, but note that seldom if ever do the requirements of this theorem hold in animal breeding. Two obvious deficiencies suggest themselves. First, the candidates for selection have differing amounts of information (number of elements in $\mathbf{y}$ differ). Second, candidates are related and consequently the $\mathbf{y}_i$ are not independent and neither are the w_i.

Properties of best predictor

1. $E(\hat{w}_i) = E(w_i)$. (2)
2. $Var(\hat{w}_i - w_i) = Var(w \mid \mathbf{y})$ averaged over the distribution of $\mathbf{y}$. (3)
3. Maximizes $r_{\hat{w}w}$ for all functions of $\mathbf{y}$. (4)

2 Best Linear Prediction

Because we seldom know the form of distribution of $(\mathbf{y}, w)$, consider a linear predictor that minimizes the squared prediction error. Find $\hat{w} = \mathbf{a}'\mathbf{y} + b$, where $\mathbf{a}'$ is a vector and b a scalar such that $E(\hat{w} - w)^2$ is minimum. Note that in contrast to BP the form of distribution of $(\mathbf{y}, w)$ is not required. We shall see that the first and second moments are needed.

Let

$$\begin{aligned} E(w) &= \gamma, \\ E(\mathbf{y}) &= \boldsymbol{\alpha}, \\ Cov(\mathbf{y}, w) &= \mathbf{c}, \text{ and} \\ Var(\mathbf{y}) &= \mathbf{V}. \end{aligned}$$

Then

$$\begin{aligned} E(\mathbf{a}'\mathbf{y} + b - w)^2 = \ & \mathbf{a}'\mathbf{V}\mathbf{a} - 2\mathbf{a}'\mathbf{c} + \mathbf{a}'\boldsymbol{\alpha}\boldsymbol{\alpha}'\mathbf{a} + b^2 \\ & + 2\mathbf{a}'\boldsymbol{\alpha}\mathbf{b} - 2\mathbf{a}'\boldsymbol{\alpha}\gamma - 2b\gamma + Var(w) + \gamma^2. \end{aligned}$$

Differentiating this with respect to $\mathbf{a}$ and b and equating to 0

$$\begin{pmatrix} \mathbf{V} + \boldsymbol{\alpha}\boldsymbol{\alpha}' & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}' & 1 \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ b \end{pmatrix} = \begin{pmatrix} \mathbf{c} + \boldsymbol{\alpha}\gamma \\ \gamma \end{pmatrix}.$$

The solution is

$$\mathbf{a} = \mathbf{V}^{-1}\mathbf{c}, b = \gamma - \boldsymbol{\alpha}'\mathbf{V}^{-1}\mathbf{c}. \quad (5)$$

Thus

$$\hat{w} = \gamma + \mathbf{c}'\mathbf{V}^{-1}(\mathbf{y} - \boldsymbol{\alpha}).$$

Note that this is $E(w \mid \mathbf{y})$ when $\mathbf{y}, w$ are jointly normally distributed. Note also that BLP is the selection index of genetics. Sewall Wright (1931) and J.L. Lush (1931) were using this selection criterion prior to the invention of selection index by Fairfield Smith (1936). I think they were invoking the conditional mean under normality, but they were not too clear in this regard.

Other properties of BLP are unbiased, that is

$$E(\hat{w}) = E(w). \tag{6}$$

$$\begin{aligned} E(\hat{w}) &= E[\gamma + \mathbf{c}'\mathbf{V}^{-1}(\mathbf{y} - \boldsymbol{\alpha})] \\ &= \gamma + \mathbf{c}'\mathbf{V}^{-1}(\boldsymbol{\alpha} - \boldsymbol{\alpha}) \\ &= \gamma = E(w). \end{aligned}$$

$$Var(\hat{w}) = Var(\mathbf{c}'\mathbf{V}^{-1}\mathbf{y}) = \mathbf{c}'\mathbf{V}^{-1}\mathbf{V}\ \mathbf{V}^{-1}\mathbf{c} = \mathbf{c}'\mathbf{V}^{-1}\mathbf{c}. \tag{7}$$

$$Cov(\hat{w}, w) = \mathbf{c}'\mathbf{V}^{-1}Cov(\mathbf{y}, w) = \mathbf{c}'\mathbf{V}^{-1}\mathbf{c} = Var(\hat{w}) \tag{8}$$

$$Var(\hat{w} - w) = Var(w) - Var(\hat{w}) \tag{9}$$

In the class of linear functions of $\mathbf{y}$, BLP maximizes the correlation,

$$r_{\hat{w}w} = \mathbf{a}'\mathbf{c}/[\mathbf{a}'\mathbf{V}\mathbf{a}\ Var(w)]^{.5}. \tag{10}$$

Maximize $\log r$.

$$\log r = \log\ \mathbf{a}'\mathbf{c} - .5\ \log\ [\mathbf{a}'\mathbf{V}\mathbf{a}] - .5\ \log\ Var(w).$$

Differentiating with respect to $\mathbf{a}$ and equating to 0.

$$\frac{\mathbf{V}\mathbf{a}}{\mathbf{a}'\mathbf{V}\mathbf{a}} = \frac{\mathbf{c}}{\mathbf{a}'\mathbf{c}} \quad \text{or} \quad \mathbf{V}\mathbf{a} = \mathbf{c}\,\frac{Var(\hat{w})}{Cov(\hat{w}, w)}.$$

The ratio on the right does not affect r. Consequently let it be one. Then $\mathbf{a} = \mathbf{V}^{-1}\mathbf{c}$. Also the constant, b, does not affect the correlation. Consequently, BLP maximizes r.

BLP of $\mathbf{m}'\mathbf{w}$ is $\mathbf{m}'\hat{\mathbf{w}}$, where $\hat{\mathbf{w}}$ is BLP of $\mathbf{w}$. Now $\mathbf{w}$ is a vector with $E(\mathbf{w}) = \boldsymbol{\gamma}$ and $Cov(\mathbf{y}, \mathbf{w}') = \mathbf{C}$. Substitute the scalar, $\mathbf{m}'\mathbf{w}$ for w in the statement for BLP. Then BLP of

$$\begin{aligned} \mathbf{m}'\mathbf{w} &= \mathbf{m}'\boldsymbol{\gamma} + \mathbf{m}'\mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \boldsymbol{\alpha}) \\ &= \mathbf{m}'[\boldsymbol{\gamma} + \mathbf{C}\mathbf{V}^{-1}(\mathbf{y} - \boldsymbol{\alpha})] \\ &= \mathbf{m}'\hat{\mathbf{w}} \end{aligned} \tag{11}$$

because

$$\hat{\mathbf{w}} = \boldsymbol{\gamma} + \mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \boldsymbol{\alpha}).$$

In the multivariate normal case, BLP maximizes the probability of selecting the better of two candidates for selection, Henderson (1963). For fixed number selected, it maximizes the expectation of the mean of the selected u_i, Bulmer (1980).

It should be noted that when the distribution of $(\mathbf{y}, w)$ is multivariate normal, BLP is the mean of w given $\mathbf{y}$, that is, the conditional mean, and consequently is BP with its desirable properties as a selection criterion. Unfortunately, however, we probably never know the mean of $\mathbf{y}$, which is $\mathbf{X}\boldsymbol{\beta}$ in our mixed model. We may, however, know $\mathbf{V}$ accurately enough to assume that our estimate is the parameter value. This leads to the derivation of best linear unbiased prediction (BLUP).

3 Best Linear Unbiased Prediction

Suppose the predictand is the random variable, w, and all we know about it is that it has mean $\mathbf{k}'\boldsymbol{\beta}$, variance $= v$, and its covariance with $\mathbf{y}'$ is $\mathbf{c}'$. How should we predict w? One possibility is to find some linear function of $\mathbf{y}$ that has expectation, $\mathbf{k}'\boldsymbol{\beta}$ (is unbiased), and in the class of such predictors has minimum variance of prediction errors. This method is called best linear unbiased prediction (BLUP).

Let the predictor be $\mathbf{a}'\mathbf{y}$. The expectation of $\mathbf{a}'\mathbf{y} = \mathbf{a}'\mathbf{X}\boldsymbol{\beta}$, and we want to choose $\mathbf{a}$ so that the expectation of $\mathbf{a}'\mathbf{y}$ is $\mathbf{k}'\boldsymbol{\beta}$. In order for this to be true for any value of $\boldsymbol{\beta}$, it is seen that $\mathbf{a}'$ must be chosen so that

$$\mathbf{a}'\mathbf{X} = \mathbf{k}'. \tag{12}$$

Now the variance of the prediction error is

$$Var(\mathbf{a}'\mathbf{y} - w) = \mathbf{a}'\mathbf{V}\mathbf{a} - 2\mathbf{a}'\mathbf{c} + v. \tag{13}$$

Consequently, we minimize (13) subject to the condition of (12). The equations to be solved to accomplish this are

$$\begin{pmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{c} \\ \mathbf{k} \end{pmatrix}. \tag{14}$$

Note the similarity to (1) in Chapter 3, the equations for finding BLUE of $\mathbf{k}'\boldsymbol{\beta}$.

Solving for $\mathbf{a}$ in the first equation of (14),

$$\mathbf{a} = -\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\theta} + \mathbf{V}^{-1}\mathbf{c}. \tag{15}$$

Substituting this value of $\mathbf{a}$ in the second equation of (14)

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\theta} = -\mathbf{k} + \mathbf{X}'\mathbf{V}^{-1}\mathbf{c}.$$

Then, if the equations are consistent, and this will be true if and only if $\mathbf{k}'\boldsymbol{\beta}$ is estimable, a solution to $\boldsymbol{\theta}$ is

$$\boldsymbol{\theta} = -(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{k} + (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{c}.$$

Substituting the solution to $\boldsymbol{\theta}$ in (15) we find

$$\mathbf{a} = \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{k} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{c} + \mathbf{V}^{-1}\mathbf{c}. \tag{16}$$

Then the predictor is

$$\mathbf{a}'\mathbf{y} = \mathbf{k}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} + \mathbf{c}'\mathbf{V}^{-1}[\mathbf{y} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}]. \tag{17}$$

But because $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} = \boldsymbol{\beta}^o$, a solution to GLS equations, the predictor can be written as

$$\mathbf{k}'\boldsymbol{\beta}^o + \mathbf{c}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o). \tag{18}$$

This result was described by Henderson (1963) and a similar result by Goldberger (1962).

Note that if $\mathbf{k}'\boldsymbol{\beta} = 0$ and if $\boldsymbol{\beta}$ is known, the predictor would be $\mathbf{c}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$. This is the usual selection index method for predicting w. Thus BLUP is BLP with $\boldsymbol{\beta}^o$ substituted for $\boldsymbol{\beta}$.

4 Alternative Derivations Of BLUP

4.1 Translation invariance

We want to predict $\mathbf{m}'\mathbf{w}$ in the situation with unknown $\boldsymbol{\beta}$. But BLP, the minimum MSE predictor in the class of linear functions of $\mathbf{y}$, involves $\boldsymbol{\beta}$. Is there a comparable predictor that is invariant to $\boldsymbol{\beta}$?

Let the predictor be

$$\mathbf{a}'\mathbf{y} + b,$$

invariant to the value of $\boldsymbol{\beta}$. For translation invariance we require

$$\mathbf{a}'\mathbf{y} + b = \mathbf{a}'(\mathbf{y} + \mathbf{X}\mathbf{k}) + b$$

for any value of $\mathbf{k}$. This will be true if and only if $\mathbf{a}'\mathbf{X} = \mathbf{0}$. We minimize

$$E(\mathbf{a}'\mathbf{y} + b - \mathbf{m}'\mathbf{w})^2 = \mathbf{a}'\mathbf{V}\mathbf{a} - \mathbf{2}\mathbf{a}'\mathbf{C}\mathbf{m} + b^2 + \mathbf{m}'\mathbf{G}\mathbf{m}$$

when $\mathbf{a}'\mathbf{X} = 0$ and where $\mathbf{G} = Var(\mathbf{w})$. Clearly b must equal 0 because b^2 is positive. Minimization of $\mathbf{a}'\mathbf{V}\mathbf{a} - \mathbf{2}\mathbf{a}'\mathbf{C}\mathbf{m}$ subject to $\mathbf{a}'\mathbf{X} = 0$ leads immediately to predictor $\mathbf{m}'\mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)$, the BLUP predictor. Under normality BLUP has, in the class of invariant predictors, the same properties as those stated for BLP.

4.2 Selection index using functions of $\mathbf{y}$ with zero means

An interesting way to compute BLUP of $\mathbf{w}$ is the following. Compute $\boldsymbol{\beta}_* = \mathbf{L}'\mathbf{y}$ such that

$$E(\mathbf{X}\boldsymbol{\beta}_*) = \mathbf{X}\boldsymbol{\beta}.$$

Then compute

$$\begin{aligned} \mathbf{y}_* &= \mathbf{y} - \mathbf{X}\boldsymbol{\beta}_* \\ &= (\mathbf{I} - \mathbf{X}\mathbf{L}')\mathbf{y} \equiv \mathbf{T}'\mathbf{y}. \end{aligned}$$

Now

$$Var(\mathbf{y}_*) = \mathbf{T}'\mathbf{V}\mathbf{T} \equiv \mathbf{V}_*, \tag{19}$$

and

$$Cov(\mathbf{y}_*, \mathbf{w}') = \mathbf{T}'\mathbf{C} \equiv \mathbf{C}_*, \tag{20}$$

where $\mathbf{C} = Cov(\mathbf{y}, \mathbf{w}')$. Then selection index is

$$\hat{\mathbf{w}} = \mathbf{C}_*^{'}\mathbf{V}_*^{-}\mathbf{y}_*. \tag{21}$$

$$Var(\hat{\mathbf{w}}) = Cov(\hat{\mathbf{w}}, \mathbf{w}') = \mathbf{C}_*^{'}\mathbf{V}^{-}\mathbf{C}_*. \tag{22}$$

$$Var(\hat{\mathbf{w}} - \mathbf{w}) = Var(\mathbf{w}) - Var(\hat{\mathbf{w}}). \tag{23}$$

Now $\hat{\mathbf{w}}$ is invariant to choice of $\mathbf{T}$ and to the g-inverse of $\mathbf{V}_*$ that is computed. $\mathbf{V}_*$ has rank $= n - r$. One choice of $\boldsymbol{\beta}_*$ is OLS $= (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y}$. In that case $\mathbf{T} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$. $\boldsymbol{\beta}_*$ could also be computed as OLS of an appropriate subset of $\mathbf{y}$, with no fewer than r elements of $\mathbf{y}$.

Under normality,

$$\hat{\mathbf{w}} = E(\mathbf{w} \mid \mathbf{y}_*), \text{ and} \tag{24}$$

$$Var(\hat{\mathbf{w}} - \mathbf{w}) = Var(\mathbf{w} \mid \mathbf{y}_*). \tag{25}$$

5 Variance Of Prediction Errors

We now state some useful variances and covariances. Let a vector of predictands be $\mathbf{w}$. Let the variance-covariance matrix of the vector be $\mathbf{G}$ and its covariance with $\mathbf{y}$ be $\mathbf{C}'$. Then the predictor of $\mathbf{w}$ is

$$\hat{\mathbf{w}} = \mathbf{K}'\boldsymbol{\beta}^o + \mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o). \tag{26}$$

$$\begin{aligned} Cov(\hat{\mathbf{w}}, \mathbf{w}') &= \mathbf{K}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{C} + \mathbf{C}'\mathbf{V}^{-1}\mathbf{C} \\ &\quad - \mathbf{C}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{C}. \end{aligned} \tag{27}$$

$$
\begin{aligned}
Var(\hat{\mathbf{w}}) &= \mathbf{K}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{K} + \mathbf{C}'\mathbf{V}^{-1}\mathbf{C} \\
&\quad - \mathbf{C}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{C}. \qquad (28)
\end{aligned}
$$

$$
\begin{aligned}
Var(\hat{\mathbf{w}} - \mathbf{w}) &= Var(\mathbf{w}) - Cov(\hat{\mathbf{w}}, \mathbf{w}') - Cov(\mathbf{w}, \hat{\mathbf{w}}') + Var(\hat{\mathbf{w}}) \\
&= \mathbf{K}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{K} - \mathbf{K}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{C} \\
&\quad - \mathbf{C}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{K} + \mathbf{G} - \mathbf{C}'\mathbf{V}^{-1}\mathbf{C} \\
&\quad + \mathbf{C}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{C}. \qquad (29)
\end{aligned}
$$

6 Mixed Model Methods

The mixed model equations, (4) of Chapter 3, often provide an easy method to compute BLUP. Suppose the predictand, $\mathbf{w}$, can be written as

$$\mathbf{w} = \mathbf{K}'\boldsymbol{\beta} + \mathbf{u}, \qquad (30)$$

where $\mathbf{u}$ are the variables of the mixed model. Then it can be proved that

$$\text{BLUP of } \mathbf{w} = \text{BLUP of } \mathbf{K}'\boldsymbol{\beta} + \mathbf{u} = \mathbf{K}'\boldsymbol{\beta}^o + \hat{\mathbf{u}}, \qquad (31)$$

where $\boldsymbol{\beta}^o$ and $\hat{\mathbf{u}}$ are solutions to the mixed model equations. From the second equation of the mixed model equations,

$$\hat{\mathbf{u}} = (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o).$$

But it can be proved that

$$(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1} = \mathbf{C}'\mathbf{V}^{-1},$$

where $\mathbf{C} = \mathbf{ZG}$, and $\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$. Also $\boldsymbol{\beta}^o$ is a GLS solution. Consequently,

$$\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) = \mathbf{K}'\boldsymbol{\beta}^o + \hat{\mathbf{u}}.$$

From (24) it can be seen that

$$\text{BLUP of } \mathbf{u} = \hat{\mathbf{u}}. \qquad (32)$$

Proof that $(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1} = \mathbf{C}'\mathbf{V}^{-1}$ follows.

$$
\begin{aligned}
\mathbf{C}'\mathbf{V}^{-1} &= \mathbf{GZ}'\mathbf{V}^{-1} \\
&= \mathbf{GZ}'[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}] \\
&= \mathbf{G}[\mathbf{Z}'\mathbf{R}^{-1} - \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}] \\
&= \mathbf{G}[\mathbf{Z}'\mathbf{R}^{-1} - (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1} \\
&\quad + \mathbf{G}^{-1}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}] \\
&= \mathbf{G}[\mathbf{Z}'\mathbf{R}^{-1} - \mathbf{Z}'\mathbf{R}^{-1} + \mathbf{G}^{-1}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}] \\
&= (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}.
\end{aligned}
$$

This result was presented by Henderson (1963). The mixed model method of estimation and prediction can be formulated as Bayesian estimation, Dempfle (1977). This is discussed in Chapter 9.

7 Variances from Mixed Model Equations

A g-inverse of the coefficient matrix of the mixed model equations can be used to find needed variances and covariances. Let a g-inverse of the matrix of the mixed model equations be

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix} \tag{33}$$

Then

$$\begin{aligned}
Var(\mathbf{K}'\boldsymbol{\beta}^o) &= \mathbf{K}'\mathbf{C}_{11}\mathbf{K}. && (34)\\
Cov(\mathbf{K}'\boldsymbol{\beta}^o, \hat{\mathbf{u}}') &= \mathbf{0}. && (35)\\
Cov(\mathbf{K}'\boldsymbol{\beta}^o, \mathbf{u}') &= -\mathbf{K}'\mathbf{C}_{12}. && (36)\\
Cov(\mathbf{K}'\boldsymbol{\beta}^o, \hat{\mathbf{u}}' - \mathbf{u}') &= \mathbf{K}'\mathbf{C}_{12}. && (37)\\
Var(\hat{\mathbf{u}}) &= \mathbf{G} - \mathbf{C}_{22}. && (38)\\
Cov(\hat{\mathbf{u}}, \mathbf{u}') &= \mathbf{G} - \mathbf{C}_{22}. && (39)\\
Var(\hat{\mathbf{u}} - \mathbf{u}) &= \mathbf{C}_{22}. && (40)\\
Var(\hat{\mathbf{w}} - \mathbf{w}) &= \mathbf{K}'\mathbf{C}_{11}\mathbf{K} + \mathbf{K}'\mathbf{C}_{12} + \mathbf{C}_{12}'\mathbf{K} + \mathbf{C}_{22}. && (41)
\end{aligned}$$

These results were derived by Henderson (1975a).

8 Prediction Of Errors

The prediction of errors (estimation of the realized values) is simple. First, consider the model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ and the prediction of the entire error vector, $\boldsymbol{\varepsilon}$. From (18)

$$\hat{\boldsymbol{\varepsilon}} = \mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o),$$

but since $\mathbf{C}' = Cov(\boldsymbol{\varepsilon}, \mathbf{y}') = \mathbf{V}$, the predictor is simply

$$\begin{aligned}
\hat{\boldsymbol{\varepsilon}} &= \mathbf{V}\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)\\
&= \mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o. && (42)
\end{aligned}$$

To predict $\boldsymbol{\varepsilon}_{n+1}$, not in the model for $\mathbf{y}$, we need to know its covariance with $\mathbf{y}$. Suppose this is $\mathbf{c}'$. Then

$$\begin{aligned}
\boldsymbol{\varepsilon}_{n+1} &= \mathbf{c}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)\\
&= \mathbf{c}'\mathbf{V}^{-1}\hat{\boldsymbol{\varepsilon}}. && (43)
\end{aligned}$$

Next consider prediction of $\mathbf{e}$ from the mixed model. Now $Cov(\mathbf{e}, \mathbf{y}') = \mathbf{R}$. Then

$$\begin{aligned}
\hat{\mathbf{e}} &= \mathbf{RV}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) \\
&= \mathbf{R}[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}](\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o), \\
& \quad \text{from the result on } \mathbf{V}^{-1}, \\
&= [\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}](\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) \\
&= \mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o - \mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) \\
&= \mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o - \mathbf{Z}\hat{\mathbf{u}}. \qquad (44)
\end{aligned}$$

To predict e^{n+1}, not in the model for $\mathbf{y}$, we need the covariance between it and $\mathbf{e}$, say $\mathbf{c}'$. Then the predictor is

$$\hat{e}_{n+1} = \mathbf{c}'\mathbf{R}^{-1}\hat{\mathbf{e}}. \qquad (45)$$

We now define $\mathbf{e}' = [\mathbf{e}_p' \quad \mathbf{e}_m']$, where $\mathbf{e}_p$ refers to errors attached to $\mathbf{y}$ and $\mathbf{e}_m$ to future errors. Let

$$\begin{pmatrix} \mathbf{e}_p \\ \mathbf{e}_m \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{pp} & \mathbf{R}_{pm} \\ \mathbf{R}_{pm}' & \mathbf{R}_{mm} \end{pmatrix} \qquad (46)$$

Then

$$\hat{\mathbf{e}}_p = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o - \mathbf{Z}\hat{\mathbf{u}},$$

and

$$\hat{\mathbf{e}}_m = \mathbf{R}_{pm}'\mathbf{R}_{pp}^{-1}\hat{\mathbf{e}}_p.$$

Some prediction error variances and covariances follow.

$$Var(\hat{\mathbf{e}}_p \; - \; \mathbf{e}_p) = \mathbf{WCW}',$$

where

$$\mathbf{W} = [\mathbf{X} \;\; \mathbf{Z}], \mathbf{C} = \begin{pmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{pmatrix}$$

where $\mathbf{C}$ is the inverse of mixed model coefficient matrix, and $\mathbf{C}_1, \; \mathbf{C}_2$ have p,q rows respectively. Additionally,

$$\begin{aligned}
Cov[(\hat{\mathbf{e}}_p \; - \; \mathbf{e}_p), \; (\boldsymbol{\beta}^o)'\mathbf{K}] &= -\mathbf{WC}_1'\mathbf{K}, \\
Cov[(\hat{\mathbf{e}}_p \; - \; \mathbf{e}_p), \; (\hat{\mathbf{u}} - \mathbf{u})'] &= -\mathbf{WC}_2', \\
Cov[(\hat{\mathbf{e}}_p \; - \; \mathbf{e}_p), \; (\hat{\mathbf{e}}_m \; - \; \mathbf{e}_m)'] &= \mathbf{WCW}'\mathbf{R}_{pp}^{-1}\mathbf{R}_{pm}, \\
Var(\hat{\mathbf{e}}_m \; - \; \mathbf{e}_m) &= \mathbf{R}_{mm} - \mathbf{R}_{pm}'\mathbf{R}_{pp}^{-1}\mathbf{WCW}'\mathbf{R}_{pp}^{-1}\mathbf{R}_{pm}, \\
Cov[(\hat{\mathbf{e}}_m \; - \; \mathbf{e}_m), \; (\boldsymbol{\beta}^o)'\mathbf{K}] &= -\mathbf{R}_{pm}'\mathbf{R}_{pp}^{-1}\mathbf{WC}_1'\mathbf{K}, \text{ and} \\
Cov[(\hat{\mathbf{e}}_m \; - \; \mathbf{e}_m), \; (\hat{\mathbf{u}} - \mathbf{u})'] &= -\mathbf{R}_{pm}'\mathbf{R}_{pp}^{-1}\mathbf{WC}_2'.
\end{aligned}$$

9 Prediction Of Missing u

Three simple methods exist for prediction of a $\mathbf{u}$ vector not in the model, say $\mathbf{u}_n$.

$$\hat{\mathbf{u}}_n = \mathbf{B}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) \tag{47}$$

where $\mathbf{B}'$ is the covariance between $\mathbf{u}_n$ and $\mathbf{y}'$. Or

$$\hat{\mathbf{u}}_n \;=\; \mathbf{C}'\mathbf{G}^{-1}\hat{\mathbf{u}}, \tag{48}$$

where $\mathbf{C}' = Cov(\mathbf{u}_n, \mathbf{u}')$, $\mathbf{G} = Var(\mathbf{u})$, and $\hat{\mathbf{u}}$ is BLUP of $\mathbf{u}$. Or write expanded mixed model equations as follows:

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{0} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{0} & \mathbf{W}_{12}' & \mathbf{W}_{22} \end{pmatrix} \left(\begin{matrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \hat{\mathbf{u}}_n \end{matrix} \right] = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{0} \end{pmatrix}, \tag{49}$$

where

$$\begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{12}' & \mathbf{W}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{G} & \mathbf{C} \\ \mathbf{C}' & \mathbf{G}_n \end{pmatrix}^{-1}$$

and $\mathbf{G} = Var(\mathbf{u}),\ \mathbf{C} = Cov(\mathbf{u},\ \mathbf{u}_n'),\ \mathbf{G}_n = Var(\mathbf{u}_n)$. The solution to (49) gives the same results as before when $\mathbf{u}_n$ is ignored. The proofs of these results are in Henderson (1977a).

10 Prediction When G Is Singular

The possibility exists that $\mathbf{G}$ is singular. This could be true in an additive genetic model with one or more pairs of identical twins. This poses no problem if one uses the method $\hat{\mathbf{u}} = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^0)$, but the mixed model method previously described cannot be used since $\mathbf{G}^{-1}$ is required. A modification of the mixed model equations does permit a solution to $\boldsymbol{\beta}^o$ and $\hat{\mathbf{u}}$. One possibility is to solve the following.

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} \tag{50}$$

The coefficient matrix has rank, $r + q$. Then $\boldsymbol{\beta}^o$ is a GLS solution to $\boldsymbol{\beta}$, and $\hat{\mathbf{u}}$ is BLUP of $\mathbf{u}$. Note that the coefficient matrix above is not symmetric. Further, a g-inverse of it does not yield sampling variances. For this we proceed as follows. Compute $\mathbf{C}$, some g-inverse of the matrix. Then

$$\mathbf{C} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{pmatrix}$$

has the same properties as the g-inverse in (33).

If we want a symmetric coefficient matrix we can modify the equations of (50) as follows.

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{ZG} \\ \mathbf{GZ}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{GZ}'\mathbf{R}^{-1}\mathbf{ZG}+\mathbf{G} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\boldsymbol{\alpha}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{GZ}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} \quad (51)$$

This coefficient matrix has rank, $r+$ rank $(\mathbf{G})$. Solve for $\boldsymbol{\beta}^o$, $\hat{\boldsymbol{\alpha}}$. Then

$$\hat{\mathbf{u}} = \mathbf{G}\hat{\boldsymbol{\alpha}}.$$

Let $\mathbf{C}$ be a g-inverse of the matrix of (51). Then

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{pmatrix} \mathbf{C} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{pmatrix}$$

has the properties of (33).

These results on singular $\mathbf{G}$ are due to Harville (1976). These two methods for singular $\mathbf{G}$ can also be used for nonsingular $\mathbf{G}$ if one wishes to avoid inverting $\mathbf{G}$, Henderson (1973).

11 Examples of Prediction Methods

Let us illustrate some of these prediction methods. Suppose

$$\mathbf{X}' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 3 & 4 \end{pmatrix}, \ \mathbf{Z}' = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix},$$

$$\mathbf{G} = \begin{pmatrix} 3 & 2 & 1 \\ & 4 & 1 \\ & & 5 \end{pmatrix}, \ \mathbf{R} = 9\mathbf{I}, \ \mathbf{y}' = (5,\ 3,\ 6,\ 7,\ 5).$$

By the basic GLS and BLUP methods

$$\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R} = \begin{pmatrix} 12 & 3 & 2 & 1 & 1 \\ & 12 & 2 & 1 & 1 \\ & & 13 & 1 & 1 \\ & & & 14 & 5 \\ & & & & 14 \end{pmatrix}.$$

Then the GLS equations, $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}^o = \mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ are

$$\begin{pmatrix} .249211 & .523659 \\ .523659 & 1.583100 \end{pmatrix} \boldsymbol{\beta}^o = \begin{pmatrix} 1.280757 \\ 2.627792 \end{pmatrix}.$$

The inverse of the coefficient matrix is

$$\begin{pmatrix} 13.1578 & -4.3522 \\ -4.3522 & 2.0712 \end{pmatrix},$$

and the solution to $\boldsymbol{\beta}^o$ is $[5.4153 \quad -.1314]'$. To predict $\mathbf{u}$,

$$\begin{aligned}
\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o &= \begin{pmatrix} -.2839 \\ -2.1525 \\ .7161 \\ 1.9788 \\ .1102 \end{pmatrix}, \\
\mathbf{G}\mathbf{Z}'\mathbf{V}^{-1} &= \begin{pmatrix} .1838 & .1838 & .0929 & .0284 & .0284 \\ .0929 & .0929 & .2747 & .0284 & .0284 \\ .0284 & .0284 & .0284 & .2587 & .2587 \end{pmatrix}, \\
\hat{\mathbf{u}} &= \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) \;=\; \begin{pmatrix} -.3220 \\ .0297 \\ .4915 \end{pmatrix}, \\
Cov(\boldsymbol{\beta}^o,\ \hat{\mathbf{u}}' - \mathbf{u}') &= -(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}\mathbf{G} \\
&= \begin{pmatrix} -3.1377 & -3.5333 & .4470 \\ .5053 & .6936 & -1.3633 \end{pmatrix}, \text{and} \\
Var(\hat{\mathbf{u}} - \mathbf{u}) &= \mathbf{G} - \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}\mathbf{Z}\mathbf{G} + \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}\mathbf{G} \\
&= \begin{pmatrix} 3 & 2 & 1 \\ & 4 & 1 \\ & & 5 \end{pmatrix} - \begin{pmatrix} 1.3456 & 1.1638 & .7445 \\ & 1.5274 & .7445 \\ & & 2.6719 \end{pmatrix} \\
&\quad + \begin{pmatrix} 1.1973 & 1.2432 & .9182 \\ & 1.3063 & .7943 \\ & & 2.3541 \end{pmatrix} \\
&= \begin{pmatrix} 2.8517 & 2.0794 & 1.1737 \\ & 3.7789 & 1.0498 \\ & & 4.6822 \end{pmatrix}.
\end{aligned}$$

The mixed model method is considerably easier.

$$\begin{aligned}
\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} &= \begin{pmatrix} .5556 & 1.2222 \\ 1.2222 & 3.4444 \end{pmatrix}, \\
\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} &= \begin{pmatrix} .2222 & .1111 & .2222 \\ .3333 & .1111 & .7778 \end{pmatrix}, \\
\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} &= \begin{pmatrix} .2222 & 0 & 0 \\ & .1111 & 0 \\ & & .2222 \end{pmatrix},
\end{aligned}$$

$$\mathbf{X}'\mathbf{R}^{-1}\mathbf{y} = \begin{pmatrix} 2.8889 \\ 6.4444 \end{pmatrix}, \ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} = \begin{pmatrix} .8889 \\ .6667 \\ 1.3333 \end{pmatrix},$$

$$\mathbf{G}^{-1} = \begin{pmatrix} .5135 & -.2432 & -.0541 \\ & .3784 & -.0270 \\ & & .2162 \end{pmatrix}.$$

Then the mixed model equations are

$$\begin{pmatrix} .5556 & 1.2222 & .2222 & .1111 & .2222 \\ & 3.4444 & .3333 & .1111 & .7778 \\ & & .7357 & -.2432 & -.0541 \\ & & & .4895 & -.0270 \\ & & & & .4384 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} 2.8889 \\ 6.4444 \\ .8889 \\ .6667 \\ 1.3333 \end{pmatrix}.$$

A g-inverse (regular inverse) is

$$\begin{pmatrix} 13.1578 & -4.3522 & -3.1377 & -3.5333 & .4470 \\ & 2.0712 & .5053 & .6936 & -1.3633 \\ & & 2.8517 & 2.0794 & 1.1737 \\ & & & 3.7789 & 1.0498 \\ & & & & 4.6822 \end{pmatrix}.$$

The upper 2 x 2 represents $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^-$, the upper 2 x 3 represents $Cov(\boldsymbol{\beta}^o, \ \hat{\mathbf{u}}' - \mathbf{u}')$, and the lower 3 x 3 $Var(\hat{\mathbf{u}} - \mathbf{u})$. These are the same results as before. The solution is (5.4153, −.1314, −.3220, .0297, .4915) as before.

Now let us illustrate with singular **G**. Let the data be the same as before except

$$\mathbf{G} = \begin{pmatrix} 2 & 1 & 3 \\ & 3 & 4 \\ & & 7 \end{pmatrix}.$$

Note that the 3rd row of **G** is the sum of the first 2 rows. Now

$$\mathbf{V} = \begin{pmatrix} 11 & 2 & 1 & 3 & 3 \\ & 11 & 1 & 3 & 3 \\ & & 12 & 4 & 4 \\ & & & 16 & 7 \\ & & & & 16 \end{pmatrix},$$

and

$$\mathbf{V}^{-1} = \begin{pmatrix} .0993 & -.0118 & .0004 & -.0115 & -.0115 \\ & .0993 & .0004 & -.0115 & -.0115 \\ & & .0943 & -.0165 & -.0165 \\ & & & .0832 & -.0280 \\ & & & & .0832 \end{pmatrix}.$$

The GLS equations are

$$\begin{pmatrix} .2233 & .3670 \\ & 1.2409 \end{pmatrix} \boldsymbol{\beta}^o = \begin{pmatrix} 1.0803 \\ 1.7749 \end{pmatrix}.$$

$$(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} = \begin{pmatrix} 8.7155 & -2.5779 \\ & 1.5684 \end{pmatrix}.$$

$$\boldsymbol{\beta}^o = \begin{pmatrix} 4.8397 \\ -.0011 \end{pmatrix}.$$

$$\hat{\mathbf{u}} = \begin{pmatrix} .1065 & .1065 & -.0032 & .1032 & .1032 \\ -.0032 & -.0032 & .1516 & .1484 & .1484 \\ .1033 & .1033 & .1484 & .2516 & .2516 \end{pmatrix} \begin{pmatrix} .1614 \\ -1.8375 \\ 1.1614 \\ 2.1636 \\ .1648 \end{pmatrix} = \begin{pmatrix} 0582 \\ .5270 \\ .5852 \end{pmatrix}.$$

Note that $\hat{u}_3 = \hat{u}_1 + \hat{u}_2$ as a consequence of the linear dependencies in $\mathbf{G}$.

$$Cov(\boldsymbol{\beta}^o, \hat{\mathbf{u}}' - \mathbf{u}') = \begin{pmatrix} -.9491 & -.8081 & -1.7572 \\ -.5564 & -.7124 & -1.2688 \end{pmatrix}.$$

$$Var(\hat{\mathbf{u}} - \mathbf{u}) = \begin{pmatrix} 1.9309 & 1.0473 & 2.9782 \\ & 2.5628 & 3.6100 \\ & & 6.5883 \end{pmatrix}.$$

By the modified mixed model methods

$$\mathbf{GZ}'\mathbf{R}^{-1}\mathbf{X} = \begin{pmatrix} 1.2222 & 3.1111 \\ 1.4444 & 3.7778 \\ 2.6667 & 6.8889 \end{pmatrix},$$

$$\mathbf{GZ}'\mathbf{R}^{-1}\mathbf{Z} = \begin{pmatrix} .4444 & .1111 & .6667 \\ .2222 & .3333 & .8889 \\ .6667 & .4444 & 1.5556 \end{pmatrix},$$

$$\mathbf{GZ}'\mathbf{R}^{-1}\mathbf{y} = \begin{pmatrix} 6.4444 \\ 8.2222 \\ 14.6667 \end{pmatrix}, \quad \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} = \begin{pmatrix} 2.8889 \\ 6.4444 \end{pmatrix}.$$

Then the non-symmetric mixed model equations (50) are

$$\begin{pmatrix} .5556 & 1.2222 & .2222 & .1111 & .2222 \\ 1.2222 & 3.4444 & .3333 & .1111 & .7778 \\ 1.2222 & 3.1111 & 1.4444 & .1111 & .6667 \\ 1.4444 & 3.7778 & .2222 & 1.3333 & .8889 \\ 2.6667 & 6.8889 & .6667 & .4444 & 2.5556 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} 2.8889 \\ 6.4444 \\ 6.4444 \\ 8.2222 \\ 14.6667 \end{pmatrix}.$$

The solution is (4.8397, −.0011, .0582, .5270, .5852) as before. The inverse of the coefficient matrix is

$$\begin{pmatrix} 8.7155 & -2.5779 & -.8666 & -.5922 & .4587 \\ -2.5779 & 1.5684 & .1737 & .1913 & -.3650 \\ -.9491 & -.5563 & .9673 & .0509 & -.0182 \\ -.8081 & -.7124 & .1843 & .8842 & -.0685 \\ -1.7572 & -1.2688 & .1516 & -.0649 & .9133 \end{pmatrix}.$$

Post multiplying this matrix by $\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{pmatrix}$ gives

$$\begin{pmatrix} 8.7155 & -2.5779 & -.9491 & -.8081 & -1.7572 \\ & 1.5684 & -.5563 & -.7124 & -1.2688 \\ & & 1.9309 & 1.0473 & 2.9782 \\ & & & 2.5628 & 3.6101 \\ & & & & 6.5883 \end{pmatrix}.$$

These yield the same variances and covariances as before. The analogous symmetric equations (51) are

$$\begin{pmatrix} .5556 & 1.2222 & 1.2222 & 1.4444 & 2.6667 \\ & 3.4444 & 3.1111 & 3.7778 & 6.8889 \\ & & 5.0 & 4.4444 & 9.4444 \\ & & & 7.7778 & 12.2222 \\ & & & & 21.6667 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\boldsymbol{\alpha}} \end{pmatrix} = \begin{pmatrix} 2.8889 \\ 6.4444 \\ 6.4444 \\ 8.2222 \\ 14.6667 \end{pmatrix}.$$

A solution is [4.8397, −.0011, −.2697, 0, .1992]. Premultiplying $\hat{\boldsymbol{\alpha}}$ by $\mathbf{G}$ we obtain $\hat{\mathbf{u}}' = (.0582,\ .5270,\ .5852)$ as before.

A g-inverse of the matrix is

$$\begin{pmatrix} 8.7155 & -2.5779 & -.2744 & 0 & -.1334 \\ & 1.5684 & -.0176 & 0 & -.1737 \\ & & 1.1530 & 0 & -.4632 \\ & & & 0 & 0 \\ & & & & .3197 \end{pmatrix}.$$

Pre-and post-multiplying this matrix by $\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{pmatrix}$, yields the same matrix as post-multiplying the non-symmetric inverse by $\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{pmatrix}$ and consequently we have the required matrix for variances and covariances.

12 Illustration Of Prediction Of Missing u

We illustrate prediction of random variables not in the model for $\mathbf{y}$ by a multiple trait example. Suppose we have 2 traits and 2 animals, the first 2 with measurements on traits 1 and 2, but the third with a record only on trait 1. We assume an additive genetic model and wish to predict breeding value of both traits on all 3 animals and also to predict the second trait of animal 3. The numerator relationship matrix for the 3 animals is

$$\begin{pmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/4 \\ 1/2 & 1/4 & 1 \end{pmatrix}.$$

The additive genetic variance-covariance and error covariance matrices are assumed to be $\mathbf{G}_0$ and $\mathbf{R}_0 = \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix}$ and $\begin{pmatrix} 4 & 1 \\ 1 & 5 \end{pmatrix}$, respectively. The records are ordered animals in traits and are [6, 8, 7, 9, 5]. Assume

$$\mathbf{X}' = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

If all 6 elements of $\mathbf{u}$ are included

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

If the last (missing $\mathbf{u}_6$) is not included delete the last column from $\mathbf{Z}$. When all $\mathbf{u}$ are included

$$\mathbf{G} = \begin{pmatrix} \mathbf{A}g_{11} & \mathbf{A}g_{12} \\ \mathbf{A}g_{12} & \mathbf{A}g_{22} \end{pmatrix},$$

where g_{ij} is the ij^{th} element of $\mathbf{G}_0$, the genetic variance-covariance matrix. Numerically this is

$$\begin{pmatrix} 2 & 1 & 1 & 2 & 1 & 1 \\ & 2 & .5 & 1 & 2 & .5 \\ & & 2 & 1 & .5 & 2 \\ & & & 3 & 1.5 & 1.5 \\ & & & & 3 & .75 \\ & & & & & 3 \end{pmatrix}.$$

If u_6 is not included, delete the 6^{th} row and column from $\mathbf{G}$.

$$\mathbf{R} = \begin{pmatrix} 4 & 0 & 0 & 1 & 0 \\ & 4 & 0 & 0 & 1 \\ & & 4 & 0 & 0 \\ & & & 5 & 0 \\ & & & & 5 \end{pmatrix}.$$

$$\mathbf{R}^{-1} = \begin{pmatrix} .2632 & 0 & 0 & -.0526 & 0 \\ & .2632 & 0 & 0 & -.0526 \\ & & .25 & 0 & 0 \\ & & & .2105 & 0 \\ & & & & .2105 \end{pmatrix}.$$

$\mathbf{G}^{-1}$ for the first 5 elements of $\mathbf{u}$ is

$$\begin{pmatrix} 2.1667 & -1. & -.3333 & -1.3333 & .6667 \\ & 2. & 0 & .6667 & -1.3333 \\ & & .6667 & 0 & 0 \\ & & & 1.3333 & -.6667 \\ & & & & 1.3333 \end{pmatrix}.$$

Then the mixed model equations for $\boldsymbol{\beta}^o$ and $\hat{u}_1, \ldots, \hat{u}_5$ are

$$\begin{pmatrix} .7763 & -.1053 & .2632 & .2632 & .25 & -.0526 & -.0526 \\ & .4211 & -.0526 & -.0526 & 0 & .2105 & .2105 \\ & & 2.4298 & -1. & -.3333 & -1.3860 & .6667 \\ & & & 2.2632 & 0 & .6667 & -1.3860 \\ & & & & 9167 & 0 & 0 \\ & & & & & 1.5439 & -.6667 \\ & & & & & & 1.5439 \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{u}_4 \\ \hat{u}_5 \end{pmatrix}$$

$$= (4.70, 2.21, 1.11, 1.84, 1.75, 1.58, .63)'.$$

The solution is (6.9909, 6.9959, .0545, -.0495, .0223, .2651, -.2601).

To predict u_6 we can use $\hat{u}_1, \ldots, \hat{u}_5$. The solution is

$$\begin{aligned} \hat{u}_6 &= [1 \;\; .5 \;\; 2 \;\; 1.5 \;\; .75] \begin{pmatrix} 2 & 1 & 1 & 2 & 1 \\ & 2 & .5 & 1 & 2 \\ & & 2 & 1 & .5 \\ & & & 3 & 1.5 \\ & & & & 3 \end{pmatrix}^{-1} \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{u}_4 \\ \hat{u}_5 \end{pmatrix} \\ &= .1276. \end{aligned}$$

We could have solved directly for $\hat{u}_6$ in mixed model equations as follows.

$$\begin{pmatrix} .7763 & -.1053 & .2632 & .2632 & .25 & .0526 & -.0526 & 0 \\ & .4211 & -.0526 & -.0526 & 0 & .2105 & .2105 & 0 \\ & & 2.7632 & -1. & -1. & -1.7193 & .6667 & .6667 \\ & & & 2.2632 & 0 & .6667 & -1.3860 & 0 \\ & & & & 2.25 & .6667 & 0 & -1.3333 \\ & & & & & 1.8772 & .6667 & -.6667 \\ & & & & & & 1.5439 & 0 \\ & & & & & & & 1.3333 \end{pmatrix}$$

$$\begin{pmatrix} \hat{\boldsymbol\beta} \\ \hat{\mathbf{u}} \end{pmatrix} = [4.70,\ 2.21,\ 1.11,\ 1.84,\ 1.75,\ 1.58,\ .63,\ 0]'.$$

The solution is (6.9909, 6.9959, .0545, -.0495, .0223, .2651, -.2601, .1276), and equals the previous solution.

The predictor of the record on the second trait on animal 3 is some new $\hat{\beta}_2 + \hat{u}_6 + \hat{e}_6$. We already have $\hat{u}_6$. We can predict $\hat{e}_6$ from $\hat{e}_1 \ldots \hat{e}_5$.

$$\begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \hat{e}_3 \\ \hat{e}_4 \\ \hat{e}_5 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{u}_4 \\ \hat{u}_5 \end{pmatrix} = \begin{pmatrix} -1.0454 \\ 1.0586 \\ -.0132 \\ 1.7391 \\ -1.7358 \end{pmatrix}.$$

Then $\hat{e}_6 = (0\ 0\ 1\ 0\ 0)\, \mathbf{R}^{-1}\, (\hat{e}_1 \ldots \hat{e}_5)' = -.0033$. The column vector above is Cov $[e_6,\ (e_1\ e_2\ e_3\ e_4\ e_4\ e_5)]$. $\mathbf{R}$ above is $Var[(e_1 \ldots \hat{e}_5)']$.

Suppose we had the same model as before but we have no data on the second trait. We want to predict breeding values for both traits in the 3 animals, that is, $u_1, \ldots, u_6$. We also want to predict records on the second trait, that is, $u_4 + e_4, u_5 + e_5, u_6 + e_6$. The mixed model equations are

$$\begin{pmatrix} .75 & .25 & .25 & .25 & 0 & 0 & 0 \\ & 2.75 & -1. & -1. & -1.6667 & .6667 & .6667 \\ & & 2.25 & 0 & .6667 & -1.3333 & 0 \\ & & & 2.25 & .6667 & 0 & -1.3333 \\ & & & & 1.6667 & -.6667 & -.6667 \\ & & & & & 1.3333 & 0 \\ & & & & & & 1.3333 \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{u}_4 \\ \hat{u}_5 \\ \hat{u}_6 \end{pmatrix} = \begin{pmatrix} 5.25 \\ 1.50 \\ 2.00 \\ 1.75 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The solution is

$$[7.0345, -.2069, .1881, -.0846, -.2069, .1881, -.0846].$$

The last 6 values represent prediction of breeding values.

$$\begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \hat{e}_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} - (\mathbf{X}\ \mathbf{Z}) \begin{pmatrix} \hat{\beta} \\ \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \end{pmatrix} = \begin{pmatrix} -.8276 \\ .7774 \\ .0502 \end{pmatrix}.$$

Then

$$\begin{pmatrix} \hat{e}_4 \\ \hat{e}_5 \\ \hat{e}_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}^{-1} \begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \hat{e}_3 \end{pmatrix} = \begin{pmatrix} -.2069 \\ .1944 \\ .0125 \end{pmatrix}.$$

Then predictions of second trait records are

$$\beta_2 + \begin{pmatrix} -.2069 \\ .1881 \\ -.0846 \end{pmatrix} + \begin{pmatrix} -.2069 \\ .1944 \\ .0125 \end{pmatrix},$$

but β_2 is unknown.

13 A Singular Submatrix In G

Suppose that $\mathbf{G}$ can be partitioned as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{22} \end{pmatrix}$$

such that $\mathbf{G}_{11}$ is non-singular and $\mathbf{G}_{22}$ is singular. A corresponding partition of $\mathbf{u}'$ is $(\mathbf{u}_1' \;\; \mathbf{u}_2')$. Then two additional methods can be used. First, solve (52)

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_2 \\ \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{Z}_1 + \mathbf{G}_{11}^{-1} & \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{Z}_2 \\ \mathbf{G}_{22}\mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{X} & \mathbf{G}_{22}\mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{G}_{22}\mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{Z}_2 + \mathbf{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}}_1 \\ \hat{\mathbf{u}}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{G}_{22}\mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{52}$$

Let a g-inverse of this matrix be $\mathbf{C}$. Then the prediction errors come from

$$\mathbf{C} \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_{22} \end{pmatrix}. \tag{53}$$

The symmetric counterpart of these equations is

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_2\mathbf{G}_{22} \\ \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{Z}_1 + \mathbf{G}_{11}^{-1} & \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{Z}_2\mathbf{G}_{22} \\ \mathbf{G}_{22}\mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{X} & \mathbf{G}_{22}\mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{G}_{22}\mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{Z}_2\mathbf{G}_{22} + \mathbf{G}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}}_1 \\ \hat{\boldsymbol{\alpha}}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{G}_{22}\mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}, \tag{54}$$

and $\hat{\mathbf{u}}_2 = \mathbf{G}_{22}\hat{\boldsymbol{\alpha}}_2$.

Let $\mathbf{C}$ be a g-inverse of the coefficient matrix of (54). Then the variances and covariances come from

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_{22} \end{pmatrix} \mathbf{C} \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_{22} \end{pmatrix}. \tag{55}$$

14 Prediction Of Future Records

Most applications of genetic evaluation are essentially problems in prediction of future records, or more precisely, prediction of the relative values of future records, the relativity arising from the fact that we may have no data available for estimation of future $\mathbf{X}\boldsymbol{\beta}$, for example, a year effect for some record in a future year. Let the model for a future record be

$$y_i = \mathbf{x}_i^{'}\boldsymbol{\beta} + \mathbf{z}_i^{'}\mathbf{u} + e_i. \tag{56}$$

Then if we have available BLUE of $\mathbf{x}_i^{'}\boldsymbol{\beta} = \mathbf{x}_i^{'}\boldsymbol{\beta}^o$ and BLUP of $\mathbf{u}$ and e_i, $\hat{\mathbf{u}}$ and $\hat{e}_i$, BLUP of this future record is

$$\mathbf{x}_i^{'}\boldsymbol{\beta}^o + \mathbf{z}_i^{'}\hat{\mathbf{u}} + \hat{e}_i.$$

Suppose however that we have information on only a subvector of $\boldsymbol{\beta}$ say $\boldsymbol{\beta}_2$. Write the model for a future record as

$$\mathbf{x}_{1i}^{'}\boldsymbol{\beta}_1 + \mathbf{x}_{2i}^{'}\boldsymbol{\beta}_2 + \mathbf{z}_i^{'}\mathbf{u} + e_i.$$

Then we can assert BLUP for only

$$\mathbf{x}_{2i}^{'}\boldsymbol{\beta}_2 + \mathbf{z}_2^{\prime}\mathbf{u} + e_i.$$

But if we have some other record we wish to compare with this one, say y_j, with model,

$$y_j = \mathbf{x}_{1j}^{'}\boldsymbol{\beta}_1 + \mathbf{x}_{2j}^{'}\boldsymbol{\beta}_2 + \mathbf{z}_j^{'}\mathbf{u} + e_j,$$

we can compute BLUP of $y_i - y_j$ provided that

$$\mathbf{x}_{1i} = \mathbf{x}_{1j}.$$

It should be remembered that the variance of the error of prediction of a future record (or linear function of a set of records) should take into account the variance of the error of prediction of the error (or linear combination of errors) and also its covariance with $\boldsymbol{\beta}^o$ and $\hat{\mathbf{u}}$. See Section 8 for these variances and covariances. An extensive discussion of prediction of future records is in Henderson (1977b).

15 When Rank of MME Is Greater Than n

In some genetic problems, and in particular individual animal multiple trait models, the order of the mixed model coefficient matrix can be much greater than n, the number of observations. In these cases one might wish to consider a method described in this section, especially if one can thereby store and invert the coefficient matrix in cases when the mixed model equations are too large for this to be done. Solve equations (57) for $\boldsymbol{\beta}^o$ and $\mathbf{s}$.

$$\begin{pmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{s} \\ \boldsymbol{\beta}^o \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}. \tag{57}$$

Then $\boldsymbol{\beta}^o$ is a GLS solution and

$$\hat{\mathbf{u}} = \mathbf{GZ}'\mathbf{s} \tag{58}$$

is BLUP of $\mathbf{u}$. It is easy to see why these are true. Eliminate $\mathbf{s}$ from equations (57). This gives

$$-(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})\boldsymbol{\beta}^o = -\mathbf{X}'\mathbf{V}^{-1}\mathbf{y},$$

which are the GLS equations. Solving for $\mathbf{s}$ in (57) we obtain

$$\mathbf{s} = \mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o).$$

Then $\mathbf{GZ}'\mathbf{s} = \mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)$, which we know to be BLUP of $\mathbf{u}$.

Some variances and covariances from a g-inverse of the matrix of (57) are shown below. Let a g-inverse be

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix}.$$

Then

$$\begin{aligned} Var(\mathbf{K}'\boldsymbol{\beta}^o) &= -\mathbf{K}'\mathbf{C}_{22}\mathbf{K}. & (59)\\ Var(\hat{\mathbf{u}}) &= \mathbf{GZ}'\mathbf{C}_{11}\mathbf{VC}_{11}\mathbf{ZG}. & (60)\\ Cov(\mathbf{K}'\boldsymbol{\beta}^o, \hat{\mathbf{u}}') &= \mathbf{K}'\mathbf{C}_{12}'\mathbf{VC}_{11}\mathbf{ZG} = \mathbf{0}. & (61)\\ Cov(\mathbf{K}'\boldsymbol{\beta}^o, \mathbf{u}') &= \mathbf{K}'\mathbf{C}_{12}'\mathbf{ZG} & (62)\\ Cov(\mathbf{K}'\boldsymbol{\beta}^o, \hat{\mathbf{u}}' - \mathbf{u}') &= -\mathbf{K}'\mathbf{C}_{12}'\mathbf{ZG}. & (63)\\ Var(\hat{\mathbf{u}} - \mathbf{u}) &= \mathbf{G} - Var(\hat{\mathbf{u}}). & (64) \end{aligned}$$

The matrix of (57) will often be too large to invert for purposes of solving $\mathbf{s}$ and $\boldsymbol{\beta}^o$. With mixed model equations that are too large we can solve by Gauss-Seidel iteration. Because this method requires diagonals that are non-zero, we cannot solve (57) by this method. But if we are interested in $\hat{\mathbf{u}}$, but not in $\boldsymbol{\beta}^o$, an iterative method can be used.

Subsection 4.2 presented a method for BLUP that is

$$\hat{\mathbf{u}} = \mathbf{C}_*'\mathbf{V}_*^-\mathbf{y}_*.$$

Now solve iteratively

$$\mathbf{V}_*\mathbf{s} = \mathbf{y}_*, \tag{65}$$

then

$$\hat{\mathbf{u}} = \mathbf{C}_*'\mathbf{s}. \tag{66}$$

Remember that $\mathbf{V}_*$ has rank $= n - r$. Nevertheless convergence will occur, but not to a unique solution. $\mathbf{V}_*$ (and $\mathbf{y}_*$) could be reduced to dimension, $n - r$, so that the reduced $\mathbf{V}_*$ would be non-singular.

Suppose that

$$\mathbf{X}' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 2 & 4 \end{pmatrix},$$

$$\mathbf{C}' = \begin{pmatrix} 1 & 1 & 2 & 0 & 3 \\ 2 & 0 & 1 & 1 & 2 \end{pmatrix},$$

$$\mathbf{V} = \begin{pmatrix} 9 & 3 & 2 & 1 & 1 \\ & 8 & 1 & 2 & 2 \\ & & 9 & 2 & 1 \\ & & & 7 & 2 \\ & & & & 8 \end{pmatrix},$$

$$\mathbf{y}' = [6 \;\; 3 \;\; 5 \;\; 2 \;\; 8].$$

First let us compute $\boldsymbol{\beta}^o$ by GLS and $\hat{\mathbf{u}}$ by $\mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)$.

The GLS equations are

$$\begin{pmatrix} .335816 & .828030 \\ .828030 & 2.821936 \end{pmatrix} \boldsymbol{\beta}^o = \begin{pmatrix} 1.622884 \\ 4.987475 \end{pmatrix}.$$

$$(\boldsymbol{\beta}^o)' = [1.717054 \;\; 1.263566].$$

From this

$$\hat{\mathbf{u}}' = [.817829 \;\; 1.027132].$$

By the method of (57) we have equations

$$\begin{pmatrix} 9 & 3 & 2 & 1 & 1 & 1 & 1 \\ & 8 & 1 & 2 & 2 & 1 & 2 \\ & & 9 & 2 & 1 & 1 & 3 \\ & & & 7 & 2 & 1 & 2 \\ & & & & 8 & 1 & 4 \\ & & & & & 0 & 0 \\ & & & & & & 0 \end{pmatrix} \begin{pmatrix} \mathbf{s} \\ \boldsymbol{\beta}^o \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 5 \\ 2 \\ 8 \\ 0 \\ 0 \end{pmatrix}.$$

The solution is $(\boldsymbol{\beta}^o)'$ = same as for GLS,

$$\mathbf{s}' = (.461240 - .296996 - .076550 - .356589 .268895).$$

Then $\hat{\mathbf{u}} = \mathbf{C}'\mathbf{s}$ = same as before. Next let us compute $\hat{\mathbf{u}}$ from different $\mathbf{y}_*$. First let $\boldsymbol{\beta}_*$ be the solution to OLS using the first two elements of $\mathbf{y}$. This gives

$$\boldsymbol{\beta}_* = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \end{pmatrix} \mathbf{y},$$

and

$$\mathbf{y}_* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 2 & -3 & 0 & 0 & 1 \end{pmatrix} \mathbf{y} = \mathbf{T}'\mathbf{y},$$

or

$$\mathbf{y}_*' = [0 \;\; 0 \;\; 5 \;\; -1 \;\; 11].$$

Using the last 3 elements of $\mathbf{y}_*$ gives

$$\mathbf{V}_*' = \begin{pmatrix} 38 & 11 & 44 \\ & 11 & 14 \\ & & 72 \end{pmatrix}, \quad \mathbf{C}_*' = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 6 \end{pmatrix}.$$

Then

$$\hat{\mathbf{u}} = \mathbf{C}_*'\mathbf{V}_*^{-1}\mathbf{y}_* = \text{same as before.}$$

Another possibility is to compute $\boldsymbol{\beta}_*$ by OLS using elements 1, 3 of $\mathbf{y}$. This gives

$$\boldsymbol{\beta}_* = \begin{pmatrix} 1.5 & 0 & -.5 & 0 & 0 \\ -.5 & 0 & .5 & 0 & 0 \end{pmatrix} \mathbf{y},$$

and

$$\mathbf{y}_*' = [0 \;\; -2.5 \;\; 0 \;\; -3.5 \;\; 3.5].$$

Dropping the first and third elements of $\mathbf{y}_*$,

$$\mathbf{V}_* = \begin{pmatrix} 9.5 & 4.0 & 6.5 \\ & 9.5 & 4.0 \\ & & 25.5 \end{pmatrix}, \quad \mathbf{C}_*' = \begin{pmatrix} -.5 & -1.5 & .5 \\ -1.5 & -.5 & 1.5 \end{pmatrix}.$$

This gives the same value for $\hat{\mathbf{u}}$.

Finally we illustrate $\boldsymbol{\beta}_*$ by GLS.

$$\boldsymbol{\beta}_* = \begin{pmatrix} .780362 & .254522 & -.142119 & .645995 & .538760 \\ -.242894 & -.036176 & .136951 & -.167959 & .310078 \end{pmatrix} \mathbf{y}.$$

$$\mathbf{y}_*' = \begin{pmatrix} 3.019380, & -1.244186, & -.507752, & -2.244186, & 1.228682 \end{pmatrix}.$$

$$\mathbf{V}_* = \begin{pmatrix} 3.268734 & -.852713 & .025840 & -2.85713 & .904393 \\ & 4.744186 & -1.658915 & -1.255814 & -.062016 \\ & & 5.656331 & -.658915 & -3.028424 \\ & & & 3.744186 & -.062016 \\ & & & & 2.005168 \end{pmatrix}.$$

$$\mathbf{C}_*' = \begin{pmatrix} .940568 & .015504 & .090439 & -.984496 & .165375 \\ .909561 & -1.193798 & -.297158 & -.193798 & .599483 \end{pmatrix}.$$

Then $\hat{\mathbf{u}} = \mathbf{C}_*'\mathbf{V}_*^-\mathbf{y}_*$. $\mathbf{V}_*$ has rank = 3, and one g-inverse is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ & .271363 & .092077 & .107220 & 0 \\ & & .211736 & .068145 & 0 \\ & & & .315035 & 0 \\ & & & & 0 \end{pmatrix}.$$

This gives $\hat{\mathbf{u}}$ the same as before.

Another g-inverse is

$$\begin{pmatrix} 1.372401 & 0 & 0 & 1.035917 & -.586957 \\ & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 \\ & & & 1.049149 & -.434783 \\ & & & & .75000 \end{pmatrix}.$$

This gives the same $\hat{\mathbf{u}}$ as before.

It can be seen that when $\boldsymbol{\beta}_* = \boldsymbol{\beta}^o$, a GLS solution, $\mathbf{C}'\mathbf{V}^{-1}\mathbf{y}_* = \mathbf{C}_*'\mathbf{V}_*^-\mathbf{y}_*$. Thus if $\mathbf{V}$ can be inverted to obtain $\boldsymbol{\beta}^o$, this is the easier method. Of course this section is really concerned with the situation in which $\mathbf{V}^{-1}$ is too difficult to compute, and the mixed model equations are also intractable.

16 Prediction When R Is Singular

If $\mathbf{R}$ is singular, the usual mixed model equations, which require $\mathbf{R}^{-1}$, cannot be used. Harville (1976) does describe a method using a particular g-inverse of $\mathbf{R}$ that can be used. Finding this g-inverse is not trivial. Consequently, we shall describe methods different from his that lead to the same results. Different situations exist depending upon whether $\mathbf{X}$ and/or $\mathbf{Z}$ are linearly independent of $\mathbf{R}$.

16.1 X and Z linearly dependent on R

If $\mathbf{R}$ has rank $t < n$, we can write $\mathbf{R}$ with possible re-ordering of rows and columns as

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & \mathbf{R}_1\mathbf{L} \\ \mathbf{L}'\mathbf{R}_1 & \mathbf{L}'\mathbf{R}_1\mathbf{L} \end{pmatrix},$$

where $\mathbf{R}_1$ is $t \times t$, and $\mathbf{L}$ is $t \times (n-t)$ with rank $(n-t)$. Then if $\mathbf{X}$, $\mathbf{Z}$ are linearly dependent upon $\mathbf{R}$,

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{L}'\mathbf{X}_1 \end{pmatrix}, \ \mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{L}'\mathbf{Z}_1 \end{pmatrix}.$$

Then it can be seen that $\mathbf{V}$ is singular, and $\mathbf{X}$ is linearly dependent upon $\mathbf{V}$. One could find $\boldsymbol{\beta}^o$ and $\hat{\mathbf{u}}$ by solving these equations

$$\begin{pmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{s} \\ \boldsymbol{\beta}^o \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}, \tag{67}$$

and $\hat{\mathbf{u}} = \mathbf{GZ}'\mathbf{s}$. See section 14. It should be noted that (67) is not a consistent set of equations unless

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{L}'\mathbf{y}_1 \end{pmatrix}.$$

If $\mathbf{X}$ has full column rank, the solution to $\boldsymbol{\beta}^o$ is unique. If $\mathbf{X}$ is not full rank, $\mathbf{K}'\boldsymbol{\beta}^o$ is unique, given $\mathbf{K}'\boldsymbol{\beta}$ is estimable. There is not a unique solution to $\mathbf{s}$ but $\hat{\mathbf{u}} = \mathbf{GZ}'\mathbf{s}$ is unique.

Let us illustrate with

$$\mathbf{X}' = (1 \ \ 2 \ \ -3), \ \mathbf{Z}' = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 1 & -3 \end{pmatrix}, \ \mathbf{y}' = (5 \ \ 3 \ \ -8),$$

$$\mathbf{R} = \begin{pmatrix} 3 & -1 & -2 \\ & 4 & -3 \\ & & 5 \end{pmatrix}, \quad \mathbf{G} = \mathbf{I}.$$

Then

$$\mathbf{V} = \mathbf{R} + \mathbf{ZGZ} = \begin{pmatrix} 8 & 3 & -11 \\ & 9 & -12 \\ & & 23 \end{pmatrix},$$

which is singular. Then we find some solution to

$$\begin{pmatrix} 8 & 3 & -11 & 1 \\ 3 & 9 & -12 & 2 \\ -11 & -12 & 23 & -3 \\ 1 & 2 & -3 & 0 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ \beta^o \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ -8 \\ 0 \end{pmatrix}.$$

Three different solution vectors are

$$\begin{array}{rrrr} (14 & -7 & 0 & 54)/29, \\ (21 & 0 & 7 & 54)/29, \\ (\ 0 & -21 & -14 & 54)/29. \end{array}$$

Each of these gives $\hat{\mathbf{u}}' = (0\ 21)/29$ and $\boldsymbol{\beta}^o = 54/29$.

We can also obtain a unique solution to $\mathbf{K}'\boldsymbol{\beta}^o$ and $\hat{\mathbf{u}}$ by setting up mixed model equations using $\mathbf{y}_1$ only or any other linearly independent subset of $\mathbf{y}$. In our example let us use the first 2 elements of $\mathbf{y}$. The mixed model equations are

$$\left(\begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right)$$

$$\begin{pmatrix} \beta^o \\ \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix}.$$

These are

$$11^{-1} \begin{pmatrix} 20 & 20 & 19 \\ 20 & 31 & 19 \\ 19 & 19 & 34 \end{pmatrix} \begin{pmatrix} \beta^o \\ \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} = \begin{pmatrix} 51 \\ 51 \\ 60 \end{pmatrix} /11.$$

The solution is $(54,\ 0,\ 21)/29$ as before.

If we use y_1, y_3 we get the same equations as above, and also the same if we use y_2, y_3

16.2 X linearly independent of V, and Z linearly dependent on R

In this case $\mathbf{V}$ is singular but with $\mathbf{X}$ independent of $\mathbf{V}$ equations (67) have a unique solution if $\mathbf{X}$ has full column rank. Otherwise $\mathbf{K}'\boldsymbol{\beta}^o$ is unique provided $\mathbf{K}'\boldsymbol{\beta}$ is estimable. In contrast to section 15.1, $\mathbf{y}$ need not be linearly dependent upon $\mathbf{V}$ and $\mathbf{R}$. Let us use the example of section 14.1 except now $\mathbf{X}' = (1\ 2\ 3)$, and $\mathbf{y}' = (5\ 3\ 4)$. Then the unique solution is $(\mathbf{s}\ \ \boldsymbol{\beta}^o)' = (1104,\ -588,\ 24,\ 4536)/2268$.

16.3 Z linearly independent of R

In this case $\mathbf{V}$ is non-singular, and $\mathbf{X}$ is usually linearly independent of $\mathbf{V}$ even though it may be linearly dependent on $\mathbf{R}$. Consequently $\mathbf{s}$ and $\mathbf{K}'\boldsymbol{\beta}^o$ are unique as in section 15.2.

17 Another Example of Prediction Error Variances

We demonstrate variances of prediction errors and predictors by the following example.

Treatment	n_{ij} Animals	
	1	2
1	2	1
2	1	3

Let

$$\mathbf{R} = 5\mathbf{I}, \quad \mathbf{G} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}.$$

The mixed model coefficient matrix is

$$\begin{pmatrix} 1.4 & .6 & .8 & .6 & .8 \\ & .6 & 0 & .4 & .2 \\ & & 8 & .2 & .6 \\ & & & 1.2 & -.2 \\ & & & & 1.2 \end{pmatrix}, \tag{68}$$

and a g-inverse of this matrix is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ & 3.33333 & 1.66667 & -1.66667 & -1.66667 \\ & & 3.19820 & -1.44144 & -2.1172 \\ & & & 1.84685 & 1.30631 \\ & & & & 2.38739 \end{pmatrix}. \tag{69}$$

Let $\mathbf{K}' = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. Then

$$Var \begin{pmatrix} \mathbf{K}'\boldsymbol{\beta}^o \\ \hat{\mathbf{u}} - \mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{K}' \\ \mathbf{I}_2 \end{pmatrix} \text{ [Matrix (69)] } (\mathbf{K} \;\; \mathbf{I}_2)$$

$$= \begin{pmatrix} 3.33333 & 1.66667 & -1.66667 & -1.66667 \\ & 3.19820 & -1.44144 & -2.1172 \\ & & 1.84685 & 1.30631 \\ & & & 2.38739 \end{pmatrix}. \tag{70}$$

$$Var \begin{pmatrix} \mathbf{K}'\boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} 3.33333 & 1.66667 & 0 & 0 \\ & 3.198198 & 0 & 0 \\ & & .15315 & -.30631 \\ & & & .61261 \end{pmatrix}. \tag{71}$$

The upper 2 x 2 is the same as in (70).

$$Cov(\mathbf{K}'\boldsymbol{\beta}^o,\ \hat{\mathbf{u}}') = \mathbf{0}.$$

$$Var(\hat{\mathbf{u}}) = \mathbf{G}\ -\ Var(\hat{\mathbf{u}}\ -\ \mathbf{u}).$$

Let us derive these results from first principles.

$$\begin{pmatrix} \mathbf{K}'\boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} .33333 & .33333 & .33333 & 0 & 0 & 0 & 0 \\ .04504 & .04504 & -.09009 & .35135 & .21622 & .21622 & .21622 \\ .03604 & .03604 & -.07207 & .08108 & -.02703 & -.02703 & -.02703 \\ -.07207 & -.07207 & .14414 & -.16216 & .05405 & .05405 & .05405 \end{pmatrix} \mathbf{y} \tag{72}$$

computed by

$$\begin{pmatrix} \mathbf{K}' \\ \mathbf{I}_2 \end{pmatrix} \text{ [matrix (71)] } \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1} \\ \mathbf{Z}'\mathbf{R}^{-1} \end{pmatrix}.$$

Contribution of $\mathbf{R}$ to $Var\begin{pmatrix} \mathbf{K}'\boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix}$

$$\begin{aligned} &= \text{[matrix (72)] } \mathbf{R} \text{ [matrix (72)]}' \\ &= \begin{pmatrix} 1.6667 & 0 & 0 & 0 \\ & 1.37935 & .10348 & -.20696 \\ & & .08278 & -.16557 \\ & & & .33114 \end{pmatrix}. \end{aligned} \tag{73}$$

For $\mathbf{u}$ in

$$\begin{aligned} \begin{pmatrix} \mathbf{K}'\boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} &= \begin{pmatrix} \mathbf{K}' \\ \mathbf{I}_2 \end{pmatrix} \text{ [matrix (72)] } \mathbf{Z} \\ &= \begin{pmatrix} .66667 & .33333 \\ .44144 & .55856 \\ .15315 & -.15315 \\ -.30631 & .30631 \end{pmatrix}. \end{aligned} \tag{74}$$

Contribution of $\mathbf{G}$ to $Var\begin{pmatrix} \mathbf{K}'\boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix}$

$$\begin{aligned} &= \text{[matrix (74)] } \mathbf{G} \text{ [matrix (74)]}' \\ &= \begin{pmatrix} 1.6667 & 1.66662 & 0 & 0 \\ & 1.81885 & -.10348 & .20696 \\ & & .07037 & -.14074 \\ & & & .28143 \end{pmatrix}. \end{aligned} \tag{75}$$

Then the sum of matrix (73) and matrix (75) = matrix (71). For variance of prediction errors we need

$$\text{Matrix (74)} \; - \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} .66667 & .33333 \\ .44144 & .55856 \\ -.84685 & -.15315 \\ -.30631 & -.69369 \end{pmatrix}. \quad (76)$$

Then contribution of $\mathbf{G}$ to prediction error variance is

$$\begin{aligned} & [\text{matrix (76)}] \; \mathbf{G} \; [\text{Matrix (76)}]', \\ = & \begin{pmatrix} 1.66667 & 1.66667 & -1.66667 & -1.66667 \\ & 1.81885 & -1.54492 & -1.91015 \\ & & 1.76406 & 1.47188 \\ & & & 2.05624 \end{pmatrix}. \end{aligned} \quad (77)$$

Then prediction error variance is matrix (73) + matrix (77) = matrix (70).

18 Prediction When u And e Are Correlated

In most applications of BLUE and BLUP it is assumed that $Cov(\mathbf{u}, \mathbf{e}') = \mathbf{0}$. If this is not the case, the mixed model equations can be modified to account for such covariances. See Schaeffer and Henderson (1983).

Let

$$Var \begin{pmatrix} \mathbf{e} \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{R} & \mathbf{S} \\ \mathbf{S}' & \mathbf{G} \end{pmatrix}. \quad (78)$$

Then

$$Var(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R} + \mathbf{ZS}' + \mathbf{SZ}'. \quad (79)$$

Let an equivalent model be

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Tu} + \boldsymbol{\varepsilon}, \quad (80)$$

where $\mathbf{T} = \mathbf{Z} + \mathbf{SG}^{-1}$,

$$Var \begin{pmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}, \quad (81)$$

and $\mathbf{B} = \mathbf{R} - \mathbf{SG}^{-1}\mathbf{S}'$. Then

$$\begin{aligned} Var(\mathbf{y}) &= Var(\mathbf{Tu} + \boldsymbol{\varepsilon}) \\ &= \mathbf{ZGZ}' + \mathbf{ZS}' + \mathbf{SZ}' + \mathbf{SG}^{-1}\mathbf{S}' + \mathbf{R} - \mathbf{SG}^{-1}\mathbf{S}' \\ &= \mathbf{ZGZ}' + \mathbf{R} + \mathbf{ZS}' + \mathbf{SZ}' \end{aligned}$$

as in the original model, thus proving equivalence. Now the mixed model equations are

$$\begin{pmatrix} \mathbf{X}'\mathbf{B}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{B}^{-1}\mathbf{T} \\ \mathbf{T}'\mathbf{B}^{-1}\mathbf{X} & \mathbf{T}'\mathbf{B}^{-1}\mathbf{T} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{B}^{-1}\mathbf{y} \\ \mathbf{T}'\mathbf{B}^{-1}\mathbf{y} \end{pmatrix}, \tag{82}$$

A g-inverse of this matrix yields the required variances and covariances for estimable functions of $\boldsymbol{\beta}^o$, $\hat{\mathbf{u}}$, and $\hat{\mathbf{u}} - \mathbf{u}$.

$\mathbf{B}$ can be inverted by a method analogous to

$$\mathbf{V}^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}$$

where $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$,

$$\mathbf{B}^{-1} = \mathbf{R}^{-1} + \mathbf{R}^{-1}\mathbf{S}(\mathbf{G} - \mathbf{S}'\mathbf{R}^{-1}\mathbf{S})^{-1}\mathbf{S}'\mathbf{R}^{-1}. \tag{83}$$

In fact, it is unnecessary to compute $\mathbf{B}^{-1}$ if we instead solve (84).

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{T} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{S} \\ \mathbf{T}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{T}'\mathbf{R}^{-1}\mathbf{T} + \mathbf{G}^{-1} & \mathbf{T}'\mathbf{R}^{-1}\mathbf{S} \\ \mathbf{S}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{S}'\mathbf{R}^{-1}\mathbf{T} & \mathbf{S}'\mathbf{R}^{-1}\mathbf{S} - \mathbf{G} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{T}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{S}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{84}$$

This may not be a good set of equations to solve iteratively since $(\mathbf{S}'\mathbf{R}^{-1}\mathbf{S} - \mathbf{G})$ is negative definite. Consequently Gauss- Seidel iteration is not guaranteed to converge, Van Norton (1959).

We illustrate the method of this section by an additive genetic model.

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \\ 1 & 4 \end{pmatrix}, \ \mathbf{Z} = \mathbf{I}_4, \mathbf{G} = \begin{pmatrix} 1. & .5 & .25 & .25 \\ & 1. & .25 & .25 \\ & & 1. & .5 \\ & & & 1. \end{pmatrix}, \ \mathbf{R} = 4\mathbf{I}_4,$$

$$\mathbf{S} = \mathbf{S}' = .9\mathbf{I}_4, \ \mathbf{y}' = (5, 6, 7, 9).$$

From these parameters

$$\mathbf{B} = \begin{pmatrix} 2.88625 & .50625 & .10125 & .10125 \\ & 2.88625 & .10125 & .10125 \\ & & 2.88625 & .50625 \\ & & & 2.88625 \end{pmatrix},$$

and

$$\mathbf{T} = \mathbf{T}' = \begin{pmatrix} 2.2375 & -.5625 & -.1125 & -.1125 \\ & 2.2375 & -.1125 & -.1125 \\ & & 2.2375 & .5625 \\ & & & 2.2375 \end{pmatrix}.$$

Then the mixed model equations of (84) are

$$\begin{pmatrix} 1.112656 & 2.225313 & .403338 & .403338 & .403338 & .403338 \\ & 6.864946 & -.079365 & 1.097106 & -.660224 & 2.869187 \\ & & 3.451184 & -1.842933 & -.261705 & -.261705 \\ & & & 3.451184 & -.261705 & -.261705 \\ & & & & 3.451184 & -1.842933 \\ & & & & & 3.451184 \end{pmatrix}$$

$$\begin{pmatrix} \beta_1^o \\ \beta_2^o \\ \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{u}_4 \end{pmatrix} = \begin{pmatrix} 7.510431 \\ 17.275150 \\ 1.389782 \\ 2.566252 \\ 2.290575 \\ 4.643516 \end{pmatrix}.$$

The solution is (4.78722, .98139, −.21423, −.21009, .31707, .10725).

We could solve this problem by the basic method

$$\boldsymbol{\beta}^o = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y},$$

and

$$\hat{\mathbf{u}} = \text{Cov }(\mathbf{u}, \mathbf{y}')\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o).$$

We illustrate that these give the same answers as the mixed model method.

$$Var(\mathbf{y}) = \mathbf{V} = \begin{pmatrix} 6.8 & .5 & .25 & .25 \\ & 6.8 & .25 & .25 \\ & & 6.8 & .5 \\ & & & 6.8 \end{pmatrix}.$$

Then the GLS equations are

$$\begin{pmatrix} .512821 & 1.025641 \\ 1.025641 & 2.991992 \end{pmatrix} \hat{\boldsymbol{\beta}} = \begin{pmatrix} 3.461538 \\ 7.846280 \end{pmatrix},$$

and

$$\hat{\boldsymbol{\beta}} = (4.78722, .98139)'$$

as before.

$$Cov(\mathbf{u}, \mathbf{y}') = \begin{pmatrix} 1.90 & .50 & .25 & .25 \\ & 1.90 & .25 & .25 \\ & & 1.90 & .50 \\ & & & 1.90 \end{pmatrix} = \mathbf{GZ}' + \mathbf{S}'.$$

$$(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = (-.768610, -.750000, 1.231390, .287221).$$

$$\hat{\mathbf{u}} = (-.21423, -.21009, .31707, .10725)' = (\mathbf{GZ}' + \mathbf{S}')\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)$$

as before.

19 Direct Solution To β And u +T β

In some problems we wish to predict $\mathbf{w} = \mathbf{u} + \mathbf{T}\boldsymbol{\beta}$. The mixed model equations can be modified to do this. Write the mixed model equations as (85). This can be done since $E(\mathbf{w} - \mathbf{T}\boldsymbol{\beta}) = \mathbf{0}$.

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{w} - \mathbf{T}\boldsymbol{\beta}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{85}$$

Re-write (85) as

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} - \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{T} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} - \mathbf{M} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{w}} \end{pmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} \tag{86}$$

where $\mathbf{M} = (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})\mathbf{T}$. To obtain symmetry premultiply the second equation by $\mathbf{T}'$ and subtract this product from the first equation. This gives

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} - \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{T} - \mathbf{T}'\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} + \mathbf{T}'\mathbf{M} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} - \mathbf{M}' \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} - \mathbf{M} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix}$$
$$\begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{w}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} - \mathbf{T}'\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{87}$$

Let a g-inverse of the matrix of (87) be $\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix}$. Then

$$\begin{aligned} Var(\mathbf{K}'\boldsymbol{\beta}^o) &= \mathbf{K}'\mathbf{C}_{11}\mathbf{K}. \\ Var(\hat{\mathbf{w}} - \mathbf{w}) &= \mathbf{C}_{22}. \end{aligned}$$

Henderson's mixed model equations for a selection model, equation (31), in Biometrics (1975a) can be derived from (86) by making the following substitutions, $\begin{pmatrix} \mathbf{X} \\ \mathbf{B} \end{pmatrix}$ for $\mathbf{X}$, $(\mathbf{0}\ \mathbf{B})$ for $\mathbf{T}$, and noting that $\mathbf{B} = \mathbf{Z}\mathbf{B}_u + \mathbf{B}_e$.

We illustrate (87) with the following example.

$$\mathbf{X} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \\ 3 & 4 \end{pmatrix}, \ \mathbf{Z} = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \\ 4 & 1 & 3 \end{pmatrix}, \ \mathbf{R} = \begin{pmatrix} 5 & 1 & 1 & 2 \\ & 6 & 2 & 1 \\ & & 7 & 1 \\ & & & 8 \end{pmatrix},$$

$$\mathbf{G} = \begin{pmatrix} 3 & 1 & 1 \\ & 4 & 2 \\ & & 5 \end{pmatrix}, \mathbf{T} = \begin{pmatrix} 3 & 1 \\ 2 & 3 \\ 2 & 4 \end{pmatrix}, \ \mathbf{y}' = (5,\ 2,\ 3,\ 6).$$

The regular mixed model equations are

$$\begin{pmatrix} 1.576535 & 1.651127 & 1.913753 & 1.188811 & 1.584305 \\ & 2.250194 & 2.088578 & .860140 & 1.859363 \\ & & 2.763701 & 1.154009 & 1.822952 \\ & & & 2.024882 & 1.142462 \\ & & & & 2.077104 \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} 2.651904 \\ 3.871018 \\ 3.184149 \\ 1.867133 \\ 3.383061 \end{pmatrix} \tag{88}$$

The solution is

$$(-2.114786, \ 2.422179, \ .086576, \ .757782, \ .580739).$$

The equations for solution to $\boldsymbol{\beta}$ and to $\mathbf{w} = \mathbf{u} + \mathbf{T}\boldsymbol{\beta}$ are

$$\begin{pmatrix} 65.146040 & 69.396108 & -12.331273 & -8.607904 & -10.323684 \\ & 81.185360 & -11.428959 & -10.938364 & -11.699391 \\ & & 2.763701 & 1.154009 & 1.822952 \\ & & & 2.024882 & 1.142462 \\ & & & & 2.077104 \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{w}} \end{pmatrix} = \begin{pmatrix} -17.400932 \\ -18.446775 \\ 3.184149 \\ 1.867133 \\ 3.383061 \end{pmatrix}. \tag{89}$$

The solution is

$$(-2.115, \ 2.422, \ -3.836, \ 3.795, \ 6.040).$$

This is the same solution to $\boldsymbol{\beta}^o$ as in (88), and $\hat{\mathbf{u}} + \mathbf{T}\boldsymbol{\beta}^o$ of the previous solution gives $\hat{\mathbf{w}}$ of this solution. Further,

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{T} & \mathbf{I} \end{pmatrix} [\text{inverse of (88)}] \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{T} & \mathbf{I} \end{pmatrix}', = [\text{inverse of (89)}]$$

20 Derivation Of MME By Maximizing $f(\mathbf{y}, \mathbf{w})$

This section describes first the method used by Henderson (1950) to derive his mixed model equations. Then a more general result is described. For the regular mixed model

$$E \begin{pmatrix} \mathbf{u} \\ \mathbf{e} \end{pmatrix} = 0, \ Var \begin{pmatrix} \mathbf{u} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix}.$$

The density function is

$$f(\mathbf{y}, \mathbf{u}) = g(\mathbf{y} \mid \mathbf{u})\ h(\mathbf{u}),$$

and under normality the log of this is

$$k[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u})'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) + \mathbf{u}'\mathbf{G}^{-1}\mathbf{u}],$$

where k is a constant. Differentiating with respect to $\boldsymbol{\beta}, \mathbf{u}$ and equating to $\mathbf{0}$ we obtain the regular mixed model equations.

Now consider a more general mixed linear model in which

$$E\begin{pmatrix} \mathbf{y} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{X}\boldsymbol{\beta} \\ \mathbf{T}\boldsymbol{\beta} \end{pmatrix}$$

with $\mathbf{T}\boldsymbol{\beta}$ estimable, and

$$Var\begin{pmatrix} \mathbf{y} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{V} & \mathbf{C} \\ \mathbf{C}' & \mathbf{G} \end{pmatrix}$$

with

$$\begin{pmatrix} \mathbf{V} & \mathbf{C} \\ \mathbf{C}' & \mathbf{G} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix}.$$

Log of $f(\mathbf{y}, \mathbf{w})$ is

$$k[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{C}_{11}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{C}_{12}(\mathbf{w} - \mathbf{T}\boldsymbol{\beta})$$
$$+(\mathbf{w} - \mathbf{T}\boldsymbol{\beta})'\mathbf{C}_{12}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + (\mathbf{w} - \mathbf{T}\boldsymbol{\beta})'\mathbf{C}_{22}(\mathbf{w} - \mathbf{T}\boldsymbol{\beta}).$$

Differentiating with respect to $\boldsymbol{\beta}$ and to $\mathbf{w}$ and equating to $\mathbf{0}$, we obtain

$$\begin{pmatrix} \mathbf{X}'\mathbf{C}_{11}\mathbf{X} + \mathbf{X}'\mathbf{C}_{12}\mathbf{T} + \mathbf{T}'\mathbf{C}_{12}'\mathbf{X} + \mathbf{T}'\mathbf{C}_{22}\mathbf{T} & -(\mathbf{X}'\mathbf{C}_{12} + \mathbf{T}'\mathbf{C}_{22}) \\ -(\mathbf{X}'\mathbf{C}_{12} + \mathbf{T}'\mathbf{C}_{22})' & \mathbf{C}_{22} \end{pmatrix}$$
$$\begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{w}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{C}_{11}\mathbf{y} + \mathbf{T}'\mathbf{C}_{12}'\mathbf{y} \\ -\mathbf{C}_{12}'\mathbf{y} \end{pmatrix}. \quad (90)$$

Eliminating $\hat{\mathbf{w}}$ we obtain

$$\mathbf{X}'(\mathbf{C}_{11} - \mathbf{C}_{12}'\mathbf{C}_{22}^{-1}\mathbf{C}_{12})\mathbf{X}\boldsymbol{\beta}^o = \mathbf{X}'(\mathbf{C}_{11} - \mathbf{C}_{12}'\mathbf{C}_{22}^{-1}\mathbf{C}_{12})\mathbf{y}. \quad (91)$$

But from partitioned matrix inverse results we know that

$$\mathbf{C}_{11} - \mathbf{C}_{12}'\ \mathbf{C}_{22}^{-1}\mathbf{C}_{12} = \mathbf{V}^{-1}.$$

Therefore (91) are GLS equations and $\mathbf{K}'\boldsymbol{\beta}^o$ is BLUE of $\mathbf{K}'\boldsymbol{\beta}$ if estimable.

Now solve for $\hat{\mathbf{w}}$ from the second equation of (90).

$$\begin{aligned} \hat{\mathbf{w}} &= -\mathbf{C}_{22}^{-1}\mathbf{C}_{12}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) + \mathbf{T}\boldsymbol{\beta}^o. \\ &= \mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) + \mathbf{T}\boldsymbol{\beta}^o. \\ &= \text{BLUP of } \mathbf{w} \text{ because } -\mathbf{C}_{22}^{-1}\mathbf{C}_{12}' = \mathbf{C}'\mathbf{V}^{-1}. \end{aligned}$$

To prove that $-\mathbf{C}_{22}^{-1}\mathbf{C}_{12}' = \mathbf{C}'\mathbf{V}^{-1}$ note that by the definition of an inverse $\mathbf{C}_{12}'\mathbf{V}+\mathbf{C}_{22}\mathbf{C}' = \mathbf{0}$. Pre-multiply this by $\mathbf{C}_{22}^{-1}$ and post-multiply by $\mathbf{V}^{-1}$ to obtain

$$\mathbf{C}_{22}^{-1}\mathbf{C}_{12}' + \mathbf{C}'\mathbf{V}^{-1} = \mathbf{0} \;\text{ or }\; -\mathbf{C}_{22}^{-1}\mathbf{C}_{12}' = \mathbf{C}'\mathbf{V}^{-1}.$$

We illustrate the method with the same example as that of section 18.

$$\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R} = \begin{pmatrix} 46 & 66 & 38 & 74 \\ & 118 & 67 & 117 \\ & & 45 & 66 \\ & & & 149 \end{pmatrix}, \mathbf{C} = \mathbf{ZG} = \begin{pmatrix} 6 & 9 & 13 \\ 11 & 18 & 18 \\ 6 & 11 & 10 \\ 16 & 14 & 21 \end{pmatrix}.$$

Then from the inverse of $\begin{pmatrix} \mathbf{V} & \mathbf{ZG} \\ \mathbf{GZ}' & \mathbf{G} \end{pmatrix}$, we obtain

$$\mathbf{C}_{11} = \begin{pmatrix} .229215 & -.023310 & -.018648 & -.052059 \\ & .188811 & -.048951 & -.011655 \\ & & .160839 & -.009324 \\ & & & .140637 \end{pmatrix},$$

$$\mathbf{C}_{12} = \begin{pmatrix} .044289 & -.069930 & -.236985 \\ -.258741 & -.433566 & -.247086 \\ -.006993 & -.146853 & .002331 \\ -.477855 & -.034965 & -.285159 \end{pmatrix},$$

and

$$\mathbf{C}_{22} = \begin{pmatrix} 2.763701 & 1.154009 & 1.822952 \\ & 2.024882 & 1.142462 \\ & & 2.077104 \end{pmatrix}.$$

Then applying (90) to these results we obtain the same equations as in (89).

The method of this section could have been used to derive the equations of (82) for $Cov(\mathbf{u},\ \mathbf{e}') \neq \mathbf{0}$.

$$f(\mathbf{y},\mathbf{u}) = g(\mathbf{y} \mid \mathbf{u})\ h(\mathbf{u}).$$

$$E(\mathbf{y} \mid \mathbf{u}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Tu},\ Var(\mathbf{y} \mid \mathbf{u}) = \mathbf{B}.$$

See section 17 for definition of $\mathbf{T}$ and $\mathbf{B}$. Then

$$\log\ g(\mathbf{y} \mid \mathbf{u})\ h(\mathbf{u}) = k(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Tu})'\mathbf{B}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Tu}) + \mathbf{u}'\mathbf{G}^{-1}\mathbf{u}.$$

This is maximized by solving (82).

This method also could be used to derive the result of section 18. Again we make use of $f(\mathbf{y}, \mathbf{w}) = g(\mathbf{y} \mid \mathbf{w})\ h(\mathbf{w})$.

$$\begin{aligned} E(\mathbf{y} \mid \mathbf{w}) &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}(\mathbf{w} - \mathbf{T}\boldsymbol{\beta}). \\ Var(\mathbf{y} \mid \mathbf{w}) &= \mathbf{R}. \end{aligned}$$

Then

$$\begin{aligned} \log\ g(\mathbf{y} \mid \mathbf{w})\ h(\mathbf{w}) &= k[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{w} - \mathbf{Z}\mathbf{T}\boldsymbol{\beta})'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{w} - \mathbf{Z}\mathbf{T}\boldsymbol{\beta})] \\ &\quad + (\mathbf{w} - \mathbf{T}\boldsymbol{\beta})'\mathbf{G}^{-1}(\mathbf{w} - \mathbf{T}\boldsymbol{\beta}). \end{aligned}$$

This is maximized by solving equations (87).

Chapter 6
G and R Known to Proportionality

C. R. Henderson

1984 - Guelph

In the preceding chapters it has been assumed that $Var(\mathbf{u}) = \mathbf{G}$ and $Var(\mathbf{e}) = \mathbf{R}$ are known. This is, of course, an unrealistic assumption, but was made in order to present estimation, prediction, and hypothesis testing methods that are exact and which may suggest approximations for the situation with unknown $\mathbf{G}$ and $\mathbf{R}$. One case does exist, however, in which BLUE and BLUP exist, and exact tests can be made even when these variances are unknown. This case is $\mathbf{G}$ and $\mathbf{R}$ known to proportionality.

Suppose that we know $\mathbf{G}$ and $\mathbf{R}$ to proportionality, that is

$$\begin{aligned} \mathbf{G} &= \mathbf{G}_* \ \sigma_e^2, \\ \mathbf{R} &= \mathbf{R}_* \ \sigma_e^2. \end{aligned} \tag{1}$$

$\mathbf{G}_*$ and $\mathbf{R}_*$ are known, but σ_e^2 is not. For example, suppose that we have a one way mixed model

$$\begin{aligned} y_{ij} &= \mathbf{x}_{ij}^{'}\boldsymbol{\beta} + a_i + e_{ij}. \\ Var(a_1 \ \ a_2 \ ...)' &= \mathbf{I}\sigma_a^2. \\ Var(e_{11} \ \ e_{12} \ ...)' &= \mathbf{I}\sigma_e^2. \end{aligned}$$

Suppose we know that $\sigma_a^2/\sigma_e^2 = \alpha$. Then

$$\begin{aligned} \mathbf{G} &= \mathbf{I}\sigma_a^2 = \mathbf{I}\alpha \ \sigma_e^2. \\ \mathbf{R} &= \mathbf{I}\sigma_e^2. \end{aligned}$$

Then by the notation of (1)

$$\mathbf{G}_* = \mathbf{I}\alpha, \ \mathbf{R}_* = \mathbf{I}.$$

1 BLUE and BLUP

Let us write the GLS equations with the notation of (1).

$$\begin{aligned} \mathbf{V} &= \mathbf{ZGZ}' + \mathbf{R} \\ &= (\mathbf{ZG}_*\mathbf{Z}' + \mathbf{R}_*)\sigma_e^2 \\ &= \mathbf{V}_*\sigma_e^2. \end{aligned}$$

Then $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}^o \; = \; \mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ can be written as

$$\sigma_e^{-2}\mathbf{X}'\mathbf{V}_*^{-1}\mathbf{X}\boldsymbol{\beta}^o \; = \; \mathbf{X}'\mathbf{V}_*^{-1}\mathbf{y}\sigma_e^{-2}. \tag{2}$$

Multiplying both sides by σ_e^2 we obtain a set of equations that can be written as,

$$\mathbf{X}'\mathbf{V}_*^{-1}\mathbf{X}\boldsymbol{\beta}^o \; = \; \mathbf{X}'\mathbf{V}_*^{-1}\mathbf{y}. \tag{3}$$

Then BLUE of $\mathbf{K}'\boldsymbol{\beta}$ is $\mathbf{K}'\boldsymbol{\beta}^o$, where $\boldsymbol{\beta}^o$ is any solution to (3).

Similarly the mixed model equations with each side multiplied by σ_e^2 are

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}_*^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}_*^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}_*^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}_*^{-1}\mathbf{Z} + \mathbf{G}_*^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} \; = \; \begin{pmatrix} \mathbf{X}'\mathbf{R}_*^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}_*^{-1}\mathbf{y} \end{pmatrix}. \tag{4}$$

$\hat{\mathbf{u}}$ is BLUP of $\mathbf{u}$ when $\mathbf{G}_*$ and $\mathbf{R}_*$ are known.

To find the sampling variance of $\mathbf{K}'\boldsymbol{\beta}^o$ we need a g-inverse of the matrix of (2). This is

$$(\mathbf{X}'\mathbf{V}_*^{-1}\mathbf{X})^- \; \sigma_e^2.$$

Consequently,

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) \; = \; \mathbf{K}'(\mathbf{X}'\mathbf{V}_*^{-1}\mathbf{X})^-\mathbf{K}\sigma_e^2. \tag{5}$$

Also

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) \; = \; \mathbf{K}'\mathbf{C}_{11}\mathbf{K}\sigma_e^2,$$

where $\mathbf{C}_{11}$ is the upper p^2 submatrix of a g-inverse of the matrix of (4). Similarly all of the results of (34) to (41) in Chapter 5 are correct if we multiply them by σ_e^2.

Of course σ_e^2 is unknown, so we can only estimate the variance by substituting some estimate of σ_e^2, say $\hat{\sigma}_e^2$, in (5). There are several methods for estimating σ_e^2, but the most frequently used one is the minimum variance, translation invariant, quadratic, unbiased estimator computed by

$$[\mathbf{y}'\mathbf{V}_*^{-1}\mathbf{y} - (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{V}_*^{-1}\mathbf{y}]/[n - rank(\mathbf{X})] \tag{6}$$

or by

$$[\mathbf{y}'\mathbf{R}_*^{-1}\mathbf{y} - (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{R}_*^{-1}\mathbf{y} - \hat{\mathbf{u}}'\mathbf{Z}'\mathbf{R}_*^{-1}\mathbf{y}]/[n - rank(\mathbf{X})]. \tag{7}$$

A more detailed account of estimation of variances is presented in Chapters 10, 11, and 12.

Next looking at BLUP of $\mathbf{u}$ under model (1), it is readily seen that $\hat{\mathbf{u}}$ of (4) is BLUP. Similarly variances and covariances involving $\hat{\mathbf{u}}$ and $\hat{\mathbf{u}} - \mathbf{u}$ are easily derived from the results for known $\mathbf{G}$ and $\mathbf{R}$. Let

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12} & \mathbf{C}_{22} \end{pmatrix}$$

be a g-inverse of the matrix of (4). Then

$$\begin{aligned}
Cov(\mathbf{K}'\boldsymbol{\beta}^o, \hat{\mathbf{u}}' - \mathbf{u}') &= \mathbf{K}'\mathbf{C}_{12}\sigma_e^2, & (8) \\
Var(\hat{\mathbf{u}}) &= (\mathbf{G}_* - \mathbf{C}_{22})\sigma_e^2, & (9) \\
Var(\hat{\mathbf{u}} - \mathbf{u}) &= \mathbf{C}_{22}\sigma_e^2. & (10)
\end{aligned}$$

2 Tests of Hypotheses

In the same way in which $\mathbf{G}$ and $\mathbf{R}$ known to proportionality pose no problems in BLUE and BLUP, exact tests of hypotheses regarding $\boldsymbol{\beta}$ can be performed, assuming as before a multivariate normal distribution. Chapter 4 describes computation of a quadratic, s, that is distributed as χ^2 with $m - a$ degrees of freedom when the null hypothesis is true, and m and a are the number of rows in $\mathbf{H}_0'$ and $\mathbf{H}_a'$ respectively. Now we compute these quadratics exactly as in these methods except that $\mathbf{V}_*$, $\mathbf{G}_*$, $\mathbf{R}_*$ are substituted for $\mathbf{V}$, $\mathbf{G}$, $\mathbf{R}$. Then when the null hypothesis is true, $s/\hat{\sigma}_e^2(m - a)$ is distributed as F with $m - a$, and $n-$rank $(\mathbf{X})$ degrees of freedom, where $\hat{\sigma}_e^2$ is computed by (6) or (7).

3 Power Of The Test Of Null Hypotheses

Two different types of errors can be made in tests of hypotheses. First, the null hypothesis may be rejected when in fact it is true. This is commonly called a Type 1 error. Second, the null hypothesis may be accepted when it is really not true. This is called a Type 2 error, and the power of the test is defined as 1 minus the probability of a Type 2 error. The results that follow regarding power assume that $\mathbf{G}_*$ and $\mathbf{R}_*$ are known.

The power of the test can be computed only if

1. The true value of $\boldsymbol{\beta}$ for which the power is to be determined is specified. Different values of $\boldsymbol{\beta}$ give different powers. Let this value be $\boldsymbol{\beta}_t$. Of course we do not know the true value, but we may be interested in the power of the test, usually for some minimum differences among elements of $\boldsymbol{\beta}$. Logically $\boldsymbol{\beta}_t$ must be true if the null and the alternative hypotheses are true. Accordingly a $\boldsymbol{\beta}_t$ must be chosen that violates neither $\mathbf{H}_0'\boldsymbol{\beta} = \mathbf{c}_0$ nor $\mathbf{H}_a'\boldsymbol{\beta} = \mathbf{c}_a$.

2. The probability of the type 1 error must be specified. This is often called the chosen significance level of the test.

3. The value of $\hat{\sigma}_e^2$ must be specified. Because the power should normally be computed prior to the experiment, this would come from prior research. Define this value as d.

4. **X** and **Z** must be specified.

Then let

$$\begin{aligned} \text{A} &= \text{significance level} \\ F_1 &= m - a \;=\; \text{numerator d.f.} \\ F_2 &= n - \text{rank}(\mathbf{X}) \;=\; \text{denominator d.f.} \end{aligned}$$

Compute $\triangle =$ the quadratic, s, but with $\mathbf{X}\boldsymbol{\beta}_t$ substituted for $\mathbf{y}$ in the computations. Compute

$$P = [\triangle/(m - a + 1)d]^{1/2} \tag{11}$$

and enter Tiku's table (1967) with $A,\ F_1,\ F_2,\ P$ to find the power of the test.

Let us illustrate computation of power by a simple one-way fixed model,

$$\begin{aligned} y_{ij} &= \mu + t_i + e_{ij}, \\ i &= 1, 2, 3. \\ Var(\mathbf{e}) &= \mathbf{I}\sigma_e^2. \end{aligned}$$

Suppose there are 3,2,4 observations respectively on the 3 treatments. We wish to test

$$\mathbf{H}_0'\boldsymbol{\beta} = \mathbf{0},$$

where

$$\mathbf{H}_0' = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

against the unrestricted hypothesis.

Suppose we want the power of the test for $\boldsymbol{\beta}_t' = [10, 2, 1, -3]$ and $\sigma_e^2 = 12$. That is, $d = 12$. Then

$$(\mathbf{X}\boldsymbol{\beta}_t)' = [12,\ 12,\ 12,\ 11,\ 11,\ 7,\ 7,\ 7,\ 7].$$

As we have shown, the reduction under the null hypothesis in this case can be found from the reduced model $E(\mathbf{y}) = \mu_0$. The OLS equations are

$$\begin{pmatrix} 9 & 3 & 2 & 4 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 4 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} \mu^o \\ t^o \end{pmatrix} = \begin{pmatrix} 86 \\ 36 \\ 22 \\ 28 \end{pmatrix}.$$

A solution is (0, 12, 11, 7), and reduction = 870. The restricted equations are

$$9\ \mu^o = 86,$$

and the reduction is 821.78. Then $s = 48.22 = \triangle$. Let us choose $A = .05$ as the significance level

$$\begin{aligned} F_1 &= 2 - 0 = 2. \\ F_2 &= 9 - 3 = 6. \\ P &= \frac{48.22}{3(12)} = 1.157. \end{aligned}$$

Entering Tiku's table we obtain the power of the test.

Chapter 7
Known Functions of Fixed Effects

C. R. Henderson

1984 - Guelph

In previous chapters we have dealt with linear relationships among $\boldsymbol{\beta}$ of the following types.

1. $\mathbf{M}'\boldsymbol{\beta}$ is a set of p-r non-estimable functions of $\boldsymbol{\beta}$, and a solution to GLS or mixed model equations is obtained such that $\mathbf{M}'\boldsymbol{\beta}^o = \mathbf{c}$.

2. $\mathbf{K}'\boldsymbol{\beta}$ is a set of r estimable functions. Then we write a set of equations, the solution to which yields directly BLUE of $\mathbf{K}'\boldsymbol{\beta}$.

3. $\mathbf{H}'\boldsymbol{\beta}$ is a set of estimable functions that are used in hypothesis testing.

In this chapter we shall be concerned with defined linear relationships of the form,

$$\mathbf{T}'\boldsymbol{\beta} = \mathbf{c}.$$

All of these are linearly independent. The consequence of these relationships is that functions of $\boldsymbol{\beta}$ may become estimable that are not estimable under a model with no such definitions concerning $\boldsymbol{\beta}$. In fact, if $\mathbf{T}'\boldsymbol{\beta}$ represents $p - r$ linearly independent non-estimable functions, all linear functions of $\boldsymbol{\beta}$ become estimable.

1 Tests of Estimability

If $\mathbf{T}'\boldsymbol{\beta}$ represents $t < p - r$ non-estimable functions the following rule can be used to determine what functions are estimable.

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{T}' \end{pmatrix} \mathbf{C} = \mathbf{0}, \tag{1}$$

where $\mathbf{C}$ is a $p \times (r - t)$ matrix with rank $r - t$. $\mathbf{C}$ always exists. Then $\mathbf{K}'\boldsymbol{\beta}$ is estimable if and only if

$$\mathbf{K}'\mathbf{C} = \mathbf{0}. \tag{2}$$

To illustrate, suppose that

$$\mathbf{X} = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 2 & 1 & 3 & 5 \\ 1 & 3 & 4 & 0 \\ 3 & 2 & 5 & 7 \\ 1 & 1 & 2 & 2 \\ 2 & 1 & 3 & 5 \end{pmatrix}.$$

with $p = 4$, and $r = 2$ becuase

$$\mathbf{X} \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} = \mathbf{0}$$

Suppose we define $\mathbf{T}' = (1\ 2\ 2\ 1)$. This is a non-estimable function because

$$(1\ 2\ 2\ 1) \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} = (1\ \ 0) \neq \mathbf{0}.$$

Now

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{T}' \end{pmatrix} \mathbf{C} = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 2 & 1 & 3 & 5 \\ 1 & 3 & 4 & 0 \\ 3 & 2 & 5 & 7 \\ 1 & 1 & 2 & 2 \\ 2 & 1 & 3 & 5 \\ 1 & 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 0 \\ -1 \end{pmatrix} = \mathbf{0}.$$

Therefore $\mathbf{K}'\boldsymbol{\beta}$ is estimable if and only if $\mathbf{K}'(3\ \ -1\ \ 0\ \ -1)' = \mathbf{0}$. If we had defined

$$\mathbf{T}' = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 3 \end{pmatrix},$$

any function of $\boldsymbol{\beta}$ would be estimable because rank $\begin{pmatrix} \mathbf{X} \\ \mathbf{T}' \end{pmatrix} = 4$. This is because $p - r = 4 - 2$ non-estimable functions are defined.

2 BLUE when β Subject to $\mathbf{T}'\beta$

One method for computing BLUE of $\mathbf{K}'\boldsymbol{\beta}$, estimable given $\mathbf{T}'\boldsymbol{\beta} = \mathbf{c}$, is $\mathbf{K}'\boldsymbol{\beta}^o$, where $\boldsymbol{\beta}^o$ is a solution to either of the following.

$$\begin{pmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{T} \\ \mathbf{T}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ \mathbf{c} \end{pmatrix}. \tag{3}$$

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{T} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{T}' & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{c} \end{pmatrix}. \tag{4}$$

If $\mathbf{T}'\boldsymbol{\beta}$ represents $p - r$ linearly independent non-estimable functions, $\boldsymbol{\beta}^o$ has a unique solution. A second method where $\mathbf{c} = \mathbf{0}$ is the following. Partition $\mathbf{T}'$, with re-ordering of columns if necessary, as

$$\mathbf{T}' = [\mathbf{T}_1' \;\; \mathbf{T}_2'],$$

the re-ordering done, if necessary, so that $\mathbf{T}_2'$ is non-singular. This of course implies that $\mathbf{T}_2$ is square. Partition $\mathbf{X} = [\mathbf{X}_1 \;\; \mathbf{X}_2]$, where $\mathbf{X}_2$ has the same number of columns as $\mathbf{T}_2'$ and with the same re-ordering of columns as in $\mathbf{T}'$. Let

$$\mathbf{W} = \mathbf{X}_1 - \mathbf{X}_2(\mathbf{T}_2')^{-1}\mathbf{T}_1'.$$

Then solve for $\boldsymbol{\beta}^o$, in either of the following two forms.

$$\mathbf{W}'\mathbf{V}^{-1}\mathbf{W}\boldsymbol{\beta}_1^o = \mathbf{W}'\mathbf{V}^{-1}\mathbf{y}. \tag{5}$$

$$\begin{pmatrix} \mathbf{W}'\mathbf{R}^{-1}\mathbf{W} & \mathbf{W}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{W} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{W}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} \tag{6}$$

In terms of the model with no definitions on the parameters,

$$E(\boldsymbol{\beta}_1^o) = (\mathbf{W}'\mathbf{V}^{-1}\mathbf{W})^{-}\mathbf{W}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}. \tag{7}$$

$$\boldsymbol{\beta}_2^o = -(\mathbf{T}_2')^{-1}\mathbf{T}_1'\boldsymbol{\beta}_1^o. \tag{8}$$

$$E(\boldsymbol{\beta}_2^o) = -(\mathbf{T}_2')^{-1}\mathbf{T}_1'\mathrm{E}(\boldsymbol{\beta}_1^o). \tag{9}$$

Let us illustrate with the same $\mathbf{X}$ used for illustrating estimability when $\mathbf{T}'\boldsymbol{\beta}$ is defined. Suppose we define

$$\mathbf{T}' = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 3 \end{pmatrix}, \; \mathbf{c} = \mathbf{0}.$$

then $\mathbf{T}'\boldsymbol{\beta}$ are non-estimable functions. Consequently the following GLS equations have a unique solution. It is assumed that $Var(\mathbf{y}) = \mathbf{I}\sigma_e^2$. The equations are

$$\sigma_e^{-2} \begin{pmatrix} 20 & 16 & 36 & 44 & 1 & 2 \\ 16 & 20 & 36 & 28 & 2 & 1 \\ 36 & 36 & 72 & 72 & 2 & 1 \\ 44 & 28 & 72 & 104 & 1 & 3 \\ 1 & 2 & 2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 3 & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} 46 \\ 52 \\ 98 \\ 86 \\ 0 \\ 0 \end{pmatrix} \sigma_e^{-2}. \tag{10}$$

and $\mathbf{y}' = [5,\ 3,\ 7,\ 2,\ 6,\ 8]$.

The solution is $[380,\ -424,\ 348,\ -228,\ 0,\ 0]/72$. If $\mathbf{T}'\boldsymbol{\beta} = \mathbf{0}$ is really true, any linear function of $\boldsymbol{\beta}$ is estimable.

By the method of (5)

$$\mathbf{X}_1' = \begin{pmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 2 & 1 & 3 & 2 & 1 & 1 \end{pmatrix},$$

$$\mathbf{X}_2' = \begin{pmatrix} 3 & 3 & 4 & 5 & 2 & 3 \\ 1 & 5 & 0 & 7 & 2 & 5 \end{pmatrix},$$

$$\mathbf{T}_1' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \mathbf{T}_2' = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix},$$

then

$$\mathbf{W}' = \begin{pmatrix} -.2 & -1.6 & .2 & -2.2 & -.6 & -1.6 \\ -1.0 & -2.0 & -1.0 & -3.0 & -1.0 & -2.0 \end{pmatrix}.$$

Equations like (5) are

$$\sigma_e^{-2} \begin{pmatrix} 10.4 & 13.6 \\ 13.6 & 20.0 \end{pmatrix} \boldsymbol{\beta}_1^o = \begin{pmatrix} -25.2 \\ -46.0 \end{pmatrix} \sigma_e^{-2}.$$

The solution is $\boldsymbol{\beta}_1^o = (380, \text{-}424)/72$. By (8)

$$\boldsymbol{\beta}_2^o = -\begin{pmatrix} .2 & 1.0 \\ .6 & 0 \end{pmatrix} \begin{pmatrix} 380 \\ -424 \end{pmatrix} /72 = \begin{pmatrix} 348 \\ -228 \end{pmatrix} /72.$$

These are identical to the result by method (3). $E(\boldsymbol{\beta}_1^o)$ by (7) is

$$\begin{pmatrix} 0 & 2.5 & 2.5 & -2.5 \\ -1.0 & -2.5 & -3.5 & -.5 \end{pmatrix} \boldsymbol{\beta}.$$

It is easy to verify that these are estimable under the restricted model.

At this point it should be noted that the computations under $\mathbf{T}'\boldsymbol{\beta} = \mathbf{c}$, where these represent $p - r$ non-estimable functions are identical with those previously described where the GLS or mixed model equation solution is restricted to $\mathbf{M}'\boldsymbol{\beta}^o = \mathbf{c}$. However, all linear functions of $\boldsymbol{\beta}$ are estimable under the restriction regarding parameters whereas they are not when these restrictions are on the solution, $\boldsymbol{\beta}^o$. Restrictions $\mathbf{M}'\boldsymbol{\beta}^o = \mathbf{c}$ are used only for convenience whereas $\mathbf{T}'\boldsymbol{\beta} = \mathbf{c}$ are used because that is part of the model.

Now let us illustrate with our same example, but with only one restriction, that being

$$(2\ \ 1\ \ 1\ \ 3)\ \boldsymbol{\beta} = 0.$$

Then equations like (3) are

$$\sigma_e^{-2}\begin{pmatrix} 20 & 16 & 36 & 44 & 2 \\ 16 & 20 & 36 & 28 & 1 \\ 36 & 36 & 72 & 72 & 1 \\ 44 & 28 & 72 & 104 & 3 \\ 2 & 1 & 1 & 3 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} 46 \\ 52 \\ 98 \\ 86 \\ 0 \end{pmatrix} \sigma_e^{-2}.$$

These do not have a unique solution, but one solution is (-88, 0, 272, -32, 0)/144. By the method of (5)

$$\begin{aligned} \mathbf{T}_1^{'} &= (2 \;\; 1 \;\; 1), \\ \mathbf{T}_2^{'} &= 3. \end{aligned}$$

This leads to

$$9^{-1}\sigma_e^{-2}\begin{pmatrix} 68 & 52 & -32 \\ 52 & 116 & 128 \\ -32 & 128 & 320 \end{pmatrix} \boldsymbol{\beta}_1^o = \begin{pmatrix} -102 \\ 210 \\ 624 \end{pmatrix} 9^{-1}\sigma_e^{-2}.$$

These do not have a unique solution but one solution is (-88 0 272)/144 as in the other method for $\boldsymbol{\beta}_1^o$.

3 Sampling Variances

If the method of (3) is used,

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) = \mathbf{K}'\mathbf{C}_{11}\mathbf{K}, \tag{11}$$

where $\mathbf{C}_{11}$ is the upper p^2 submatrix of a g-inverse of the coefficient matrix. The same is true for (4).

If the method of (5) is used

$$Var(\mathbf{K}_1^{'}\boldsymbol{\beta}_1^o) = \mathbf{K}_1^{'}(\mathbf{W}'\mathbf{V}^{-1}\mathbf{W})^{-}\mathbf{K}_1. \tag{12}$$

$$Cov(\mathbf{K}_1^{'}\boldsymbol{\beta}_1^o, \boldsymbol{\beta}_2^{o'}\mathbf{K}_2) = -\mathbf{K}_1^{'}(\mathbf{W}'\mathbf{V}^{-1}\mathbf{W})^{-}\mathbf{T}_1\mathbf{T}_2^{-1}\mathbf{K}_2. \tag{13}$$

$$Var(\mathbf{K}_2^{'}\boldsymbol{\beta}_2^o) = \mathbf{K}_2^{'}(\mathbf{T}_2^{'})^{-1}\mathbf{T}_1^{'}(\mathbf{W}'\mathbf{V}^{-1}\mathbf{W})^{-}\mathbf{T}_1\mathbf{T}_2^{-1}\mathbf{K}_2. \tag{14}$$

If the method of (6) is used, the upper part of a g-inverse of the coefficient matrix is used in place of (11).

Let us illustrate with the same example and with one restriction. A g-inverse of the coefficient matrix is

$$\frac{\sigma_e^2}{576}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 80 & -32 & -16 & 576 \\ 0 & -32 & 29 & 1 & -432 \\ 0 & -16 & 1 & 5 & 144 \\ 0 & 576 & -432 & 144 & 0 \end{pmatrix}.$$

Then

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) = \frac{\sigma_e^2}{576} \ \mathbf{K}' \begin{pmatrix} 0 & 80 & -32 & -16 \\ 0 & -32 & 29 & 1 \\ 0 & -16 & 1 & 5 \end{pmatrix} \mathbf{K}.$$

Using the method of (5) a g-inverse of $\mathbf{W}'\mathbf{V}^{-1}\mathbf{W}$ is

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 80 & -32 \\ 0 & -32 & 29 \end{pmatrix} \frac{\sigma_e^2}{576},$$

which is the same as the upper 3×3 of the matrix above. From (13)

$$-(\mathbf{W}'\mathbf{V}^{-1}\mathbf{W})^{-}\mathbf{T}_1\mathbf{T}_2^{-1} = -\frac{1}{576} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 80 & -32 \\ 0 & -32 & 29 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \frac{1}{3} = \begin{pmatrix} 0 \\ 16 \\ 1 \end{pmatrix} \frac{1}{576}$$

and for (14) $(\mathbf{T}_2')^{-1}\mathbf{T}_1'(\mathbf{W}'\mathbf{V}^{-1}\mathbf{W})^{-}\mathbf{T}_1\mathbf{T}_2^{-1} = 5/576$, thus verifying that the sampling variances are the same by the two methods.

4 Hypothesis Testing

As before let $\mathbf{H}_0'\boldsymbol{\beta} = \mathbf{c}_0$ be the null hypothesis and $\mathbf{H}_a'\boldsymbol{\beta} = \mathbf{c}_a$ be the alternative hypothesis, but now we have defined $\mathbf{T}'\boldsymbol{\beta} = \mathbf{c}$. Consequently $\mathbf{H}_0'\boldsymbol{\beta}$ and $\mathbf{H}_a'\boldsymbol{\beta}$ need be estimable only when $\mathbf{T}'\boldsymbol{\beta} = \mathbf{c}$ is assumed.

Then the tests proceed as in the unrestricted model except that for the null hypothesis computations we substitute

$$\begin{pmatrix} \mathbf{H}_0' \\ \mathbf{T}' \end{pmatrix} \boldsymbol{\beta} - \begin{pmatrix} \mathbf{c}_0 \\ \mathbf{c} \end{pmatrix} \text{ for } \mathbf{H}_0\boldsymbol{\beta} - \mathbf{c}_0. \quad (15)$$

and for the alternative hypothesis we substitute

$$\begin{pmatrix} \mathbf{H}_a' \\ \mathbf{T}' \end{pmatrix} \boldsymbol{\beta} - \begin{pmatrix} \mathbf{c}_a \\ \mathbf{c} \end{pmatrix} \text{ for } \mathbf{H}_a'\boldsymbol{\beta} - \mathbf{c}_a. \quad (16)$$

To illustrate suppose the unrestrained GLS equations are

$$\begin{pmatrix} 6 & 3 & 2 & 1 \\ 3 & 7 & 1 & 2 \\ 2 & 1 & 8 & 1 \\ 1 & 2 & 1 & 9 \end{pmatrix} \boldsymbol{\beta}^o = \begin{pmatrix} 9 \\ 12 \\ 15 \\ 16 \end{pmatrix}.$$

Suppose that we define $\mathbf{T}'\boldsymbol{\beta} = 0$ where $\mathbf{T}' = (3\ 1\ 2\ 3)$.

We wish to test $\mathbf{H}_0'\boldsymbol{\beta} = 0$, where

$$\mathbf{H}_0' = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}$$

against $\mathbf{H}_a'\boldsymbol{\beta} = 0$, where $\mathbf{H}_a' = [1\ \text{-}1\ \text{-}1\ 1]$. Note that (-1 1) $\mathbf{H}_0' = \mathbf{H}_a'$ and both are estimable. Therefore these are valid hypotheses. Using the reduction method we solve

$$\begin{pmatrix} 6 & 3 & 2 & 1 & 1 & 3 \\ 3 & 7 & 1 & 2 & -1 & 1 \\ 2 & 1 & 8 & 1 & -1 & 2 \\ 1 & 2 & 1 & 9 & 1 & 3 \\ 1 & -1 & -1 & 1 & 0 & 0 \\ 3 & 1 & 2 & 3 & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_a \\ \boldsymbol{\theta}_a \end{pmatrix} = \begin{pmatrix} 9 \\ 12 \\ 15 \\ 16 \\ 0 \\ 0 \end{pmatrix}.$$

The solution is [-1876, 795, -636, 2035, -20035, 20310]/3643, and the reduction under $\mathbf{H}_a'$ is 15676/3643 = 4.3030. Then solve

$$\begin{pmatrix} 6 & 3 & 2 & 1 & 1 & 2 & 3 \\ 3 & 7 & 1 & 2 & 2 & 1 & 1 \\ 2 & 1 & 8 & 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 9 & 0 & 1 & 3 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 3 & 1 & 2 & 3 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_0 \\ \boldsymbol{\theta}_0 \end{pmatrix} = \begin{pmatrix} 9 \\ 12 \\ 15 \\ 16 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The solution is [-348, 290, -232, 406, 4380, -5302, 5088]/836, and the reduction is 3364/836 = 4.0239. Then we test 4.3030 - 4.0239 = .2791 entering χ^2 with 1 degree of freedom coming from the differences between the number of rows in $\mathbf{H}_0'$ and $\mathbf{H}_a'$.

By the method involving $Var(\mathbf{H}_o'\boldsymbol{\beta})$ and $Var(\mathbf{H}_a'\boldsymbol{\beta})$ we solve the following equations and find a g-inverse of the coefficient matrix.

$$\begin{pmatrix} 6 & 3 & 2 & 1 & 3 \\ 3 & 7 & 1 & 2 & 1 \\ 2 & 1 & 8 & 1 & 2 \\ 1 & 2 & 1 & 9 & 3 \\ 3 & 1 & 2 & 3 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \boldsymbol{\theta}_0 \end{pmatrix} = \begin{pmatrix} 9 \\ 12 \\ 15 \\ 16 \\ 0 \end{pmatrix}.$$

The solution is [-7664, 8075, 5561, 1265, 18040]/4972. The inverse is

$$\begin{pmatrix} 624 & -276 & -336 & -308 & 1012 \\ & 887 & 53 & -55 & -352 \\ & & 659 & -121 & 220 \\ & & & 407 & 616 \\ & & & & -2024 \end{pmatrix} /4972.$$

Now

$$\mathbf{H}_0^{'}\boldsymbol{\beta}^o = [2.82522 \;\; -1.20434]',$$

$$\mathbf{H}_0^{'}\mathbf{C}_{11}\mathbf{H}_0 \;\; = \;\; \begin{pmatrix} .65708 & .09735 \\ & .27031 \end{pmatrix},$$

and

$$(\mathbf{H}_0^{'}\mathbf{C}_{11}\mathbf{H}_0)^{-1} \;\; = \;\; \begin{pmatrix} 1.60766 & -.57895 \\ & 3.90789 \end{pmatrix} = \;\mathbf{B},$$

where $\mathbf{C}_{11}$ is the upper 4×4 submatrix of the inverse of the coefficient matrix. Then

$$[2.82522 \;\; -1.20434] \; \mathbf{B} \; [2.8522 \;\; -1.20434]' \; = \; 22.44007.$$

Similarly computations with $\mathbf{H}_a^{'}$ = (1 -1 -1 1), give $\mathbf{H}_a^{'}\boldsymbol{\beta}_a$ = -4.02957, $\mathbf{B}$ = 1.36481, and $(-4.02957)\mathbf{B}(-4.02957)$ = 22.16095. Then 22.44007 - 22.16095 = .2791 as before.

Chapter 8
Unbiased Methods for G and R Unknown

C. R. Henderson

1984 - Guelph

Previous chapters have dealt with known $\mathbf{G}$ and $\mathbf{R}$ or known proportionality of these matrices. In these cases BLUE, BLUP, exact sampling variances, and exact tests of hypotheses exist. In this chapter we shall be concerned with the unsolved problem of what are "best" estimators and predictors when $\mathbf{G}$ and $\mathbf{R}$ are unknown even to proportionality. We shall construct many unbiased estimators and predictors and under certain circumstances compute their variances. Tests of hypotheses pose more serious problems, for only approximate tests can be made. We shall be concerned with three different situations regarding estimation and prediction. These are described in Henderson and Henderson (1979) and in Henderson, Jr. (1982).

1. Methods of estimation and prediction not involving $\mathbf{G}$ and $\mathbf{R}$.

2. Methods involving $\mathbf{G}$ and $\mathbf{R}$ in which assumed values, say $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ are used in the computations and these are regarded as constants.

3. The same situation as 2, but $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ are regarded more realistically as estimators from data and consequently are random variables.

1 Unbiased Estimators

Many unbiased estimators of $\mathbf{K}'\boldsymbol{\beta}$ can be computed. Some of these are much easier than GLS or mixed models with $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ used. Also some of them are invariant to $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$. The first, and one of the easiest, is ordinary least squares (OLS) ignoring $\mathbf{u}$.

Solve for $\boldsymbol{\beta}^o$ in

$$\mathbf{X}'\mathbf{X}\boldsymbol{\beta}^o \;=\; \mathbf{X}'\mathbf{y}. \tag{1}$$

Then $E(\mathbf{K}'\boldsymbol{\beta}^o) = E[\mathbf{K}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{y}] = \mathbf{K}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{K}'\boldsymbol{\beta}$ if $\mathbf{K}'\boldsymbol{\beta}$ is estimable. The variance of $\mathbf{K}'\boldsymbol{\beta}^o$ is

$$\mathbf{K}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{K}, \tag{2}$$

and this can be evaluated easily for chosen $\tilde{\mathbf{G}}$, $\tilde{\mathbf{R}}$, but it is valid only if $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ are regarded as fixed.

A second estimator is analogous to weighted least squares. Let $\mathbf{D}$ be a diagonal matrix formed from the diagonals of $(\mathbf{Z}\tilde{\mathbf{G}}\mathbf{Z}' + \tilde{\mathbf{R}})$. Then solve

$$\mathbf{X}'\mathbf{D}^{-1}\mathbf{X}\boldsymbol{\beta}^o = \mathbf{X}'\mathbf{D}^{-1}\mathbf{y}. \tag{3}$$

$\mathbf{K}'\boldsymbol{\beta}^o$ is an unbiased estimator of $\mathbf{K}'\boldsymbol{\beta}$ if estimable.

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) = \mathbf{K}'(\mathbf{X}'\mathbf{D}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{D}^{-1}(\mathbf{ZGZ}' + \mathbf{R})\mathbf{D}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{D}^{-1}\mathbf{X})^-\mathbf{K}. \tag{4}$$

A third possibility if $\tilde{\mathbf{R}}^{-1}$ is easy to compute, but $\tilde{\mathbf{V}}^{-1}$ is not easy, is to solve

$$\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X}\boldsymbol{\beta}^o = \mathbf{X}\tilde{\mathbf{R}}^{-1}\mathbf{y}. \tag{5}$$

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) = \mathbf{K}'(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^-\mathbf{X}'\tilde{\mathbf{R}}^{-1}(\mathbf{ZGZ}' + \mathbf{R})\tilde{\mathbf{R}}^{-1}\mathbf{X}(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^-\mathbf{K}. \tag{6}$$

These methods all would seem to imply that the diagonals of $\mathbf{G}^{-1}$ are large relative to diagonals of $\mathbf{R}^{-1}$.

Other methods would seem to imply just the opposite, that is, the diagonals of $\mathbf{G}^{-1}$ are small relative to $\mathbf{R}^{-1}$. One of these is OLS regarding $\mathbf{u}$ as fixed for purposes of computation. That is solve

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}. \tag{7}$$

Then if $\mathbf{K}'\boldsymbol{\beta}$ is estimable under a fixed $\mathbf{u}$ model, $\mathbf{K}'\boldsymbol{\beta}^o$ is an unbiased estimator of $\mathbf{K}'\boldsymbol{\beta}$. However, if $\mathbf{K}'\boldsymbol{\beta}$ is estimable under a random $\mathbf{u}$ model, but is not estimable under a fixed $\mathbf{u}$ model, $\mathbf{K}'\boldsymbol{\beta}^o$ may be biased. To forestall this, find a function $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ that is estimable under a fixed $\mathbf{u}$ model. Then $\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}^o$ is an unbiased estimator of $\mathbf{K}'\boldsymbol{\beta}$.

$$Var(\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}^o) = [\mathbf{K}' \;\; \mathbf{M}']\mathbf{CW}'(\mathbf{ZGZ}' + \mathbf{R})\mathbf{WC}\begin{pmatrix} \mathbf{K} \\ \mathbf{M} \end{pmatrix} \tag{8}$$

$$= (\mathbf{K}' \;\; \mathbf{M}')\mathbf{CW}'\mathbf{RWC}\begin{pmatrix} \mathbf{K} \\ \mathbf{M} \end{pmatrix} + \mathbf{M}'\mathbf{GM}, \tag{9}$$

where $\mathbf{C}$ is a g-inverse of the matrix of (7) and $\mathbf{W} = (\mathbf{X} \;\; \mathbf{Z})$.

The method of (9) is simpler than (8) if $\mathbf{R}$ has a simple form compared to $\mathbf{ZGZ}'$. In fact, if $\mathbf{R} = \mathbf{I}\sigma_e^2$, the first term of (9) becomes

$$(\mathbf{K}' \;\; \mathbf{M}')\mathbf{C} \begin{pmatrix} \mathbf{K} \\ \mathbf{M} \end{pmatrix} \sigma_e^2. \tag{10}$$

Analogous estimators would come from solving

$$\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \end{pmatrix}. \tag{11}$$

Another one would use $\mathbf{D}^{-1}$ in place of $\mathbf{R}^{-1}$ where $\mathbf{D}$ is a diagonal matrix formed from the diagonals of $\tilde{\mathbf{R}}$. In both of these last two methods $\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}^o$ would be the estimator of $\mathbf{K}'\boldsymbol{\beta}$, and we require that $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ be estimable under a fixed $\mathbf{u}$ model.

From (11)

$$\begin{aligned} Var(\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}^o) &= (\mathbf{K}' \ \mathbf{M}')\mathbf{C}\mathbf{W}'\tilde{\mathbf{R}}^{-1}(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\tilde{\mathbf{R}}^{-1}\mathbf{W}\mathbf{C}\begin{pmatrix} \mathbf{K} \\ \mathbf{M} \end{pmatrix} \\ &= (\mathbf{K}' \ \mathbf{M}')\mathbf{C}\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{R}\tilde{\mathbf{R}}^{-1}\mathbf{W}\mathbf{C}\begin{pmatrix} \mathbf{K} \\ \mathbf{M} \end{pmatrix} \\ &\quad + \mathbf{M}'\mathbf{G}\mathbf{M}. \end{aligned} \tag{12}$$

When $\mathbf{D}^{-1}$ is substituted for $\tilde{\mathbf{R}}^{-1}$ the expression in (12) is altered by making this same substitution.

Another method which is a compromise between (1) and (11) is to ignore a subvector of $\mathbf{u}$, say $\mathbf{u}_2$, then compute by OLS regarding the remaining subvector of $\mathbf{u}$, say $\mathbf{u}_1$, as fixed. The resulting equations are

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_1 \\ \mathbf{Z}_1'\mathbf{X} & \mathbf{Z}_1'\mathbf{Z}_1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}_1^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}_1'\mathbf{y} \end{pmatrix}. \tag{13}$$

$(\mathbf{Z}_1 \ \mathbf{Z}_2)$ is a partitioning of $\mathbf{Z}$ corresponding to $\mathbf{u}' = (\mathbf{u}_1' \ \mathbf{u}_2')$. Now to insure unbiasedness of the estimator of $\mathbf{K}'\boldsymbol{\beta}$ we need to find a function,

$$\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}_1,$$

that is estimable under a fixed $\mathbf{u}_1$ model. Then the unbiased estimator of $\mathbf{K}'\boldsymbol{\beta}$ is

$$\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}_1^o,$$

The variance of this estimator is

$$(\mathbf{K}' \ \mathbf{M}')\mathbf{C}\mathbf{W}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\mathbf{W}\mathbf{C} \begin{pmatrix} \mathbf{K} \\ \mathbf{M} \end{pmatrix}. \tag{14}$$

$\mathbf{W} = (\mathbf{X} \ \ \mathbf{Z}_1)$, and $\mathbf{Z}\mathbf{G}\mathbf{Z}'$ refers to the entire $\mathbf{Z}\mathbf{u}$ vector, and $\mathbf{C}$ is some g-inverse of the matrix of (13).

Let us illustrate some of these methods with a simple example.

$$\mathbf{X}' = [1 \ \ 1 \ \ 1 \ \ 1 \ \ 1],$$

$$\mathbf{Z}' = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{R} = 15\mathbf{I}, \quad \mathbf{G} = 2\mathbf{I},$$

$$\mathbf{y}' = [6\ \ 8\ \ 7\ \ 5\ \ 7].$$

$$Var(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R} = \begin{pmatrix} 17 & 2 & 0 & 0 & 0 \\ & 17 & 0 & 0 & 0 \\ & & 17 & 2 & 0 \\ & & & 17 & 0 \\ & & & & 17 \end{pmatrix}.$$

$\boldsymbol{\beta}$ is estimable. By the method of (1) we solve

$$5\boldsymbol{\beta}^o = 33.$$

$$\boldsymbol{\beta}^o = 6.6.$$

$$Var(\boldsymbol{\beta}^o) = .2\ (1\ \ 1\ \ 1\ \ 1\ \ 1)\ \ \text{Var}\ (\mathbf{y}) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}\ .2 = 3.72.$$

By the method of (7) the equations to be solved are

$$\begin{pmatrix} 5 & 2 & 2 & 1 \\ & 2 & 0 & 0 \\ & & 2 & 0 \\ & & & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} 33 \\ 14 \\ 12 \\ 7 \end{pmatrix}.$$

A solution is (0, 7, 6, 7). Because $\boldsymbol{\beta}$ is not estimable when $\mathbf{u}$ is fixed, we need some function with $\mathbf{k}' = 1$ and $\mathbf{m}'$ such that $(\mathbf{k}'\ \ \mathbf{m}')\begin{pmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{pmatrix}$ is estimable. A possibility is $(3\ \ 1\ \ 1\ \ 1)/3$. The resulting estimate is $20/3 \neq 6.6$, our previous estimate. To find the variance of the estimator by method (8) we can use a g-inverse.

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ & .5 & 0 & 0 \\ & & .5 & 0 \\ & & & 1 \end{pmatrix}.$$

$$(\mathbf{k}'\ \mathbf{m}')\mathbf{CW}' = \frac{1}{3}\ (3\ 1\ 1\ 1) \begin{pmatrix} 0 & 0 & 0 & 0 \\ & .5 & 0 & 0 \\ & & .5 & 0 \\ & & & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \frac{1}{6}\ (1\ 1\ 1\ 1\ 2).$$

Then $Var(\boldsymbol{\beta}^o) \ = \ 4 \ \neq \ 3.72$ of previous result. By the method of (9) we obtain $3.333 \ + \ .667 \ = \ 4$ also.

BLUE would be obtained by using the mixed model equations with $\mathbf{R} \ = \ 15\mathbf{I}$, $\mathbf{G} \ = \ 2\mathbf{I}$ if these are the true values of $\mathbf{R}$ and $\mathbf{G}$. The resulting equations are

$$\frac{1}{15}\begin{pmatrix} 5 & 2 & 2 & 1 \\ 2 & 9.5 & 0 & 0 \\ 2 & 0 & 9.5 & 0 \\ 1 & 0 & 0 & 8.5 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} 33 \\ 14 \\ 12 \\ 7 \end{pmatrix} /15.$$

$$\boldsymbol{\beta}^o \ = \ 6.609.$$

The upper 1×1 of a g-inverse is 3.713, which is less than for any other methods, but of course depends upon true values of $\mathbf{G}$ and $\mathbf{R}$.

2 Unbiased Predictors

The method for prediction of $\mathbf{u}$ used by most animal breeders prior to the recent general acceptance of the mixed model equations was selection index (BLP) with some estimate of $\mathbf{X}\boldsymbol{\beta}$ regarded as a parameter value. Denote the estimate of $\mathbf{X}\boldsymbol{\beta}$ by $\mathbf{X}\tilde{\boldsymbol{\beta}}$. Then the predictor of $\mathbf{u}$ is

$$\tilde{\mathbf{u}} \ = \ \tilde{\mathbf{G}}\mathbf{Z}'\tilde{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}). \tag{15}$$

$\tilde{\mathbf{G}}$ and $\tilde{\mathbf{V}}$ are estimated $\mathbf{G}$ and $\mathbf{V}$.

This method utilizes the entire data vector and the entire variance-covariance structure to predict. More commonly a subset of $\mathbf{y}$ was chosen for each individual element of $\mathbf{u}$ to be predicted, and (15) involved this reduced set of matrices and vectors.

Now if $\mathbf{X}\tilde{\boldsymbol{\beta}}$ is an unbiased estimator of $\mathbf{X}\boldsymbol{\beta}$, $E(\tilde{\mathbf{u}}) \ = \ \mathbf{0} = E(\mathbf{u})$ and is unbiased. Even if $\mathbf{G}$ and $\mathbf{R}$ were known, (15) would not represent a predictor with minimum sampling variance. We have already found that for this $\tilde{\boldsymbol{\beta}}$ should be a GLS solution. Further, in selection models (discussed in chapter 13), usual estimators for $\boldsymbol{\beta}$ such as OLS or estimators ignoring $\mathbf{u}$ are biased, so $\tilde{\mathbf{u}}$ is no longer an unbiased predictor.

Another unbiased predictor, if computed correctly, is "regressed least squares" first reported by Henderson (1948). Solve for $\mathbf{u}^o$ in equations (16).

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix} \tag{16}$$

Take a solution for which $E(\mathbf{u}^o) = \mathbf{0}$ in a fixed $\boldsymbol{\beta}$ but random $\mathbf{u}$ model. This can be done by "absorbing" $\boldsymbol{\beta}^o$ to obtain a set of equations

$$\mathbf{Z}'\mathbf{PZ}\,\mathbf{u}^o = \mathbf{Z}'\mathbf{Py}, \tag{17}$$

where

$$\mathbf{P} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'].$$

Then any solution to $\mathbf{u}^o$, usually not an unique solution, has expectation $\mathbf{0}$, because $E[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}']\mathbf{y} = (\mathbf{X} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X})\boldsymbol{\beta} = (\mathbf{X} - \mathbf{X})\boldsymbol{\beta} = \mathbf{0}$. Thus $\mathbf{u}^o$ is an unbiased predictor, but not a good one for selection, particularly if the amount of information differs greatly among individuals.

Let some g-inverse of $\mathbf{Z}'\mathbf{PZ}$ be defined as $\mathbf{C}$. Then

$$Var(\mathbf{u}^o) = \mathbf{CZ}'\mathbf{P}(\mathbf{ZGZ}' + \mathbf{R})\mathbf{PZC}, \tag{18}$$

$$Cov(\mathbf{u}, \mathbf{u}^o) = \mathbf{GZ}'\mathbf{PZC}. \tag{19}$$

Let the i^{th} diagonal of (18) be v_i, and the i^{th} diagonal of (19) be c_i, both evaluated by some estimate of $\mathbf{G}$ and $\mathbf{R}$. Then the regressed least square prediction of u_i is

$$c_i u_i^o / v_i. \tag{20}$$

This is BLP of u_i when the only observation available for prediction is u_i^o. Of course other data are available, and we could use the entire $\mathbf{u}^o$ vector for prediction of each u_i. That would give a better predictor because (18) and (19) are not diagonal matrices.

In fact, BLUP of $\mathbf{u}$ can be derived from $\mathbf{u}^o$. Denote (18) by $\mathbf{S}$ and (19) by $\mathbf{T}$. Then BLUP of $\mathbf{u}$ is

$$\mathbf{TS}^-\mathbf{u}^o, \tag{21}$$

provided $\mathbf{G}$ and $\mathbf{R}$ are known. Otherwise it would be approximate BLUP.

This is a cumbersome method as compared to using the mixed model equations, but it illustrates the reason why regressed least squares is not optimum. See Henderson (1978b) for further discussion of this method.

3 Substitution Of Fixed Values For G And R

In the methods presented above it appears that some assumption is made concerning the relative values of $\mathbf{G}$ and $\mathbf{R}$. Consequently it seems logical to use a method that approaches optimality as $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ approach $\mathbf{G}$ and $\mathbf{R}$. This would be to substitute $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ for the corresponding parameters in the mixed model equations. This is a procedure which requires no choice among a variety of unbiased methods. Further, it has the desirable property that if $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ are fixed, the estimated sampling variance and prediction error variances are simple to express. Specifically the variances and covariances estimated for $\mathbf{G} = \tilde{\mathbf{G}}$ and $\mathbf{R} = \tilde{\mathbf{R}}$ are precisely the results in (34) to (41) in Chapter 5.

It also is true that the estimators and predictors are unbiased. This is easy to prove for fixed $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ but for estimated (random) $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ we need to invoke a result by Kackar and Harville (1981) presented in Section 4. For fixed $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ note that after "absorbing" $\mathbf{u}$ from the mixed model equations we have

$$\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X}\boldsymbol{\beta}^o = \mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{y}.$$

Then

$$\begin{aligned} E(\mathbf{K}'\boldsymbol{\beta}^o) &= E(\mathbf{K}'(\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X})^-\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{y}) \\ &= \mathbf{K}'(\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X})^-\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{K}'\boldsymbol{\beta}. \end{aligned}$$

Also

$$\hat{\mathbf{u}} = (\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} + \tilde{\mathbf{G}}^{-1})^{-1}\mathbf{Z}'\tilde{\mathbf{R}}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o).$$

But $\mathbf{X}\boldsymbol{\beta}^o$ is an unbiased estimator of $\mathbf{X}\boldsymbol{\beta}$, $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o$ with expectation $\mathbf{0}$ and consequently $E(\hat{\mathbf{u}}) = \mathbf{0}$ and is unbiased.

4 Mixed Model Equations With Estimated G and R

It is not a trivial problem to find the expectations of $\mathbf{K}'\boldsymbol{\beta}^o$ and $\hat{\mathbf{u}}$ from mixed model equations with estimated $\mathbf{G}$ and $\mathbf{R}$. Kackar and Harville (1981) derived a very important result for this case. They prove that if $\mathbf{G}$ and $\mathbf{R}$ are estimated by a method having the following properties and substituted in mixed model equations, the resulting estimators and predictors are unbiased. This result requires that

1. $\mathbf{y}$ is symmetrically distributed, that is, $f(\mathbf{y}) = f(-\mathbf{y})$.

2. The estimators of $\mathbf{G}$ and $\mathbf{R}$ are translation invariant.

3. The estimators of $\mathbf{G}$ and $\mathbf{R}$ are even functions of $\mathbf{y}$.

These are not very restrictive requirements because they include a variety of distributions of $\mathbf{y}$ and most of the presently used methods for estimation of variances and covariances.

An interesting consequence of substituting ML estimates of $\mathbf{G}$ and $\mathbf{R}$ for the corresponding parameters of mixed model equations is that the resulting $\mathbf{K}'\boldsymbol{\beta}^o$ are ML and the $\hat{\mathbf{u}}$ are ML of $(\mathbf{u} \mid \mathbf{y})$.

5 Tests Of Hypotheses Concerning β

We have seen that unbiased estimators and predictors can be obtained even though $\mathbf{G}$ and $\mathbf{R}$ are unknown. When it comes to testing hypotheses regarding $\boldsymbol{\beta}$ little is known except that exact tests do not exist apart from a special case that is described below. The problem is that quadratics in $\mathbf{H}'\boldsymbol{\beta}^o - \mathbf{c}$ appropriate for exact tests when $\mathbf{G}$ and $\mathbf{R}$ are known, do not have a χ^2 or any other tractable distribution when $\tilde{\mathbf{G}}$, $\tilde{\mathbf{R}}$ replace $\mathbf{G}$, $\mathbf{R}$ in the computation. What should be done? One possibility is to estimate, if possible $\mathbf{G}$, $\mathbf{R}, \boldsymbol{\beta}$ by ML and then invoke a likelihood ratio test, in which under normality assumptions and large samples, -2 log likelihood ratio is approximated by χ^2. This raises the question of what is a large sample of unbalanced data. Certainly $n \rightarrow \infty$ is not a sufficient condition. Consideration needs to be given to the number of levels of each subvector of $\mathbf{u}$ and to the proportion of missing subclasses. Consequently the value of a χ^2 approximation to the likelihood ratio test is uncertain.

A second and easier approximation is to pretend that $\tilde{\mathbf{G}} = \mathbf{G}$ and $\tilde{\mathbf{R}} = \mathbf{R}$ and proceed to an approximate test using χ^2 as described in Chapter 4 for hypothesis testing with known $\mathbf{G}$, $\mathbf{R}$ and normality assumptions. The validity of this test must surely depend, as it does in the likelihood ratio approximation, upon the number of levels of $\mathbf{u}$ and the balance and lack of missing subclasses.

One interesting case exists in which exact tests of $\boldsymbol{\beta}$ can be made even when we do not know $\mathbf{G}$ and $\mathbf{R}$ to proportionality. The requirements are as follows

1. $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$, and

2. $\mathbf{H}_0'\boldsymbol{\beta}$ is estimable under a fixed $\mathbf{u}$ model.

Solve for $\boldsymbol{\beta}^o$ in equations (7). Then

$$Var(\mathbf{H}_0'\boldsymbol{\beta}^o) = \mathbf{H}_0'\mathbf{C}_{11}\mathbf{H}_0\sigma_e^2 \tag{22}$$

where $\mathbf{C}_{11}$ is the upper $p \times p$ submatrix of a g-inverse of the coefficient matrix. Then under the null hypothesis versus the unrestricted hypothesis

$$(\mathbf{H}_0'\boldsymbol{\beta}^o)' \ [\mathbf{H}_0'\mathbf{C}_{11}\ \mathbf{H}_0]^{-1} \ \mathbf{H}_0'\boldsymbol{\beta}^o/s\hat{\boldsymbol{\sigma}}_e^2 \tag{23}$$

is distributed as F with degrees of freedom s, $n-$rank $(\mathbf{X} \ \ \mathbf{Z})$. $\hat{\sigma}_e^2$ is an estimate of σ_e^2 computed by

$$(\mathbf{y}'\mathbf{y} - (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{y} - (\mathbf{u}^o)'\mathbf{Z}'\mathbf{y})/[n - rank(\mathbf{X} \ \ \mathbf{Z})], \tag{24}$$

and s is the number of rows, linearly independent, in $\mathbf{H}_0'$.

Chapter 9
Biased Estimation and Prediction

C. R. Henderson

1984 - Guelph

All methods for estimation and prediction in previous chapters have been unbiased. In this chapter we relax the requirement of unbiasedness and attempt to minimize the mean squared error of estimation and prediction. Mean squared error refers to the sum of prediction error variance plus squared bias. In general, biased predictors and estimators exist that have smaller mean squared errors than BLUE and BLUP. Unfortunately, we never know what are truly minimum mean squared error estimators and predictors because we do not know some of the parameters required for deriving them. But even for BLUE and BLUP we must know $\mathbf{G}$ and $\mathbf{R}$ at least to proportionality. Additionally for minimum mean squared error we need to know squares and products of $\boldsymbol{\beta}$ at least proportionally to $\mathbf{G}$ and $\mathbf{R}$.

1 Derivation Of BLBE And BLBP

Suppose we want to predict $\mathbf{k}_1'\boldsymbol{\beta}_1 + \mathbf{k}_2'\boldsymbol{\beta}_2 + \mathbf{m}'\mathbf{u}$ by a linear function of $\mathbf{y}$, say $\mathbf{a}'\mathbf{y}$, such that the predictor has expectation $\mathbf{k}_1'\boldsymbol{\beta}_1$ plus some function of $\boldsymbol{\beta}_2$, and in the class of such predictors, has minimum mean squared error of prediction, which we shall call BLBP (best linear biased predictor).

The mean squared error (MSE) is

$$\mathbf{a}'\mathbf{R}\mathbf{a} + (\mathbf{a}'\mathbf{X}_2 - \mathbf{k}_2')\boldsymbol{\beta}_2\boldsymbol{\beta}_2'(\mathbf{X}_2'\mathbf{a} - \mathbf{k}_2) + (\mathbf{a}'\mathbf{Z} - \mathbf{m}')\mathbf{G}(\mathbf{Z}'\mathbf{a} - \mathbf{m}). \quad (1)$$

In order that $E(\mathbf{a}'\mathbf{y})$ contains $\mathbf{k}_1'\boldsymbol{\beta}_1$ it is necessary that $\mathbf{a}'\mathbf{X}_1\boldsymbol{\beta}_1 = \mathbf{k}_1'\boldsymbol{\beta}_1$, and this will be true for any $\boldsymbol{\beta}_1$ if $\mathbf{a}'\mathbf{X}_1 = \mathbf{k}_1'$. Consequently we minimize (1) subject to this condition. Differentiating (1) with respect to $\mathbf{a}$ and to an appropriate Lagrange Multiplier, we have equations (2) to solve.

$$\begin{pmatrix} \mathbf{V} + \mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2' & \mathbf{X}_1 \\ \mathbf{X}_1' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{Z}\mathbf{G}\mathbf{m} + \mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{k}_2 \\ \mathbf{k}_1 \end{pmatrix}. \quad (2)$$

$\mathbf{a}$ has a unique solution if and only if $\mathbf{k}_1'\boldsymbol{\beta}_1$ is estimable under a model in which $E(\mathbf{y})$ contains $\mathbf{X}_1\boldsymbol{\beta}_1$. The analogy to GLS of $\boldsymbol{\beta}_1$ is a solution to (3).

$$\mathbf{X}_1'(\mathbf{V} + \mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2')^{-1}\mathbf{X}_1\boldsymbol{\beta}_1^* = \mathbf{X}_1'(\mathbf{V} + \mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2')^{-1}\mathbf{y}. \quad (3)$$

Then if $\mathbf{K}_1'\boldsymbol{\beta}_1$ is estimable under a model, $E(\mathbf{y})$ containing $\mathbf{X}_1\boldsymbol{\beta}_1$, $\mathbf{K}_1'\boldsymbol{\beta}_1^*$ is unique and is the minimum MSE estimator of $\mathbf{K}_1'\boldsymbol{\beta}_1$. The BLBE of $\boldsymbol{\beta}_2$ is

$$\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2'(\mathbf{V} + \mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2')^{-1}(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1^*) \tag{4}$$

$\equiv \boldsymbol{\beta}_2^*$, and this is unique provided $\mathbf{K}_1\boldsymbol{\beta}_1$ is estimable when $E(\mathbf{y})$ contains $\mathbf{X}_1\boldsymbol{\beta}_1$. The BLBP of $\mathbf{u}$ is

$$\mathbf{u}^* = \mathbf{G}\mathbf{Z}'(\mathbf{V} + \mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2')^{-1}(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1^*), \tag{5}$$

and this is unique. Furthermore BLBP of

$$\mathbf{K}_1'\boldsymbol{\beta}_1 + \mathbf{K}_2'\boldsymbol{\beta}_2 + \mathbf{M}'\mathbf{u} \text{ is } \mathbf{K}_1'\boldsymbol{\beta}_1^* + \mathbf{K}_2'\boldsymbol{\beta}_2^* + \mathbf{M}'\mathbf{u}^*. \tag{6}$$

We know that BLUE and BLUP can be computed from mixed model equations. Similarly $\boldsymbol{\beta}_1^*$, $\boldsymbol{\beta}_2^*$, and $\mathbf{u}^*$ can be obtained from modified mixed model equations (7), (8), or (9). Let $\boldsymbol{\beta}_2\boldsymbol{\beta}_2' = \mathbf{P}$. Then with $\mathbf{P}$ singular we can solve (7).

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{P}\mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{P}\mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_2 + \mathbf{I} & \mathbf{P}\mathbf{X}_2'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1^* \\ \boldsymbol{\beta}_2^* \\ \mathbf{u}^* \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{P}\mathbf{X}_2'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} \tag{7}$$

The rank of this coefficient matrix is rank $(\mathbf{X}_1) + p_2 + q$, where $p_2 =$ the number of elements in $\boldsymbol{\beta}_2$. The solution to β_2^* and u^* is unique but β_1^* is not unless $\mathbf{X}_1$ has full column rank. Note that the coefficient matrix is non-symmetric. If we prefer a symmetric matrix, we can use equations (8).

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_2\mathbf{P} & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{P}\mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{P}\mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_2\mathbf{P} + \mathbf{P} & \mathbf{P}\mathbf{X}_2'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_2\mathbf{P} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1^* \\ \boldsymbol{\alpha}_2^* \\ \mathbf{u}^* \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{P}\mathbf{X}_2'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} \tag{8}$$

Then $\boldsymbol{\beta}_2^* = \mathbf{P}\boldsymbol{\alpha}_2^*$. The rank of this coefficient matrix is rank $(\mathbf{X}_1)$ + rank $(\mathbf{P}) + \mathbf{q}$. $\mathbf{K}_1'\boldsymbol{\beta}_1^*$, $\boldsymbol{\beta}_2^*$, and $\mathbf{u}^*$ are identical to the solution from (7). If $\mathbf{P}$ were non-singular we could use equations (9).

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_2 + \mathbf{P}^{-1} & \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1^* \\ \boldsymbol{\beta}_2^* \\ \mathbf{u}^* \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} \tag{9}$$

The rank of this coefficient matrix is rank $(\mathbf{X}_1) + p_2 + q$.

Usually $\mathbf{R}$, $\mathbf{G}$, and $\mathbf{P}$ are unknown, so we need to use guesses or estimates of them, say $\tilde{\mathbf{R}}$, $\tilde{\mathbf{G}}$, and $\tilde{\mathbf{P}}$. These would be used in place of the parameter values in (2) through (9).

In all of these except (9) the solution to $\boldsymbol{\beta}_2^*$ has a peculiar and seemingly undesirable property, namely $\boldsymbol{\beta}_2^* = k\tilde{\boldsymbol{\beta}}_2$, where k is some constant. That is, the elements of $\boldsymbol{\beta}_2^*$ are proportional to the elements of $\tilde{\boldsymbol{\beta}}_2$. Also it should be noted that if, as should always be the case, $\mathbf{P}$ is positive definite or positive semi-definite, the elements of $\boldsymbol{\beta}_2^*$ are "shrunken" (are nearer to 0) compared to the elements of the GLS solution to $\boldsymbol{\beta}_2$ when $\mathbf{X}_2$ is full column rank. This is comparable to the fact that BLUP of elements of $\mathbf{u}$ are smaller in absolute value than are the corresponding GLS computed as though $\mathbf{u}$ were fixed. This last property of course creates bias due to $\boldsymbol{\beta}_2$ but may reduce mean squared errors.

2 Use Of An External Estimate Of β

We next consider methods for utilizing an external estimate of $\boldsymbol{\beta}$ in order to obtain a better unbiased estimator from a new data set. For this purpose it will be simplest to assume that in both the previous experiments and the present one the rank of $\mathbf{X}$ is $r \leq p$ and that the same linear dependencies among columns of $\mathbf{X}$ existed in both cases. With possible re-ordering the full rank subset is denoted by $\mathbf{X}_1$ and the corresponding $\boldsymbol{\beta}$ by $\boldsymbol{\beta}_1$. Suppose we have a previous solution to $\boldsymbol{\beta}_1$ denoted by $\boldsymbol{\beta}_1^*$ and $E(\boldsymbol{\beta}_1^*) = \boldsymbol{\beta}_1 + \mathbf{L}\boldsymbol{\beta}_2$ where $\mathbf{X} = (\mathbf{X}_1 \quad \mathbf{X}_2)$ and $\mathbf{X}_2 = (\mathbf{X}_1\mathbf{L})$. Further $Var(\boldsymbol{\beta}_1^*) = \mathbf{V}_1$. Assuming logically that the prior estimator is uncorrelated with the present data vector, $\mathbf{y}$, the GLS equations are

$$(\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_1 + \mathbf{V}_1^{-1})\hat{\boldsymbol{\beta}}_1 = \mathbf{X}_1'\mathbf{V}^{-1}\mathbf{y} + \mathbf{V}_1^{-1}\boldsymbol{\beta}_1^*. \tag{10}$$

Then BLUE of $\mathbf{K}'\boldsymbol{\beta}$, where $\mathbf{K}'$ has the form $(\mathbf{K}_1' \quad \mathbf{K}_1'\mathbf{L})$ is $\mathbf{K}_1'\hat{\boldsymbol{\beta}}_1$, and its variance is

$$\mathbf{K}_1'(\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_1 + \mathbf{V}_1^{-1})^{-1}\mathbf{K}_1. \tag{11}$$

The mixed model equations corresponding to (10) are

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1 + \mathbf{V}_1^{-1} & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{y} + \mathbf{V}_1^{-1}\boldsymbol{\beta}_1^* \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{12}$$

3 Assumed Pattern Of Values Of β

The previous methods of this chapter requiring prior values of every element of $\boldsymbol{\beta}$ and resulting estimates with the same proportionality as the prior is rather distasteful. A possible alternative solution is to assume a pattern of values of $\boldsymbol{\beta}$ with less than p

parameters. For example, with two way, fixed, cross-classified factors with interaction we might assume in some situations that there is no logical pattern of values for interactions. Defining for convenience that the interactions sum to 0 across each row and each column, and then considering all possible permutations of the labelling of rows and columns, the following is true for the average squares and products of these interactions. Define the interaction for the ij^{th} cell as α_{ij} and define the number of rows as r and the number of columns as c. The average values are as follows.

$$\begin{aligned} \alpha_{ij}^2 &= \gamma, & (13) \\ \alpha_{ij}\alpha_{ij'} &= -\gamma/(c-1), & (14) \\ \alpha_{ij}\alpha_{i'j} &= -\gamma/(r-1), & (15) \\ \alpha_{ij}\alpha_{i'j'} &= \gamma/(c-1)(r-1). & (16) \end{aligned}$$

Then if we have some prior value of γ we can proceed to obtain locally minimum mean squared error estimators and predictors as follows. Let $\mathbf{P}$ = estimated average value of $\boldsymbol{\beta}_2\ \boldsymbol{\beta}_2^{'}$. Then solve equations (7), (8) or (9).

4 Evaluation Of Bias

If we are to consider biased estimation and prediction, we should know how to evaluate the bias. We do this by looking at expectations. A method applied to (7) is as follows. Remember that $\mathbf{K}_1\boldsymbol{\beta}_1^*$ is required to have expectation, $\mathbf{K}_1^{'}\boldsymbol{\beta}_1+$ some linear function of $\boldsymbol{\beta}_2$. For this to be true $\mathbf{K}_1^{'}\boldsymbol{\beta}_1$ must be estimable under a model with $\mathbf{X}_2\boldsymbol{\beta}_2$ not existing. $\boldsymbol{\beta}_2^*$ and $\mathbf{u}^*$ are required to have expectation that is some linear function of $\boldsymbol{\beta}_2$.

Let some g-inverse of the matrix of (7) be

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{C}_{33} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \mathbf{C}_3 \end{pmatrix}. \quad (17)$$

Then

$$E(\mathbf{K}_1^{'}\boldsymbol{\beta}_1^*) = \mathbf{K}_1^{'}\boldsymbol{\beta}_1 + \mathbf{K}_1^{'}\mathbf{C}_1\mathbf{T}\boldsymbol{\beta}_2, \quad (18)$$

where

$$\mathbf{T} = \begin{pmatrix} \mathbf{X}_1^{'}\tilde{\mathbf{R}}^{-1}\mathbf{X}_2 \\ \mathbf{P}\mathbf{X}_2^{'}\tilde{\mathbf{R}}^{-1}\mathbf{X}_2 \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X}_2 \end{pmatrix}.$$

$$\begin{aligned} E(\boldsymbol{\beta}_2^*) &= \mathbf{C}_2\mathbf{T}\boldsymbol{\beta}_2. & (19) \\ E(\mathbf{u}^*) &= \mathbf{C}_3\mathbf{T}\boldsymbol{\beta}_2. & (20) \end{aligned}$$

Then the biases are as follows.

$$\begin{aligned}\text{For } \mathbf{K}_1\boldsymbol{\beta}_1^*, \text{ bias} &= \mathbf{K}_1'\mathbf{C}_1\mathbf{T}\boldsymbol{\beta}_2. && (21)\\ \text{For } \boldsymbol{\beta}_2^*, \text{ bias} &= (\mathbf{C}_2\mathbf{T} - \mathbf{I})\boldsymbol{\beta}_2. && (22)\\ \text{For } \mathbf{u}^*, \text{ bias} &= \mathbf{C}_3\mathbf{T}\boldsymbol{\beta}_2. && (23)\end{aligned}$$

If the equations (8) are used, the biases are the same as in (21), (22), and (23) except that (22) is premultiplied by $\mathbf{P}$, and $\mathbf{C}$ refers to a g-inverse of the matrix of (8). If the equations of (9) are used, the second term of $\mathbf{T}$ is $\mathbf{X}_2'\tilde{\mathbf{R}}^{-1}\mathbf{X}_2$, and $\mathbf{C}$ refers to the inverse of the matrix of (9).

5 Evaluation Of Mean Squared Errors

If we are to use biased estimation and prediction, we should know how to estimate mean squared errors of estimation and prediction. For the method of (7) proceed as follows. Let

$$\mathbf{T} = \begin{pmatrix} \mathbf{X}_1'\tilde{\mathbf{R}}^{-1}\mathbf{X}_2 \\ \mathbf{P}\mathbf{X}_2'\tilde{\mathbf{R}}^{-1}\mathbf{X}_2 \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X}_2 \end{pmatrix}. \tag{24}$$

Note the similarity to the second "column" of the matrix of (7). Let

$$\mathbf{S} = \begin{pmatrix} \mathbf{X}_1'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{P}\mathbf{X}_2'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \end{pmatrix}. \tag{25}$$

Note the similarity to the third "column" of the matrix of (7). Let

$$\mathbf{H} = \begin{pmatrix} \mathbf{X}_1'\tilde{\mathbf{R}}^{-1} \\ \mathbf{P}\mathbf{X}_2'\tilde{\mathbf{R}}^{-1} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1} \end{pmatrix}. \tag{26}$$

Note the similarity to the right hand side of (7). Then compute

$$\begin{aligned}&\begin{pmatrix} \mathbf{C}_1\mathbf{T} \\ \mathbf{C}_2\mathbf{T} - \mathbf{I} \\ \mathbf{C}_3\mathbf{T} \end{pmatrix} \boldsymbol{\beta}_2\boldsymbol{\beta}_2'(\mathbf{T}'\mathbf{C}_1' \quad \mathbf{T}'\mathbf{C}_2' - \mathbf{I} \quad \mathbf{T}'\mathbf{C}_3') \\ &+ \begin{pmatrix} \mathbf{C}_1\mathbf{S} \\ \mathbf{C}_2\mathbf{S} \\ \mathbf{C}_3\mathbf{S} - \mathbf{I} \end{pmatrix} \mathbf{G} \quad (\mathbf{S}'\mathbf{C}_1' \quad \mathbf{S}'\mathbf{C}_2' \quad \mathbf{S}'\mathbf{C}_3' - \mathbf{I}) \\ &+ \begin{pmatrix} \mathbf{C}_1\mathbf{H} \\ \mathbf{C}_2\mathbf{H} \\ \mathbf{C}_3\mathbf{H} \end{pmatrix} \mathbf{R} \quad (\mathbf{H}'\mathbf{C}_1' \quad \mathbf{H}'\mathbf{C}_2' \quad \mathbf{H}'\mathbf{C}_3')\end{aligned}$$

$$= \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} \end{pmatrix} = \mathbf{B}. \tag{27}$$

Then mean squared error of

$$(\mathbf{M}_1' \;\; \mathbf{M}_2' \;\; \mathbf{M}_3') \begin{pmatrix} \boldsymbol{\beta}_1^* \\ \boldsymbol{\beta}_2^* \\ \mathbf{u}^* - \mathbf{u} \end{pmatrix}$$

$$= (\mathbf{M}_1' \;\; \mathbf{M}_2' \;\; \mathbf{M}_3') \;\; \mathbf{B} \begin{pmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \\ \mathbf{M}_3 \end{pmatrix}. \tag{28}$$

Of course this cannot be evaluated numerically except for assumed values of $\boldsymbol{\beta}$, $\mathbf{G}$, $\mathbf{R}$. The result simplifies remarkably if we evaluate at the same values used in (7), namely $\boldsymbol{\beta}_2\boldsymbol{\beta}_2' = \tilde{\mathbf{P}}$, $\mathbf{G} = \tilde{\mathbf{G}}$, $\mathbf{R} = \tilde{\mathbf{R}}$. Then $\mathbf{B}$ is simply

$$\mathbf{C} \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12}\mathbf{P} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{C}_{22}\mathbf{P} & \mathbf{C}_{23} \\ \mathbf{C}_{31} & \mathbf{C}_{32}\mathbf{P} & \mathbf{C}_{33} \end{pmatrix}. \tag{29}$$

$\mathbf{C}$ and $\mathbf{C}_{ij}$ are defined in (9.17).

When the method of (8) is used, modify the result for (7) as follows. Let a g-inverse of the matrix of (8) be

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{C}_{33} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \mathbf{C}_3 \end{pmatrix} = \mathbf{C}. \tag{30}$$

Substitute $\tilde{\mathbf{P}}\mathbf{C}_2\mathbf{T} - \mathbf{I}$ for $\mathbf{C}_2\mathbf{T} - \mathbf{I}$, $\tilde{\mathbf{P}}\mathbf{C}_2\mathbf{S}$ for $\mathbf{C}_2\mathbf{S}$, and $\tilde{\mathbf{P}}\mathbf{C}_2\mathbf{H}$ for $\mathbf{C}_2\mathbf{H}$ and proceed as in (28) using the $\mathbf{C}_i$ from (29). If $\mathbf{P} = \tilde{\mathbf{P}}$, $\mathbf{G} = \tilde{\mathbf{G}}$, $\mathbf{R} = \tilde{\mathbf{R}}$, $\mathbf{B}$ simplifies to

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix} \mathbf{C} \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix}. \tag{31}$$

If the method of (9) is used, delete $\mathbf{P}$ from $\mathbf{T}$, $\mathbf{S}$, and $\mathbf{H}$ in (24), (25), and (26), let $\mathbf{C}$ be a g-inverse of the matrix of (9), and then proceed as for method (7). When $\mathbf{P} = \tilde{\mathbf{P}}$, $\mathbf{G} = \tilde{\mathbf{G}}$, and $\mathbf{R} = \tilde{\mathbf{R}}$, the simple result, $\mathbf{B} = \mathbf{C}$ can be used.

6 Estimability In Biased Estimation

The traditional understanding of estimability in the linear model is that $\mathbf{K}'\boldsymbol{\beta}$ is defined as estimable if some linear function of $\mathbf{y}$ exists that has expectation $\mathbf{K}'\boldsymbol{\beta}$, and thus this linear function is an unbiased estimator. But if we relax the requirement of unbiasedness, is the above an appropriate definition of estimability? Is any function of $\boldsymbol{\beta}$ now estimable? It seems reasonable to me to restrict estimation to functions that could be estimated if we had no missing subclasses. Otherwise we could estimate elements of $\boldsymbol{\beta}$ that have no relevance to the experiment in question. For example, treatments involve levels of protein in the ration. Just because we invoke biased estimation of treatments would hardly seem to warrant estimation of some treatment that has nothing to do with level of protein. Consequently we state these rules for functions that can be estimated biasedly.

1. We want to estimate $\mathbf{K}_1'\boldsymbol{\beta}_1 + \mathbf{K}_2'\boldsymbol{\beta}_2$, where a prior on $\boldsymbol{\beta}_2$ is used.

2. If $\mathbf{K}_1'\boldsymbol{\beta}_1+\mathbf{K}_2'\boldsymbol{\beta}_2$ were estimable with no missing subclasses, this function is a candidate for estimation.

3. $\mathbf{K}_1'\boldsymbol{\beta}_1$ must be estimable under a model in which $E(\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta}_1$.

4. $\mathbf{K}_1'\boldsymbol{\beta}_1 + \mathbf{K}_2'\boldsymbol{\beta}_2$ does not need to be estimable in the sample, but must be estimable in the filled subclass case.

Then $\mathbf{K}_1'\boldsymbol{\beta}_1^o + \mathbf{K}_2'\boldsymbol{\beta}_2^o$ is invariant to the solution to (7),(8), or (9). Let us illustrate with a model

$$y_{ij} = \mu + t_i + e_{ij} \quad , \quad i = 1, 2, 3.$$

Suppose that the numbers of observations per treatment are (5, 3, 0). However, we are willing to assume prior values for squares and products of $t_1,\ t_2,\ t_3$ even though we have no data on t_3. The following functions would be estimable if $n_3 > 0$,

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ t_1 \\ t_2 \\ t_3 \end{pmatrix}.$$

Further with $\boldsymbol{\beta}_1$ being just μ, and $\mathbf{K}_1'$ being 1, and $\mathbf{X}_1' = (1\ 1\ 1)$, $\mathbf{K}_1'\boldsymbol{\beta}_1$ is estimable under a model $E(y_{ij}) = \mu$.

Suppose in contrast that we want to impose a prior on just t_3 . Then $\boldsymbol{\beta}_1' = (\mu\ t_1\ t_2)$ and $\boldsymbol{\beta}_2 = t_3$. Now

$$\mathbf{K}_1'\boldsymbol{\beta}_1' = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ t_1 \\ t_2 \end{pmatrix}.$$

But the third row represents a non-estimable function. That is, μ is not estimable under the model with $\beta_1' = (\mu\ t_1\ t_2)$. Consequently $\mu + t_3$ should not be estimated in this way.

As another example suppose we have a 2×3 fixed model with $n_{23} = 0$ and all other $n_{ij} > 0$. We want to estimate all six $\mu_{ij} = \mu + a_i + b_j + \gamma_{ij}$. With no missing subclasses these are estimable, so they are candidates for estimation. Suppose we use priors on $\boldsymbol{\gamma}$. Then

$$(\mathbf{K}_1' \;\; \mathbf{K}_2') \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ a_1 \\ a_2 \\ b_1 \\ b_2 \\ b_3 \\ \boldsymbol{\gamma} \end{pmatrix}.$$

Now $\mathbf{K}_1'\boldsymbol{\beta}_1$ is estimable under a model, $E(y_{ijk}) = \mu + a_i + b_j$. Consequently we can by our rules estimate all six μ_{ij} . These will have expectations as follows.

$$E(\hat{\mu}_{ij}) = \mu + a_i + b_j + \text{ some function of } \gamma \;\neq\; \mu + a_i + b_j + \gamma_{ij}.$$

Now suppose we wish to estimate by using a prior only on γ_{23}. Then the last row of $\mathbf{K}_1'\boldsymbol{\beta}$ is $\mu + a_2 + b_3$ but this is not estimable under a model

$$E \begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{pmatrix} = \begin{pmatrix} \mu + a_1 + b_1 + \gamma_{11} \\ \mu + a_1 + b_2 + \gamma_{12} \\ \mu + a_1 + b_3 + \gamma_{13} \\ \mu + a_2 + b_1 + \gamma_{21} \\ \mu + a_2 + b_2 + \gamma_{22} \\ \mu + a_2 + b_3 \end{pmatrix}.$$

Consequently we should not use a prior on just γ_{23}.

7 Tests Of Hypotheses

Exact tests of hypotheses do not exist when biased estimation is used, but one might wish to use the following approximate tests that are based on using mean squared error of $\mathbf{K}'\boldsymbol{\beta}^o$ rather than $Var(\mathbf{K}'\boldsymbol{\beta}^o)$.

7.1 $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$

When $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$ write (7) as (32) or (8) as (33). Using the notation of Chapter 6, $\mathbf{G} = \mathbf{G}_*\sigma_e^2$ and $\mathbf{P} = \mathbf{P}_*\sigma_e^2$.

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 & \mathbf{X}_1'\mathbf{Z} \\ \tilde{\mathbf{P}}_*\mathbf{X}_2'\mathbf{X}_1 & \tilde{\mathbf{P}}_*\mathbf{X}_2'\mathbf{X}_2 + \mathbf{I} & \tilde{\mathbf{P}}_*\mathbf{X}_2'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X}_1 & \mathbf{Z}'\mathbf{X}_2 & \mathbf{Z}'\mathbf{Z} + \mathbf{G}_*^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1^* \\ \boldsymbol{\beta}_2^* \\ \mathbf{u}^* \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{y} \\ \tilde{\mathbf{P}}_*\mathbf{X}_2'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}. \quad (32)$$

The corresponding equations with symmetric coefficient matrix are in (33).

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2\tilde{\mathbf{P}}_* & \mathbf{X}_1'\mathbf{Z} \\ \tilde{\mathbf{P}}_*\mathbf{X}_2'\mathbf{X}_1 & \tilde{\mathbf{P}}_*\mathbf{X}_2'\mathbf{X}_2\tilde{\mathbf{P}}_* + \tilde{\mathbf{P}}_* & \tilde{\mathbf{P}}_*\mathbf{X}_2'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X}_1 & \mathbf{Z}'\mathbf{X}_2\tilde{\mathbf{P}}_* & \mathbf{Z}'\mathbf{Z} + \mathbf{G}_*^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1^* \\ \boldsymbol{\alpha}^* \\ \mathbf{u}^* \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{y} \\ \tilde{\mathbf{P}}_*\mathbf{X}_2'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix} \tag{33}$$

Then $\boldsymbol{\beta}_2^* = \tilde{\mathbf{P}}_*\boldsymbol{\alpha}_2^*$.

Let a g-inverse of the matrix of (32) post-multiplied by

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix} \equiv \mathbf{Q}$$

or a g-inverse of the matrix (33) pre-multipled and post-multiplied by $\mathbf{Q}$ be denoted by

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix},$$

where $\mathbf{C}_{11}$ has order $p \times p$ and $\mathbf{C}_{22}$ has order $q \times q$. Then if $\tilde{\mathbf{P}}_* = \mathbf{P}_*$, mean squared error of $\mathbf{K}'\boldsymbol{\beta}^*$ is $\mathbf{K}'\mathbf{C}_{11}\mathbf{K}\sigma_e^2$. Then

$$(\mathbf{K}'\boldsymbol{\beta}^* - \mathbf{c})'[\mathbf{K}'\mathbf{C}_{11}\mathbf{K}]^{-1}(\mathbf{K}'\hat{\boldsymbol{\beta}}^* - \mathbf{c})/s\ \hat{\sigma}_e^2$$

is distributed under the null hypothesis approximately as F with $s,\ t$ degrees of freedom, where s = number of rows (linearly independent) in $\mathbf{K}'$, and $\hat{\sigma}_e^2$ is estimated unbiasedly with t degrees of freedom.

7.2 $Var(\mathbf{e}) = \mathbf{R}$

Let g-inverse of (7) post-multiplied by

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix} \equiv \mathbf{Q}$$

or a g-inverse of (8) pre-multiplied and post-multiplied by $\mathbf{Q}$ be denoted by

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix}.$$

Then if $\tilde{\mathbf{R}} = \mathbf{R}, \tilde{\mathbf{G}} = \mathbf{G}$, and $\tilde{\mathbf{P}} = \mathbf{P}$, $\mathbf{K}'\mathbf{C}_{11}\mathbf{K}$ is the mean squared error of $\mathbf{K}'\boldsymbol{\beta}^*$, and $(\mathbf{K}'\boldsymbol{\beta}^* - \mathbf{c})'(\mathbf{K}'\mathbf{C}_{11}\mathbf{K})^{-1}(\mathbf{K}'\boldsymbol{\beta}^* - \mathbf{c})$ is distributed approximately as χ^2 with s degrees of freedom under the null hypothesis, $\mathbf{K}'\boldsymbol{\beta} = \mathbf{c}$.

8 Estimation of P

If one is to use biased estimation and prediction, one would usually have to estimate $\mathbf{P}$, ordinarily a singular matrix. If the elements of $\boldsymbol{\beta}_2$ are thought to have no particular pattern, permutation theory might be used to derive average values of squares and products of elements of $\boldsymbol{\beta}_2$, that is the value of $\mathbf{P}$. We might then formulate this as estimation of a variance covariance matrix, usually with fewer parameters than $t(t+1)/2$, where t is the order of $\mathbf{P}$. I think I would estimate these parameters by the MIVQUE method for singular $\mathbf{G}$ described in Section 9 of Chapter 11 or by REML of Chapter 12.

9 Illustration

We illustrate biased estimation by a 3-way mixed model. The model is

$$y_{hijk} = r_h + c_i + \gamma_{hi} + u_j + e_{ijk},$$
$$r, c, \gamma \text{ are fixed}, \ Var(\mathbf{u}) = \mathbf{I}/10, \ \ Var(\mathbf{e}) = 2\mathbf{I}.$$

The data are as follows:

	Levels of j			
hi subclasses	1	2	3	$y_{hi..}$
11	2	1	0	18
12	0	1	1	13
13	1	0	0	7
21	1	2	1	26
22	0	0	1	9
$y_{..j.}$	25	27	21	

We want to estimate using prior values of the squares and products of γ_{hi}. Suppose this is as follows, ordering i within h, and including γ_{23}.

$$\begin{pmatrix} .1 & -.05 & -.05 & -.1 & .05 & .05 \\ & .1 & -.05 & .05 & -.1 & .05 \\ & & .1 & .05 & .05 & -.1 \\ & & & .1 & -.05 & -.05 \\ & & & & .1 & -.05 \\ & & & & & .1 \end{pmatrix}.$$

The equations of the form

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}$$

are presented in (34).

$$\frac{1}{2}\begin{pmatrix} 6 & 0 & 3 & 2 & 1 & 3 & 2 & 1 & 0 & 0 & 0 & 3 & 2 & 1 \\ & 5 & 4 & 1 & 0 & 0 & 0 & 0 & 4 & 1 & 0 & 1 & 2 & 2 \\ & & 7 & 0 & 0 & 3 & 0 & 0 & 4 & 0 & 0 & 3 & 3 & 1 \\ & & & 3 & 0 & 0 & 2 & 0 & 0 & 1 & 0 & 0 & 1 & 2 \\ & & & & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ & & & & & 3 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ & & & & & & 2 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ & & & & & & & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ & & & & & & & & 4 & 0 & 0 & 1 & 2 & 1 \\ & & & & & & & & & 1 & 0 & 0 & 0 & 1 \\ & & & & & & & & & & 0 & 0 & 0 & 0 \\ & & & & & & & & & & & 4 & 0 & 0 \\ & & & & & & & & & & & & 4 & 0 \\ & & & & & & & & & & & & & 3 \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ \mathbf{c} \\ \boldsymbol{\gamma} \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} 38 \\ 35 \\ 44 \\ 22 \\ 7 \\ 18 \\ 13 \\ 7 \\ 26 \\ 9 \\ 0 \\ 25 \\ 27 \\ 21 \end{pmatrix} \frac{1}{2} \qquad (34)$$

Note that γ_{23} is included even though no observation on it exists.

Pre-multiplying these equations by

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix} \equiv \mathbf{T}$$

and adding $\mathbf{I}$ to the diagonals of equations (6)-(11) and $10\mathbf{I}$ to the diagonals of equations (12)-(14) we obtain the coefficient matrix to solve for the biased estimators and predictors. The right hand side vector is

$$(19, 17.5, 22, 11, 3.5, -.675, .225, .45, .675, -.225, -.45, 12.5, 13.5, 10.5)'.$$

This gives a solution of

$$\begin{aligned} \mathbf{r}^* &= (3.6899, 4.8607), \\ \mathbf{c}^* &= (1.9328, 3.3010, 3.3168), \\ \boldsymbol{\gamma}^* &= (.11406, -.11406, 0, -.11406, .11406, 0), \\ \mathbf{u}^* &= (-.00664, .04282, -.03618). \end{aligned}$$

Note that

$$\begin{aligned} \textstyle\sum_i \gamma^*_{ij} &= 0 \text{ for i } = 1,\ 2,\ \text{and} \\ \textstyle\sum_j \gamma^*_{ij} &= 0 \text{ for j } = 1,\ 2,\ 3. \end{aligned}$$

These are the same relationships that were defined for γ.

Post-multiplying the g-inverse of the coefficient matrix by $\mathbf{T}$ we get (35) ... (38) and the matrix for computing mean squared errors for $\mathbf{M}'(\mathbf{r}^*, \mathbf{c}^*, \boldsymbol{\gamma}^*, \mathbf{u}^*)$. The lower 9×9 submatrix is symmetric and invariant reflecting the fact that $\boldsymbol{\gamma}^*$, and $\mathbf{u}^*$ are invariant to the g-inverse taken.

Upper left 7×7

$$\begin{pmatrix}
.26181 & -.10042 & .02331 & 0 & .15599 & -.02368 & .00368 \\
-.05313 & .58747 & -.22911 & 0 & .54493 & .07756 & .00244 \\
-.05783 & -.26296 & .41930 & 0 & -.35232 & -.02259 & -.00741 \\
.56640 & .61368 & -.64753 & 0 & -1.02228 & .00080 & -.03080 \\
-.29989 & .13633 & .02243 & 0 & 2.07553 & .07567 & .04433 \\
-.02288 & .07836 & -.02339 & 0 & .07488 & .08341 & -.03341 \\
-.02712 & -.02836 & .02339 & 0 & .07512 & -.03341 & .08341
\end{pmatrix} \tag{35}$$

Upper right 7×7

$$\begin{pmatrix}
.02 & .02368 & -.00368 & -.02 & -.03780 & -.01750 & -.00469 \\
-.08 & -.07756 & -.00244 & .08 & -.01180 & -.01276 & -.03544 \\
.03 & .02259 & .00741 & -.03 & -.01986 & -.02631 & .00617 \\
.03 & -.00080 & .03080 & -.03 & .02588 & -.01608 & -.04980 \\
-.12 & -.07567 & -.04433 & .12 & -.05563 & .01213 & .00350 \\
-.05 & -.08341 & .03341 & .05 & -.00199 & .00317 & -.00118 \\
-.05 & .03341 & -.08341 & .05 & .00199 & -.00317 & .00118
\end{pmatrix} \tag{36}$$

Lower left 7×7

$$\begin{pmatrix}
.05 & -.05 & 0 & 0 & -.15 & -.05 & -.05 \\
.02288 & -.07836 & .02339 & 0 & -.07488 & -.08341 & .03341 \\
.02712 & .02836 & -.02339 & 0 & -.07512 & .03341 & -.08341 \\
-.05 & .05 & 0 & 0 & .15 & .05 & .05 \\
-.01192 & .01408 & -.04574 & 0 & -.08l51 & -.00199 & .00199 \\
-.03359 & -.02884 & -.01023 & 0 & .02821 & .00317 & -.00317 \\
-.05450 & -.08524 & .05597 & 0 & .05330 & -.00118 & .00118
\end{pmatrix} \tag{37}$$

Lower right 7×7

$$\begin{pmatrix}
.10 & .05 & .05 & -.10 & 0 & 0 & 0 \\
 & .08341 & -.03341 & -.05 & .00199 & -.00317 & .00118 \\
 & & .08341 & -.05 & -.00199 & .00317 & -.00118 \\
 & & & .10 & 0 & 0 & 0 \\
 & & & & .09343 & .00537 & .00120 \\
 & & & & & .09008 & .00455 \\
 & & & & & & .09425
\end{pmatrix} \tag{38}$$

A g-inverse of the coefficient matrix of equations like (8) is in (39) ... (41).

This gives a solution (−1.17081, 0, 6.79345, 8.16174, 8.17745, 0, 0, .76038, −1.52076, 0, 0, −.00664, .04282, −.03618). Premultiplying this solution by $\mathbf{T}$ we obtain for $\boldsymbol{\beta}_1^*$, (-1.17081, 0, 6.79345, 8.16174, 8.17745), and the same solution as before for $\boldsymbol{\beta}_2^*$ and $\mathbf{u}^*$; $\boldsymbol{\beta}_1$ is not estimable so $\boldsymbol{\beta}_1^*$ is not invariant and differs from the previous solution. But estimable functions of $\boldsymbol{\beta}_1$ are the same.

Pre and post-multiplying (39) ... (41) by $\mathbf{T}$ gives the matrix (42) ... (43). The lower 9×9 submatrix is the same as that of (38) associated with the fact that $\boldsymbol{\beta}_2^*$ and $\mathbf{u}^*$ are unique to whatever g-inverse is obtained.

Upper left 7×7

$$\begin{pmatrix} 1.00283 & 0 & -.43546 & -.68788 & -1.07683 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 0 & 0 \\ & & .51469 & .32450 & .51712 & 0 & 0 \\ & & & 1.20115 & .72380 & 0 & 0 \\ & & & & 3.34426 & 0 & 0 \\ & & & & & 0 & 0 \\ & & & & & & 0 \end{pmatrix} \qquad (39)$$

Upper right 7×7 and (lower left 7×7)'

$$\begin{pmatrix} .65838 & .68324 & 0 & 0 & -.026 & -.00474 & .03075 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.30020 & -.39960 & 0 & 0 & -.03166 & -.03907 & -.02927 \\ -.14427 & -.71147 & 0 & 0 & .01408 & -.02884 & -.08524 \\ -1.64509 & -.70981 & 0 & 0 & -.06743 & -.00063 & -.03l94 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad (40)$$

Lower right 7×7

$$\begin{pmatrix} 12.59603 & -5.19206 & 0 & 0 & -.01329 & .02112 & -.00784 \\ & 10.38413 & 0 & 0 & .02657 & -.04224 & .01567 \\ & & 0 & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 \\ & & & & .09343 & .00537 & .00120 \\ & & & & & .09008 & .00455 \\ & & & & & & .09425 \end{pmatrix} \qquad (41)$$

Upper left 7×7

$$\begin{pmatrix} 1.00283 & 0 & -.43546 & -.68788 & -1.07683 & -.10124 & .00124 \\ & 0 & 0 & 0 & 0 & 0 & 0 \\ & & .51469 & .32450 & .51712 & .05497 & -.00497 \\ & & & 1.20115 & .72380 & .07836 & -.02836 \\ & & & & 3.34426 & .15324 & .04676 \\ & & & & & .08341 & -.03341 \\ & & & & & & .08341 \end{pmatrix} \quad (42)$$

Upper right 7×7 and (lower left 7×7)'

$$\begin{pmatrix} .1 & .10124 & -.00124 & -.1 & -.026 & -.00474 & .03075 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.05 & -.05497 & .00497 & .05 & -.03166 & -.03907 & -.02927 \\ -.05 & -.07836 & .02836 & .05 & .01408 & -.02884 & -.08524 \\ -.2 & -.15324 & -.04676 & .2 & -.06743 & -.00063 & -.03194 \\ -.05 & -.08341 & .03341 & .05 & -.00199 & .00317 & -.00118 \\ -.05 & .03341 & -.08341 & .05 & .00199 & -.00317 & .00118 \end{pmatrix} \quad (43)$$

Lower right 7×7 is the same as in (38).

Suppose we wish to estimate $\mathbf{K}'(\boldsymbol{\beta}_1' \ \ \boldsymbol{\beta}_2')'$, which is estimable when the $r \times c$ subclasses are all filled, and

$$\mathbf{K}' = \begin{pmatrix} 6 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 0 & 0 & 0 \\ 0 & 6 & 2 & 2 & 2 & 0 & 0 & 0 & 2 & 2 & 2 \\ 3 & 3 & 6 & 0 & 0 & 3 & 0 & 0 & 3 & 0 & 0 \\ 3 & 3 & 0 & 6 & 0 & 0 & 3 & 0 & 0 & 3 & 0 \\ 3 & 3 & 0 & 0 & 6 & 0 & 0 & 3 & 0 & 0 & 3 \end{pmatrix} /6.$$

Pre-multiplying the upper 11x11 submatrix of either (35) to (38) or (42) to (43) by $\mathbf{K}'$ gives identical results shown in (44).

$$\begin{pmatrix} .44615 & .17671 & .15136 & .19665 & .58628 \\ & .91010 & .08541 & .38312 & 1.16170 \\ & & .32993 & .01354 & .01168 \\ & & & .76397 & .09215 \\ & & & & 2.51814 \end{pmatrix} \quad (44)$$

This represents the estimated mean squared error matrix of these 5 functions of $\boldsymbol{\beta}$.

Next we illustrate with another set of data the relationships of (3), (4), and (5) to (7). We have a design with 3 treatments and 2 random sires. The subclass numbers are

	Sires	
Treatments	1	2
1	2	1
2	1	2
3	2	0

The model is

$$y_{ijk} = \mu + t_i + s_j + x_{ijk}\beta + e_{ijk}.$$

where β is a regression and x_{ijk} the associated covariate.

$$\begin{aligned}\mathbf{y}' &= (5\ 3\ 6\ 4\ 7\ 5\ 4\ 8),\\ \text{Covariates} &= (1\ 2\ 1\ 3\ 2\ 4\ 2\ 3).\end{aligned}$$

The data are ordered sires in treatments. We shall use a prior on treatments of

$$\begin{pmatrix} 2 & -1 & -1 \\ & 2 & -1 \\ & & 2 \end{pmatrix}.$$

$$Var(\mathbf{e}) = 5\mathbf{I}, \text{ and } Var(\mathbf{s}) = \mathbf{I}.$$

We first illustrate the equations of (8),

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_1 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}_2 & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} \end{pmatrix} =$$

$$\begin{pmatrix} 1.6 & 3.6 & .6 & .6 & .4 & 1.0 & .6 \\ & 9.6 & .8 & 1.8 & 1.0 & 2.2 & 1.4 \\ & & .6 & 0 & 0 & .4 & .2 \\ & & & .6 & 0 & .2 & .4 \\ & & & & .4 & .4 & 0 \\ & & & & & 1.0 & 0 \\ & & & & & & .6 \end{pmatrix}. \tag{45}$$

and

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{X}_2'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} = (8.4\ \ 19.0\ \ 2.8\ \ 3.2\ \ 2.4\ \ 4.8\ \ 3.6)'. \tag{46}$$

These are ordered, μ, $\boldsymbol{\beta}$, $\mathbf{t}$, $\mathbf{s}$. Premultiplying (45) and (46) by

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 & 0 \\ & & 2 & -1 & -1 & 0 & 0 \\ & & & 2 & -1 & 0 & 0 \\ & & & & 2 & 0 & 0 \\ & & & & & 1 & 0 \\ & & & & & & 1 \end{pmatrix}$$

we get

$$\begin{pmatrix} 1.6 & 3.6 & .6 & .6 & .4 & 1.0 & .6 \\ 3.6 & 9.6 & .8 & 1.8 & 1.0 & 2.2 & 1.4 \\ .2 & -1.2 & 1.2 & -.6 & -.4 & .2 & 0 \\ .2 & 1.8 & -.6 & 1.2 & -.4 & -.4 & .6 \\ -.4 & -.6 & -.6 & -.6 & .8 & .2 & -.6 \\ 1.0 & 2.2 & .4 & .2 & .4 & 1.0 & 0 \\ .6 & 1.4 & .2 & .4 & 0 & 0 & .6 \end{pmatrix}, \tag{47}$$

and

$$(8.4 \;\; 19.0 \;\; 0 \;\; 1.2 \;\; -1.2 \;\; 4.8 \;\; 3.6)'. \tag{48}$$

The vector (48) is the right hand side of equations like (8). Then the coefficient matrix is matrix (47) + dg(0 0 1 1 1 1 1). The solution is

$$\begin{aligned} \mu^* &= 5.75832, \\ \beta^* &= -.16357, \\ (\mathbf{t}^*)' &= (-.49697 \;\; -.02234 \;\; .51931), \\ (\mathbf{s}^*)' &= (-.30146 \;\; .30146). \end{aligned}$$

Now we set up equations (3).

$$\mathbf{V} = (\mathbf{ZGZ}' + \mathbf{R}) = \begin{pmatrix} 6 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ & 6 & 0 & 1 & 0 & 0 & 1 & 1 \\ & & 6 & 0 & 1 & 1 & 0 & 0 \\ & & & 6 & 0 & 0 & 1 & 1 \\ & & & & 6 & 1 & 0 & 0 \\ & & & & & 6 & 0 & 0 \\ & & & & & & 6 & 1 \\ & & & & & & & 6 \end{pmatrix}. \tag{49}$$

$$\mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2' = \begin{pmatrix} 2 & 2 & 2 & -1 & -1 & -1 & -1 & -1 \\ & 2 & 2 & -1 & -1 & -1 & -1 & -1 \\ & & 2 & -1 & -1 & -1 & -1 & -1 \\ & & & 2 & 2 & 2 & -1 & -1 \\ & & & & 2 & 2 & -1 & -1 \\ & & & & & 2 & -1 & -1 \\ & & & & & & 2 & 2 \\ & & & & & & & 2 \end{pmatrix}. \tag{50}$$

$$(\mathbf{V} + \mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2')^{-1} =$$

$$\begin{pmatrix} .1525 & -.0475 & -.0275 & -.0090 & .0111 & .0111 & -.0005 & -.0005 \\ & .1525 & -.0275 & -.0090 & .0111 & .0111 & -.0005 & -.0005 \\ & & .1444 & .0214 & -.0067 & -.0067 & .0119 & .0119 \\ & & & .1417 & -.0280 & -.0280 & -.0031 & -.0031 \\ & & & & .1542 & -.0458 & .0093 & .0093 \\ & & & & & .1542 & .0093 & .0093 \\ & & & & & & .1482 & -.0518 \\ & & & & & & & .1482 \end{pmatrix}. \tag{51}$$

The equations like (4) are

$$\begin{pmatrix} .857878 & 1.939435 \\ 1.939435 & 5.328690 \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu}^* \\ \boldsymbol{\beta}^* \end{pmatrix} = \begin{pmatrix} 4.622698 \\ 10.296260 \end{pmatrix}. \tag{52}$$

The solution is (5.75832 - .163572) as in the mixed model equations.

$$(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta}_1^*) = \mathbf{y} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \\ 1 & 3 \\ 1 & 2 \\ 1 & 4 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 5.75832 \\ -.163572 \end{pmatrix} = \begin{pmatrix} -.59474 \\ -2.43117 \\ .40526 \\ -1.26760 \\ 1.56883 \\ -.10403 \\ -1.43117 \\ 2.73240 \end{pmatrix}.$$

$$\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2'(\mathbf{V} + \mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2')^{-1} =$$

$$\begin{pmatrix} .1426 & .1426 & .1471 & -.0725 & -.0681 & -.0681 & -.0899 & -.0899 \\ -.0501 & -.0501 & -.0973 & .1742 & .1270 & .1270 & -.0766 & -.0766 \\ -.0925 & -.0925 & -.0497 & -.1017 & -.0589 & -.0589 & .1665 & .1665 \end{pmatrix}.$$

Then $\mathbf{t}^* =$ (-.49697 -.02234 .51931)$'$ as before.

$$\mathbf{G}\mathbf{Z}'(\mathbf{V} + \mathbf{X}_2\boldsymbol{\beta}_2\boldsymbol{\beta}_2'\mathbf{X}_2')^{-1} =$$

$$\begin{pmatrix} .0949 & .0949 & -.0097 & .1174 & .0127 & .0127 & .0923 & .0923 \\ -.0053 & -.0053 & .1309 & -.0345 & .1017 & .1017 & .0304 & .0304 \end{pmatrix}.$$

Then $\mathbf{u}^* =$ (-.30146 .30146)$'$ as before.

Sections 9 and 10 of Chapter 15 give details concerning use of a diagonal matrix in place of $\mathbf{P}$.

10 Relationships Among Methods

BLUP, Bayesian estimation, and minimum mean squared error estimation are quite similar, and in fact are identical under certain assumptions.

10.1 Bayesian estimation

Let $(\mathbf{X} \;\; \mathbf{Z}) = \mathbf{W}$ and $(\boldsymbol{\beta}' \;\; \mathbf{u}') = \boldsymbol{\gamma}'$. Then the linear model is

$$\mathbf{y} = \mathbf{W}\boldsymbol{\gamma} + \mathbf{e}.$$

Let $\mathbf{e}$ have multivariate normal distribution with null means and $Var(\mathbf{e}) = \mathbf{R}$. Let the prior distribution of $\boldsymbol{\gamma}$ be multivariate normal with $E(\boldsymbol{\gamma}) = \boldsymbol{\mu}$, $Var(\boldsymbol{\gamma}) = \mathbf{C}$, and $Cov(\boldsymbol{\gamma}, \mathbf{e}') = \mathbf{0}$. Then for any of the common loss functions, that is, squared loss function, absolute loss function, or uniform loss function the Bayesian estimator of γ is the solution to (53).

$$(\mathbf{W}'\mathbf{R}^{-1}\mathbf{W} + \mathbf{C}^{-1})\;\hat{\boldsymbol{\gamma}} = \mathbf{W}'\mathbf{R}^{-1}\mathbf{y} + \mathbf{C}^{-1}\boldsymbol{\mu}. \tag{53}$$

Note that $\hat{\boldsymbol{\gamma}}$ is an unbiased estimator of $\boldsymbol{\gamma}$ if estimable and $E(\boldsymbol{\gamma}) = \boldsymbol{\mu}$. See Lindley and Smith (1972) for a discussion of Bayesian estimation for linear models. Equation (53) can be derived by maximizing $f(\mathbf{y}, \boldsymbol{\gamma})$ for variations in $\boldsymbol{\gamma}$. This might be called a MAP (maximum a posteriori) estimator, Melsa and Cohn (1978).

Now suppose that

$$\mathbf{C}^{-1} \rightarrow \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} \end{pmatrix}$$

and prior on $\boldsymbol{\mu} = \mathbf{0}$. Then (53) becomes the mixed model equations for BLUE and BLUP.

10.2 Minimum mean squared error estimation

Using the same notation as in Section 10.1, the minimum mean squared error estimator is

$$(\mathbf{W}'\mathbf{R}^{-1}\mathbf{W} + \mathbf{Q}^{-1})\gamma^o = \mathbf{W}'\mathbf{R}^{-1}\mathbf{y}, \tag{54}$$

where $\mathbf{Q} = \mathbf{C} + \boldsymbol{\mu}\boldsymbol{\mu}'$. Note that if $\boldsymbol{\mu} = \mathbf{0}$ this and the Bayesian estimator are identical. The essential difference is that the Bayesian estimator uses prior $E(\boldsymbol{\beta})$, whereas minimum MSE uses only squares and products of $\boldsymbol{\beta}$.

To convert (54) to the situation with prior on $\boldsymbol{\beta}_2$ but not on $\boldsymbol{\beta}_1$, let

$$\mathbf{Q}^- = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}^{-1} \end{pmatrix}.$$

The upper left partition is square with order equal to the number of elements in $\boldsymbol{\beta}_1$.

To convert (54) to the BLUP, mixed model equations let

$$\mathbf{Q}^- = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} \end{pmatrix},$$

where the upper left submatix is square with order p, the number of elements in $\boldsymbol{\beta}$. In the above results $\mathbf{P}$ may be singular. In that case use the technique described in previous sections for singular $\mathbf{G}$ and $\mathbf{P}$.

10.3 Invariance property of Bayesian estimator

Under normality and with absolute deviation as the loss function, the Bayesian estimator of f($\boldsymbol{\beta}$, $\mathbf{u}$) is $f(\boldsymbol{\beta}^o, \hat{\mathbf{u}})$, where $(\boldsymbol{\beta}^o, \hat{\mathbf{u}})$ is the Bayesian solution (also the BLUP solution when the priors are on $\mathbf{u}$ only), and f is any function. This was noted by Gianola (1982) who made use of a result reported by DeGroot (1981). Thus under normality any function of the BLUP solution is the Bayesian estimator of that function when the loss function is absolute deviation.

10.4 Maximum likelihood estimation

If the prior distribution on the parameters to be estimated is the uniform distribution and the mode of the posterior distribution is to be maximized, the resulting estimator is ML. When $\mathbf{Zu} + \mathbf{e} = \boldsymbol{\epsilon}$ has the multivariate normal distribution the MLE of $\boldsymbol{\beta}$, assumed estimable, is the maximizing value of $k \exp[-.5\ (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})]$. The maximizing value of this is the solution to

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{y},$$

the GLS equations. Now we know that the conditional mean of $\mathbf{u}$ given $\mathbf{y}$ is

$$\mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

Under fairly general conditions the ML estimator of a function of parameters is that same function of the ML estimators of those same parameters. Thus ML of the conditional mean of $\mathbf{u}$ under normality is

$$\mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o),$$

which we recognize as BLUP of $\mathbf{u}$ for any distribution.

11 Pattern Of Values Of P

When $\mathbf{P}$ has the structure described above and consequently is singular, a simpler method can be used. A diagonal, non-singular $\mathbf{P}$ can be written, which when used in mixed model equations results in the same estimates and predictions of estimable and predictable functions. See Chapter 15.

Chapter 10
Quadratic Estimation of Variances

C. R. Henderson

1984 - Guelph

Estimation of $\mathbf{G}$ and $\mathbf{R}$ is a crucial part of estimation and tests of significance of estimable functions of $\boldsymbol{\beta}$ and of prediction of $\mathbf{u}$. Estimators and predictors with known desirable properties exist when $\mathbf{G}$ and $\mathbf{R}$ are known, but realistically that is never the case. Consequently we need to have good estimates of them if we are to obtain estimators and predictors that approach BLUE and BLUP. This chapter is concerned with a particular class of estimators namely translation invariant, unbiased, quadratic estimators. First a model will be described that appears to include all linear models proposed for animal breeding problems.

1 A General Model For Variances And Covariances

The model with which we have been concerned is

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}. \\ Var(\mathbf{u}) &= \mathbf{G}, Var(\mathbf{e}) = \mathbf{R}, Cov(\mathbf{u},\ \mathbf{e}') = \mathbf{0}. \end{aligned}$$

The dimensions of vectors and matrices are

$\mathbf{y} : n \times 1,\ \mathbf{X} : n \times p,\ \boldsymbol{\beta} : p \times 1,\ \mathbf{Z} : n \times q,\ \mathbf{u} : q \times 1,\ \mathbf{e} : n \times 1,\ \mathbf{G} : q \times q,$ and $\mathbf{R} : n \times n.$

Now we characterize $\mathbf{u}$ and $\mathbf{e}$ in more detail. Let

$$\mathbf{Z}\mathbf{u} = \sum_{i=1}^{b} \mathbf{Z}_i \mathbf{u}_i. \tag{1}$$

$\mathbf{Z}_i$ has dimension $n \times q_i$, and $\mathbf{u}_i$ is $q_i \times 1$.

$$\begin{aligned} \sum_{i=1}^{b} q_i &= q, \\ Var(\mathbf{u}_i) &= \mathbf{G}_{ii} g_{ii}. && (2) \\ Cov(\mathbf{u}_i, \mathbf{u}_j') &= \mathbf{G}_{ij} g_{ij}. && (3) \end{aligned}$$

g_{ii} represents a variance and g_{ij} a covariance. Let

$$\begin{aligned} \mathbf{e}' &= (\mathbf{e}_1'\ \mathbf{e}_2' \ldots \mathbf{e}_c'). \\ Var(\mathbf{e}_i) &= \mathbf{R}_{ii} r_{ii}. && (4) \\ Cov(\mathbf{e}_i, \mathbf{e}_j') &= \mathbf{R}_{ij} r_{ij}. && (5) \end{aligned}$$

r_{ii} and r_{ij} represent variances and covariances respectively. With this model $Var(\mathbf{y})$ is

$$\mathbf{V} = \sum_{i=1}^{b} \sum_{j=1}^{b} \mathbf{Z}_i \mathbf{G}_{ij} \mathbf{Z}_j' \mathbf{g}_{ij} + \mathbf{R}, \tag{6}$$

$$Var(\mathbf{u}) = \mathbf{G} = \begin{pmatrix} \mathbf{G}_{11} g_{11} & \mathbf{G}_{12} g_{12} & \cdots & \mathbf{G}_{1b} g_{1b} \\ \mathbf{G}_{12}' g_{12} & \mathbf{G}_{22} g_{22} & \cdots & \mathbf{G}_{2b} g_{2b} \\ \vdots & \vdots & & \vdots \\ \mathbf{G}_{1b}' g_{1b} & \mathbf{G}_{2b}' g_{22} & \cdots & \mathbf{G}_{bb} g_{bb} \end{pmatrix}, \tag{7}$$

and

$$Var(\mathbf{e}) = \mathbf{R} = \begin{pmatrix} \mathbf{R}_{11} r_{11} & \mathbf{R}_{12} r_{12} & \cdots & \mathbf{R}_{1c} r_{1c} \\ \mathbf{R}_{12}' r_{12} & \mathbf{R}_{22} r_{22} & \cdots & \mathbf{R}_{2c} r_{2c} \\ \vdots & \vdots & & \vdots \\ \mathbf{R}_{1c}' r_{1c} & \mathbf{R}_{2c}' r_{2c} & \cdots & \mathbf{R}_{cc} r_{cc} \end{pmatrix}. \tag{8}$$

We illustrate this general model with two different specific models, first a traditional mixed model for variance components estimation, and second a two trait model with missing data. Suppose we have a random sire by fixed treatment model with interaction. The numbers of observations per subclass are

	Sires		
Treatment	1	2	3
1	2	1	2
2	1	3	0

Let the scalar model be

$$y_{ijk} = \mu + t_i + s_j + (ts)_{ij} + e_{ijk}.$$

The s_j have common variance, σ_s^2, and are uncorrelated. The $(ts)_{ij}$ have common variance, σ_{st}^2, and are uncorrelated. The s_j and $(ts)_{ij}$ are uncorrelated. The e_{ijk} have common variance, σ_e^2, and are uncorrelated. The corresponding vector model, for $b = 2$, is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \mathbf{e}.$$

$$\mathbf{X}\boldsymbol{\beta} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ t_1 \\ t_2 \end{pmatrix}, \ \mathbf{Z}_1\mathbf{u}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix},$$

$$\mathbf{Z}_2\mathbf{u}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ts_{11} \\ ts_{12} \\ ts_{13} \\ ts_{21} \\ ts_{22} \end{pmatrix},$$

and

$$\mathbf{G}_{11}g_{11} = \mathbf{I}_3\ \sigma^2_s,\ \ \mathbf{G}_{22}g_{22} = \mathbf{I}_5\ \sigma^2_{ts}.$$

$\mathbf{G}_{12}g_{12}$ does not exist, $c = 1$, and $\mathbf{R}_{11}r_{11} = \mathbf{I}_9\ \sigma^2_e$.

For a two trait model suppose that we have the following data on progeny of two related sires

Sire	Progeny	Trait 1	Trait 2
1	1	X	X
1	2	X	X
1	3	X	0
2	4	X	X
2	5	X	0

X represents a record and 0 represents a missing record. Let us assume an additive genetic sire model. Order the records by columns, that is animals within traits. Let $\mathbf{u}_1, \mathbf{u}_2$ represent sire values for traits 1 and 2 respectively. These are breeding values divided by 2. Let $\mathbf{e}_1, \mathbf{e}_2$ represent "errors" for traits 1 and 2 respectively. Sire 2 is a son of sire 1, both non-inbred.

$$n = 8,\ q_1 = 2,\ q_2 = 2.$$

$$\mathbf{Z}_1\mathbf{u}_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{u}_1,\ \ \mathbf{Z}_2\mathbf{u}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{u}_2,$$

$$\mathbf{G}_{11}g_{11} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} g^*_{11},\ \ \mathbf{G}_{12}g_{12} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} g^*_{12},$$

$$\mathbf{G}_{22} g_{22} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} g^*_{22},$$

where

$$\begin{pmatrix} g^*_{11} & g^*_{12} \\ g^*_{12} & g^*_{22} \end{pmatrix}$$

is the additive genetic variance-covariance matrix divided by 4. Also,

$$\mathbf{R}_{11} r_{11} = \mathbf{I}_5 r^*_{11},\ \mathbf{R}_{22} r_{22} = \mathbf{I}_3 r^*_{22},\ \mathbf{R}_{12} r_{12} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} r^*_{12},$$

where

$$\begin{pmatrix} r^*_{11} & r^*_{12} \\ r^*_{12} & r^*_{22} \end{pmatrix}$$

is the error variance-covariance matrix for the 2 traits. Then $h^2_1 = 4\ g^*_{11}/(g^*_{11} + r^*_{11})$. Genetic correlation between traits 1 and 2 is $g^*_{12}/(g^*_{11}\ \ g^*_{22})^{1/2}$.

Another method for writing **G** and **R** is the following

$$\mathbf{G} = \mathbf{G}^*_{11} g_{11} + \mathbf{G}^*_{12} g_{12}\ + ... +\ \mathbf{G}^*_{bb} g_{bb}, \tag{9}$$

where

$$\mathbf{G}^*_{11} = \begin{pmatrix} \mathbf{G}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},\ \mathbf{G}^*_{12} = \begin{pmatrix} \mathbf{0} & \mathbf{G}_{12} & \mathbf{0} \\ \mathbf{G}'_{12} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},\ ...,\ \mathbf{G}^*_{bb} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{bb} \end{pmatrix}.$$

Every $\mathbf{G}^*_{ij}$ has order, q, and

$$\mathbf{R} = \mathbf{R}^*_{11} r_{11} + \mathbf{R}^*_{12} r_{12}\ + ... +\ \mathbf{R}^*_{cc} r_{cc}, \tag{10}$$

where

$$\mathbf{R}^*_{11} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},\ \mathbf{R}^*_{12} = \begin{pmatrix} \mathbf{0} & \mathbf{R}_{12} & \mathbf{0} \\ \mathbf{R}'_{12} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},\ \text{etc.}$$

and every $\mathbf{R}^*_{ij}$ has order, n.

2 Quadratic Estimators

Many methods commonly used for estimation of variances and covariances are quadratic, unbiased, and translation invariant. They include among others, ANOVA estimators for balanced designs, unweighted means and weighted squares of means estimators for filled subclass designs, Henderson's methods 1, 2 and 3 for unequal numbers, MIVQUE, and MINQUE. Searle (1968, 1971a) describes in detail some of these methods.

A quadratic estimator is defined as $\mathbf{y}'\mathbf{Q}\mathbf{y}$ where for convenience $\mathbf{Q}$ can be specified as a symmetric matrix. If we derive a quadratic with a non-symmetric matrix, say $\mathbf{P}$, we can convert this to a quadratic with a symmetric matrix by the following identity.

$$\begin{aligned} \mathbf{y}'\mathbf{Q}\mathbf{y} &= (\mathbf{y}'\mathbf{P}\mathbf{y} + \mathbf{y}'\mathbf{P}'\mathbf{y})/2 \\ \text{where} \quad \mathbf{Q} &= (\mathbf{P} + \mathbf{P}')/2. \end{aligned}$$

A translation invariant quadratic estimator satisfies

$$\begin{aligned} \mathbf{y}'\mathbf{Q}\mathbf{y} &= (\mathbf{y} + \mathbf{X}\mathbf{k})'\mathbf{Q}(\mathbf{y} + \mathbf{X}\mathbf{k}) \text{ for any vector, } \mathbf{k}. \\ \mathbf{y}'\mathbf{Q}\mathbf{y} &= \mathbf{y}'\mathbf{Q}\mathbf{y} + 2\mathbf{y}'\mathbf{Q}\mathbf{X}\mathbf{k} + \mathbf{k}'\mathbf{X}'\mathbf{Q}\mathbf{X}\mathbf{k}. \end{aligned}$$

From this it is apparent that for equality it is required that

$$\mathbf{Q}\mathbf{X} = \mathbf{0}. \tag{11}$$

For unbiasedness we examine the expectation of $\mathbf{y}'\mathbf{Q}\mathbf{y}$ intended to estimate, say g_{gh}.

$$E(\mathbf{y}'\mathbf{Q}\mathbf{y}) = \boldsymbol{\beta}'\mathbf{X}'\mathbf{Q}\mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{b}\sum_{j=1}^{b} tr(\mathbf{Q}\mathbf{Z}_i\mathbf{G}^*_{ij}\mathbf{Z}'_j)g_{ij}$$

$$+ \sum_{i=1}^{c}\sum_{j=i}^{c} tr(\mathbf{Q}\mathbf{R}^*_{ij})r_{ij}.$$

We require that the expectation equals g_{gh}. Now if the estimator is translation invariant, the first term in the expectation is 0 because $\mathbf{Q}\mathbf{X} = \mathbf{0}$. Further requirements are that

$$\begin{aligned} tr(\mathbf{Q}\mathbf{Z}\mathbf{G}^*_{ij}\mathbf{Z}') &= 1 \text{ if } i = g \text{ and } j = h \\ &= 0, \text{ otherwise and} \\ tr(\mathbf{Q}\mathbf{R}^*_{ij}) &= 0 \text{ for all } i, j. \end{aligned}$$

3 Variances Of Estimators

Searle(1958) showed that the variance of a quadratic estimator $\mathbf{y}'\mathbf{Q}\mathbf{y}$, that is unbiased and translation invariant is

$$2\ tr(\mathbf{Q}\mathbf{V}\mathbf{Q}\mathbf{V}), \tag{12}$$

and the covariance between two estimators $\mathbf{y}'\mathbf{Q}_1\mathbf{y}$ and $\mathbf{y}'\mathbf{Q}_2\mathbf{y}$ is

$$2\ tr(\mathbf{Q}_1\mathbf{V}\mathbf{Q}_2\mathbf{V}) \tag{13}$$

where $\mathbf{y}$ is multivariate normal, and $\mathbf{V}$ is defined in (6). Then it is seen that (12) and (13) are quadratics in the g_{ij} and r_{ij}, the unknown parameters that are estimated. Consequently the results are in terms of these parameters, or they can be evaluated numerically for assumed values of $\mathbf{g}$ and $\mathbf{r}$. In the latter case it is well to evaluate $\mathbf{V}$ numerically for assumed $\mathbf{g}$ and $\mathbf{r}$ and then to proceed with the methods of (12) and (13).

4 Solutions Not In The Parameter Space

Unbiased estimators of variances and covariances with only one exception have positive probabilities of solutions not in the parameter space. The one exception is estimation of error variance from least squares or mixed model residuals. Otherwise estimates of variances can be negative, and functions of estimates of covariances and variances can result in estimated correlations outside the permitted range -1 to 1. In Chapter 12 the condition required for an estimated variance-covariance matrix to be in the parameter space is that there be no negative eigenvalues.

An inevitable price to pay for quadratic unbiasedness is non-zero probability that the estimated variance-covariance matrix will not fall in the parameter space. All such estimates are obtained by solving a set of linear equations obtained by equating a set of quadratics to their expectations. We could, if we knew how, impose side conditions on these equations that would force the solution into the parameter space. Having done this the solution would no longer yield unbiased estimators. What should be done in practice? It is sometimes suggested that we estimate unbiasedly, report all such results and then ultimately we can combine these into a better set of estimates that do fall in the parameter space. On the other hand, if the purpose of estimation is to provide $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ for immediate use in mixed model estimation and prediction, it would be very foolish to use estimates not in the parameter space. For example, suppose that in a sire evaluation situation we estimate σ_e^2/σ_s^2 to be negative and use this in mixed model equations. This would result in predicting a sire with a small number of progeny to be more different from zero than the adjusted progeny mean if $-\hat{\sigma}_e^2/\hat{\sigma}_s^2$ is less than the corresponding diagonal element of the sire. If the absolute value of this ratio is greater than the diagonal element, the sign of $\hat{s}_i$ is reversed as compared to the adjusted progeny mean. These consequences are of course contrary to selection index and BLUP principles.

Another problem in estimation should be recognized. The fact that estimated variance-covariance matrices fall in the parameter space does not necessarily imply that functions of these have that same property. For example, in an additive genetic sire model it is often assumed that $4\hat{\sigma}_s^2/(\hat{\sigma}_s^2 + \hat{\sigma}_e^2)$ is an estimate of h^2. But it is entirely possible that this computed function is greater than one even when $\hat{\sigma}_s^2$ and $\hat{\sigma}_e^2$ are both greater than 0. Of course if $\hat{\sigma}_s^2 < 0$ and $\hat{\sigma}_e^2 > 0$, the estimate of h^2 would be negative. Side conditions to solution of $\hat{\sigma}_s^2$ and $\hat{\sigma}_e^2$ that will insure that $\hat{\sigma}_s^2$, $\hat{\sigma}_e^2$, and h^2 (computed as above) fall in the parameter space are

$$\hat{\sigma}_s^2 > 0,\ \hat{\sigma}_e^2 > 0,\ \text{and}\ \hat{\sigma}_s^2/\hat{\sigma}_e^2 < 1/3.$$

Another point that should be made is that even though $\hat{\sigma}_s^2$ and $\hat{\sigma}_e^2$ are unbiased, $\hat{\sigma}_s^2/\hat{\sigma}_e^2$ is a biased estimator of σ_s^2/σ_e^2, and $4\hat{\sigma}_s^2/(\hat{\sigma}_s^2 + \hat{\sigma}_e^2)$ is a biased estimator of h^2.

5 Form Of Quadratics

Except for MIVQUE and MINQUE most quadratic estimators in models with all $g_{ij} = 0$ for $i \neq j$ and with $\mathbf{R} = \mathbf{I}\sigma_e^2$ can be expressed as linear functions of $\mathbf{y}'\mathbf{y}$ and of reductions in sums of squares that will now be defined.

Let OLS equations in $\boldsymbol{\beta}, \mathbf{u}$ be written as

$$\mathbf{W}'\mathbf{W}\boldsymbol{\alpha}^o = \mathbf{W}'\mathbf{y} \tag{14}$$

where $\mathbf{W} = (\mathbf{X}\ \ \mathbf{Z})$ and

$$\boldsymbol{\alpha}^o = \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix}.$$

Then reduction under the full model is

$$(\boldsymbol{\alpha}^o)'\mathbf{W}'\mathbf{y} \tag{15}$$

Partition with possible re-ordering of columns

$$\mathbf{W} = (\mathbf{W}_1\ \ \mathbf{W}_2) \tag{16}$$

and correspondingly

$$\boldsymbol{\alpha}^o = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix}.$$

$\boldsymbol{\alpha}_1$ should always contain $\boldsymbol{\beta}$ and from 0 to $b-1$ of the $\mathbf{u}_i$. Solve for $\boldsymbol{\alpha}_1^*$ in

$$\mathbf{W}_1'\mathbf{W}_1\boldsymbol{\alpha}_1^* = \mathbf{W}_1'\mathbf{y}. \tag{17}$$

Then reduction under the reduced model is

$$(\boldsymbol{\alpha}_1^*)'\mathbf{W}_1'\mathbf{y}. \tag{18}$$

6 Expectations of Quadratics

Let us derive the expectations of these ANOVA type quadratics.

$$E(\mathbf{y}'\mathbf{y}) = tr\ Var(\mathbf{y}) + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \tag{19}$$

$$= \sum_{i=1}^{b} tr(\mathbf{Z}_i\mathbf{G}_{ii}\mathbf{Z}_i')g_{ii} + n\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{20}$$

In traditional variance components models every $\mathbf{G}_{ii} = \mathbf{I}$. Then

$$E(\mathbf{y}'\mathbf{y}) = \sum_{i=1}^{b} \ n \ g_{ii} + n \ \sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{21}$$

It can be seen that (15) and (18) are both quadratics in $\mathbf{W}'\mathbf{y}$. Consequently we use $Var(\mathbf{W}'\mathbf{y})$ in deriving expectations. The random part of $\mathbf{W}'\mathbf{y}$ is

$$\sum_i \ \mathbf{W}'\mathbf{Z}_i\mathbf{u}_i + \mathbf{W}'\mathbf{e}. \tag{22}$$

The matrix of the quadratic in $\mathbf{W}'\mathbf{y}$ for the reduction under the full model is $(\mathbf{W}'\mathbf{W})^-$. Therefore the expectation is

$$\sum_{i=1}^{b} \ tr(\mathbf{W}'\mathbf{W})^-\mathbf{W}'\mathbf{Z}_i\mathbf{G}_{ii}\mathbf{Z}_l'\mathbf{W}g_{ii} \ + \ \text{rank } (\mathbf{W})\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{W}(\mathbf{W}'\mathbf{W})^-\mathbf{W}'\mathbf{X}\boldsymbol{\beta}. \tag{23}$$

When all $\mathbf{G}_{ii} \ = \ \mathbf{I}$, (23) reduces to

$$\sum_{i=1}^{b} \ n \ g_{ii} + r(\mathbf{W})\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{24}$$

For the reduction due to $\boldsymbol{\alpha}_1$, the matrix of the quadratic in $\mathbf{W}'\mathbf{y}$ is

$$\begin{pmatrix} (\mathbf{W}_1'\mathbf{W}_1)^- & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Then the expectation of the reduction is

$$\sum_{i=1}^{h} \ tr(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{Z}_i\mathbf{G}_{ii}\mathbf{Z}_i'\mathbf{W}_1 g_{ii} \ + \ \text{rank } (\mathbf{W}_1)\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{X}\boldsymbol{\beta}. \tag{25}$$

When all $\mathbf{G}_{ii} \ = \ \mathbf{I}$, (25) and when $\mathbf{X}$ is included in $\mathbf{W}_1$ simplifies to

$$\sum_i \ n \ g_{ii} + \sum_j \ tr(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{Z}_j\mathbf{Z}_j'\mathbf{W}_1 g_{jj} \ + \text{rank } (\mathbf{W}_1)\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{26}$$

where i refers to $\mathbf{u}_i$ included in $\boldsymbol{\alpha}_1$, and j refers to $\mathbf{u}_j$ not included in $\boldsymbol{\alpha}_1$. If $\mathbf{Z}_j$ is a linear function of $\mathbf{W}_1$, the coefficient of g_{jj} is n also.

7 Quadratics in $\hat{\mathbf{u}}$ and $\hat{\mathbf{e}}$

MIVQUE computations can be formulated as we shall see in Chapter 11 as quadratics in $\hat{\mathbf{u}}$ and $\hat{\mathbf{e}}$, BLUP of $\mathbf{u}$ and $\mathbf{e}$ when $\mathbf{g} = \tilde{\mathbf{g}}$ and $\mathbf{r} = \tilde{\mathbf{r}}$. The mixed model equations are

$$\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} + \tilde{\mathbf{G}}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \end{pmatrix}. \tag{27}$$

Let some quadratic in $\hat{\mathbf{u}}$ be $\hat{\mathbf{u}}'\mathbf{Q}\hat{\mathbf{u}}$. The expectation of this is

$$tr\mathbf{Q}\ Var(\hat{\mathbf{u}}). \tag{28}$$

To find $Var(\hat{\mathbf{u}})$, define a g-inverse of the coefficient matrix of (27) as

$$\begin{pmatrix} \mathbf{C}_{00} & \mathbf{C}_{01} \\ \mathbf{C}_{10} & \mathbf{C}_{11}. \end{pmatrix} \equiv \begin{pmatrix} \mathbf{C}_0 \\ \mathbf{C}_1 \end{pmatrix} \equiv \mathbf{C}. \tag{29}$$

$\hat{\mathbf{u}} = \mathbf{C}_1\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{y}$. See (16) for definition of $\mathbf{W}$. Then

$$Var(\hat{\mathbf{u}}) = \mathbf{C}_1\ [Var(\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{y})]\ \mathbf{C}_1', \tag{30}$$

and

$$Var(\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{y}) = \sum_{i=1}^{b}\sum_{j=1}^{b} \mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}_i\mathbf{G}_{ij}\mathbf{Z}_j'\tilde{\mathbf{R}}^{-1}\mathbf{W}g_{ij} \tag{31}$$

$$+\sum_{i=1}^{c}\sum_{j=1}^{c} \mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{R}_{ij}^*\tilde{\mathbf{R}}^{-1}\mathbf{W}r_{ij}. \tag{32}$$

Let some quadratic in $\hat{\mathbf{e}}$ be $\hat{\mathbf{e}}'\mathbf{Q}\hat{\mathbf{e}}$. The expectation of this is

$$tr\mathbf{Q}\ Var(\hat{\mathbf{e}}). \tag{33}$$

But $\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o - \mathbf{Z}\hat{\mathbf{u}} = \mathbf{y} - \mathbf{W}\boldsymbol{\alpha}^o$, where $(\boldsymbol{\alpha}^o)' = [(\boldsymbol{\beta}^o)'\ \ \hat{\mathbf{u}}']$ and $\mathbf{W} = (\mathbf{X}\ \ \mathbf{Z})$, giving

$$\hat{\mathbf{e}} = [\mathbf{I} - \mathbf{WCW}'\tilde{\mathbf{R}}^{-1}]\mathbf{y}. \tag{34}$$

Therefore,

$$Var(\hat{\mathbf{e}}) = (\mathbf{I} - \mathbf{WCW}'\tilde{\mathbf{R}}^{-1})\ [Var(\mathbf{y})]\ (\mathbf{I} - \mathbf{WCW}'\tilde{\mathbf{R}}^{-1})', \tag{35}$$

and

$$Var(\mathbf{y}) = \sum_{i=1}^{b}\sum_{j=1}^{b} \mathbf{Z}_i\mathbf{G}_{ij}\mathbf{Z}_j'\mathbf{g}_{ij} \tag{36}$$

$$+\sum_{i=1}^{c}\sum_{j=1}^{c} \mathbf{R}_{ij}^* r_{ij}. \tag{37}$$

When

$$\mathbf{G} = \tilde{\mathbf{G}},$$

$$\mathbf{R} = \tilde{\mathbf{R}},$$

$$Var(\hat{\mathbf{u}}) = \mathbf{G} - \mathbf{C}_{11}, \quad \text{and} \tag{38}$$

$$Var(\hat{\mathbf{e}}) = \mathbf{R} - \mathbf{WCW}'. \tag{39}$$

(38) and (39) are used for REML and ML methods to be described in Chapter 12.

8 Henderson's Method 1

We shall now present several methods that have been used extensively for estimation of variances (and in some cases with modifications for covariances). These are modelled after balanced ANOVA methods of estimation. The model for these methods is usually

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{b} \mathbf{Z}_i\mathbf{u}_i + \mathbf{e}, \tag{40}$$

where $Var(\mathbf{u}_i) = \mathbf{I}\sigma_i^2$, $Cov(\mathbf{u}_i, \mathbf{u}_j') = \mathbf{0}$ for all $i \neq j$, and $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$. However, it is relatively easy to modify these methods to deal with

$$Var(\mathbf{u}_i) = \mathbf{G}_{ii}\sigma_i^2.$$

For example, $\mathbf{G}_{ii}$ might be $\mathbf{A}$, the numerator relationship matrix.

Method 1, Henderson(1953), requires for unbiased estimation that $\mathbf{X}' = [1...1]$. The model is usually called a random model. The following reductions in sums of squares are computed

$$\mathbf{y}'\mathbf{Z}_i(\mathbf{Z}_i\mathbf{Z}_i)^{-1}\mathbf{Z}_i'\mathbf{y}(i = 1, \ldots, b), \tag{41}$$

$$(\mathbf{1}'\mathbf{y}\mathbf{y}'\mathbf{1})/n, \tag{42}$$

and

$$\mathbf{y}'\mathbf{y}. \tag{43}$$

The first b of these are simply uncorrected sums of squares for the various factors and interactions. The next one is the "correction factor", and the last is the uncorrected sum of squares of the individual observations.

Then these $b+2$ quadratics are equated to their expectations. The quadratics of (41) are easy to compute and their expectations are simple because $\mathbf{Z}_i'\mathbf{Z}_i$ is always diagonal. Advantage should therefore be taken of this fact. Also one should utilize the fact that the coefficient of σ_i^2 is n, as is the coefficient of any σ_j^2 for which $\mathbf{Z}_j$ is linearly dependent upon $\mathbf{Z}_i$. That is $\mathbf{Z}_j = \mathbf{Z}_i\mathbf{K}$. For example the reduction due to sires $\times$ herds has coefficient n for σ_{sh}^2, σ_s^2, σ_h^2 in a model with random sires and herds. The coefficient of σ_e^2 in the expectation is the rank of $\mathbf{Z}_i'\mathbf{Z}_i$, which is the number of elements in $\mathbf{u}_i$.

Because Method 1 is so easy, it is often tempting to use it on a model in which $\mathbf{X}' \neq (1...1)$, but to pretend that one or more fixed factors is random. This leads to biased estimators, but the bias can be evaluated in terms of unknown $\boldsymbol{\beta}\boldsymbol{\beta}'$. In balanced designs no bias results from using this method.

We illustrate Method 1 with a treatment $\times$ sire design in which treatments are regarded as random. The data are arranged as follows.

Number of Observations					
	Sires				
Treatment	1	2	3	4	Sums
1	8	3	2	5	18
2	7	4	1	0	12
3	6	2	0	1	9
Sums	21	9	3	6	39

Sums of Observations					
	Sires				
Treatment	1	2	3	4	Sums
1	54	21	13	25	113
2	55	33	8	0	96
3	44	17	0	9	70
Sums	153	71	21	34	279

$\mathbf{y}'\mathbf{y} = 2049$.

The ordinary least squares equations for these data are useful for envisioning Method 1 as well as some others. The coefficient matrix is in (44). The right hand side vector is (279, 113, 96, 70, 153, 71, 21, 34, 54, 21, 13, 25, 55, 33, 8, 44, 17, 9)$'$.

$$\begin{pmatrix}
39 & 18 & 12 & 9 & 21 & 9 & 3 & 6 & 8 & 3 & 2 & 5 & 7 & 4 & 1 & 6 & 2 & 1 \\
 & 18 & 0 & 0 & 8 & 3 & 2 & 5 & 8 & 3 & 2 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & 12 & 0 & 7 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 7 & 4 & 1 & 0 & 0 & 0 \\
 & & & 9 & 6 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 2 & 1 \\
 & & & & 21 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 7 & 0 & 0 & 6 & 0 & 0 \\
 & & & & & 9 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 4 & 0 & 0 & 2 & 0 \\
 & & & & & & 3 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 & & & & & & & 6 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 1 \\
 & & & & & & & & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & 7 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & 4 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & & 1 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & & & 6 & 0 & 0 \\
 & & & & & & & & & & & & & & & & 2 & 0 \\
 & & & & & & & & & & & & & & & & & 1
\end{pmatrix} \quad (44)$$

$$\begin{aligned}
\text{Red } (\mathbf{ts}) &= \frac{54^2}{8} + \frac{21^2}{3} \ldots + \frac{9^2}{1} = 2037.56. \\
\text{Red } (\mathbf{t}) &= \frac{113^2}{18} + \frac{96^2}{12} + \frac{70^2}{9} = 2021.83. \\
\text{Red } (\mathbf{s}) &= \frac{153^2}{21} + ... + \frac{34^2}{6} = 2014.49. \\
\text{C.F.} &= 279^2/39 = 1995.92. \\
E[\text{Red } (\mathbf{ts})] &= 10\sigma_s^2 + 39(\sigma_s^2 + \sigma_t^2 + \sigma_{ts}^2) + 39\ \mu^2.
\end{aligned}$$

For the expectations of other reductions as well as for the expectations of quadratics used in other methods including MIVQUE we need certain elements of $\mathbf{W}'\mathbf{Z}_1\mathbf{Z}_1'\mathbf{W}$, $\mathbf{W}'\mathbf{Z}_2\mathbf{Z}_2'\mathbf{W}$, and $\mathbf{W}'\mathbf{Z}_3\mathbf{Z}_3'\mathbf{W}$, where $\mathbf{Z}_1$, $\mathbf{Z}_2$, $\mathbf{Z}_3$ refer to incidence matrices for $\mathbf{t}$, $\mathbf{s}$, and $\mathbf{ts}$, respectively, and $\mathbf{W} = [\mathbf{1} \;\; \mathbf{Z}]$. The coefficients of $\mathbf{W}'\mathbf{Z}_1\mathbf{Z}_1'\mathbf{W}$ are in (45), (46), and (47).

Upper left 9×9

$$\begin{pmatrix} 549 & 324 & 144 & 81 & 282 & 120 & 48 & 99 & 144 \\ & 324 & 0 & 0 & 144 & 54 & 36 & 90 & 144 \\ & & 144 & 0 & 84 & 48 & 12 & 0 & 0 \\ & & & 81 & 54 & 18 & 0 & 9 & 0 \\ & & & & 149 & 64 & 23 & 46 & 64 \\ & & & & & 29 & 10 & 17 & 24 \\ & & & & & & 5 & 10 & 16 \\ & & & & & & & 26 & 40 \\ & & & & & & & & 64 \end{pmatrix} \tag{45}$$

Upper right 9×9 and (lower left 9×9)'

$$\begin{pmatrix} 54 & 36 & 90 & 84 & 48 & 12 & 54 & 18 & 9 \\ 54 & 36 & 90 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 84 & 48 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 54 & 18 & 9 \\ 24 & 16 & 40 & 49 & 28 & 7 & 36 & 12 & 6 \\ 9 & 6 & 15 & 28 & 16 & 4 & 12 & 4 & 2 \\ 6 & 4 & 10 & 7 & 4 & 1 & 0 & 0 & 0 \\ 15 & 10 & 25 & 0 & 0 & 0 & 6 & 2 & 1 \\ 24 & 16 & 40 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{46}$$

Lower right 9×9

$$\begin{pmatrix} 9 & 6 & 15 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 4 & 10 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 25 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 49 & 28 & 7 & 0 & 0 & 0 \\ & & & & 16 & 4 & 0 & 0 & 0 \\ & & & & & 1 & 0 & 0 & 0 \\ & & & & & & 36 & 12 & 6 \\ & & & & & & & 4 & 2 \\ & & & & & & & & 1 \end{pmatrix} \tag{47}$$

The coefficients of $\mathbf{W}'\mathbf{Z}_3\mathbf{Z}_3'\mathbf{W}$ are in (48), (49), and (50).

Upper left 9×9

$$\begin{pmatrix} 209 & 102 & 66 & 41 & 149 & 29 & 5 & 26 & 64 \\ & 102 & 0 & 0 & 64 & 9 & 4 & 25 & 64 \\ & & 66 & 0 & 49 & 16 & 1 & 0 & 0 \\ & & & 41 & 36 & 4 & 0 & 1 & 0 \\ & & & & 149 & 0 & 0 & 0 & 64 \\ & & & & & 29 & 0 & 0 & 0 \\ & & & & & & 5 & 0 & 0 \\ & & & & & & & 26 & 0 \\ & & & & & & & & 64 \end{pmatrix} \quad (48)$$

Upper right 9×9 and (lower left 9×9)'

$$\begin{pmatrix} 9 & 4 & 25 & 49 & 16 & 1 & 36 & 4 & 1 \\ 9 & 4 & 25 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 49 & 16 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 36 & 4 & 1 \\ 0 & 0 & 0 & 49 & 0 & 0 & 36 & 0 & 0 \\ 9 & 0 & 0 & 0 & 16 & 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 25 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (49)$$

Lower right 9×9

$$\text{dg } (9, 4, 25, 49, 16, 1, 36, 4, 1) \quad (50)$$

The coefficients of $\mathbf{W}'\mathbf{Z}_2\mathbf{Z}_2'\mathbf{W}$ are in (51), (52), and (53).

Upper left 9×9

$$\begin{pmatrix} 567 & 231 & 186 & 150 & 441 & 81 & 9 & 36 & 168 \\ & 102 & 70 & 59 & 168 & 27 & 6 & 30 & 64 \\ & & 66 & 50 & 147 & 36 & 3 & 0 & 56 \\ & & & 41 & 126 & 18 & 0 & 6 & 48 \\ & & & & 441 & 0 & 0 & 0 & 168 \\ & & & & & 81 & 0 & 0 & 0 \\ & & & & & & 9 & 0 & 0 \\ & & & & & & & 36 & 0 \\ & & & & & & & & 64 \end{pmatrix} \quad (51)$$

Upper right 9×9 and (lower left 9×9)'

$$\begin{pmatrix} 27 & 6 & 30 & 147 & 36 & 3 & 126 & 18 & 6 \\ 9 & 4 & 25 & 56 & 12 & 2 & 48 & 6 & 5 \\ 12 & 2 & 0 & 49 & 16 & 1 & 42 & 8 & 0 \\ 6 & 0 & 5 & 42 & 8 & 0 & 36 & 4 & 1 \\ 0 & 0 & 0 & 147 & 0 & 0 & 126 & 0 & 0 \\ 27 & 0 & 0 & 0 & 36 & 0 & 0 & 18 & 0 \\ 0 & 6 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 30 & 0 & 0 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 56 & 0 & 0 & 48 & 0 & 0 \end{pmatrix} \tag{52}$$

Lower right 9×9

$$\begin{pmatrix} 9 & 0 & 0 & 0 & 12 & 0 & 0 & 6 & 0 \\ & 4 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ & & 25 & 0 & 0 & 0 & 0 & 0 & 5 \\ & & & 49 & 0 & 0 & 42 & 0 & 0 \\ & & & & 16 & 0 & 0 & 8 & 0 \\ & & & & & 1 & 0 & 0 & 0 \\ & & & & & & 36 & 0 & 0 \\ & & & & & & & 4 & 0 \\ & & & & & & & & 1 \end{pmatrix} \tag{53}$$

$$\begin{aligned} E[\text{Red }(\mathbf{t})] &= 3\sigma_e^2 + 39\sigma_t^2 + k_1(\sigma_s^2 + \sigma_{ts}^2) + 39\ \mu^2. \\ k_1 &= \frac{102}{18} + \frac{66}{12} + \frac{41}{9} \ = \ 15.7222. \end{aligned}$$

The numerators above are the 2nd, 3rd, and 4th diagonals of (48) and (51). The denominators are the corresponding diagonals of the least squares coefficient matrix of (44). Also note that

$$\begin{aligned} 102 &= \sum_j n_{1j}^2 = 8^2 + 3^2 + 2^2 + 5^2, \\ 66 &= 7^2 + 4^2 + 1^2, \\ 41 &= 6^2 + 2^2 + 1^2. \\ E[\text{Red }(\mathbf{s})] &= 4\sigma_e^2 + 39\sigma_s^2 + k_2(\sigma_t^2 + \sigma_{ts}^2) + 39\ \mu^2. \\ k_2 &= \frac{149}{21} + \frac{29}{9} + \frac{5}{3} + \frac{26}{6} \ = \ 16.3175. \\ E(\text{ C.F.}) &= \sigma_e^2 + k_3\sigma_{ts}^2 + k_4\sigma_t^2 + k_5\sigma_s^2 + 39\ \mu^2. \\ k_3 &= \frac{209}{39} = 5.3590,\ k_4 = \frac{549}{39} = 14.0769,\ k_5 = \frac{567}{39} = 14.5385. \end{aligned}$$

It turns out that

$$\begin{aligned}\hat{\sigma}_e^2 &= [\mathbf{y}'\mathbf{y} - \text{Red }(\mathbf{ts})]/(39-10) \\ &= (2049 - 2037.56)/29 = .3945.\end{aligned}$$

$$E\begin{pmatrix}\hat{\sigma}_e^2 \\ \text{R}(\mathbf{ts}) \\ \text{R}(\mathbf{s}) \\ \text{R}(\mathbf{t}) \\ \text{CF}\end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 10 & 39 & 39 & 39 & 1 \\ 4 & 16.3175 & 39 & 16.3175 & 1 \\ 3 & 15.7222 & 15.7222 & 39 & 1 \\ 1 & 5.3590 & 14.5385 & 14.0769 & 1 \end{pmatrix}\begin{pmatrix}\sigma_e^2 \\ \sigma_{ts}^2 \\ \sigma_s^2 \\ \sigma_t^2 \\ 39\mu^2\end{pmatrix}.$$

$$\begin{pmatrix}\hat{\sigma}_e^2 \\ \hat{\sigma}_{ts}^2 \\ \hat{\sigma}_s^2 \\ \hat{\sigma}_t^2 \\ 39\hat{\mu}^2\end{pmatrix} =$$

$$\begin{pmatrix} 1. & 0 & 0 & 0 & 0 \\ -.31433 & .07302 & -.06979 & -.06675 & .06352 \\ .01361 & -.03006 & .06979 & .02379 & -.06352 \\ .04981 & -.02894 & .02571 & .06675 & -.06352 \\ -.21453 & .45306 & -1.00251 & -.92775 & 2.47720 \end{pmatrix}\begin{pmatrix} .3945 \\ 2037.56 \\ 2014.49 \\ 2021.83 \\ 1995.92 \end{pmatrix}$$

= [.3945, -.1088, .6660, 1.0216, 1972.05]'.

The 5×5 matrix just above is the inverse of the expectation matrix.

What if $\mathbf{t}$ is fixed but we estimate by Method 1 nevertheless? We can evaluate the bias in $\hat{\sigma}_{ts}^2$ and $\hat{\sigma}_s^2$ by noting that

$$\hat{\sigma}_{ts}^2 = \mathbf{y}'\mathbf{W}\mathbf{Q}_1\mathbf{W}'\mathbf{y} - .31433\ \hat{\sigma}_e^2$$

where $\mathbf{Q}_1$ is a matrix formed from these elements of the inverse just above, (.07302, -.06979, -.06675, 06352) and the matrices of quadratics in right hand sides representing Red (ts), Red (s), Red (t), C.F.

$\mathbf{Q}_1$ is dg [.0016, -.0037, -.0056, -.0074, -.0033, -0.0078, -.0233, -.0116, .0091, .0243, .0365, .0146, .0104, .0183, .0730, .0122, .0365, .0730]. dg refers to the diagonal elements of a matrix. Then the contribution of $\mathbf{tt}'$ to the expectation of $\hat{\sigma}_{ts}^2$ is

$$tr(\mathbf{Z}_1'\mathbf{W}\mathbf{Q}_1\mathbf{W}'\mathbf{Z}_1)\ (\mathbf{tt}')$$

where $\mathbf{Z}_1$ is the incidence matrix for $\mathbf{t}$ and $\mathbf{W} = (\mathbf{1}\ \ \mathbf{Z})$.

This turns out to be

$$tr \begin{pmatrix} -.0257 & .0261 & -.0004 \\ & -.0004 & -.0257 \\ & & .0261 \end{pmatrix} \mathbf{tt}',$$

that is, $-.0257\ t_1^2 + 2(.0261)\ t_1\ t_2\ - 2(.0004)\ t_1 t_3\ - .0004\ t_2^2\ - 2(.0257)\ t_2 t_3\ +\ .0261\ t_2^2$. This is the bias due to regarding $\mathbf{t}$ as random. Similarly the quadratic in right hand sides for estimation of σ_s^2 is

dg [-.0016, .0013, .0020, .0026, .0033, .0078, .0233, .0116, -.0038, -.0100, -.0150, -.0060, -.0043, -.0075, -.0301, -.0050, -.0150, -.0301].

The bias in $\hat{\sigma}_s^2$ is

$$tr \begin{pmatrix} .0257 & -.0261 & .0004 \\ & .0004 & .0257 \\ & & -.0261 \end{pmatrix} \mathbf{tt}'.$$

This is the negative of the bias in $\hat{\sigma}_{ts}^2$.

9 Henderson's Method 3

Method 3 of Henderson(1953) can be applied to any general mixed model for variance components. Usually the model assumed is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}\ +\ \sum\ \mathbf{Z}_i\mathbf{u}_i + \mathbf{e}. \tag{54}$$

$Var(\mathbf{u}_i) = \mathbf{I}\sigma_i^2$, $Cov(\mathbf{u}_i\mathbf{u}_j^{'}) = \mathbf{0}$, $Var(\mathbf{e}) = \mathbf{I}\sigma_i^2$. In this method $b+1$ different quadratics of the following form are computed.

$$\text{Red } (\boldsymbol{\beta} \text{ with from 0 to } b \text{ included } \mathbf{u}_i) \tag{55}$$

Then σ_e^2 is estimated usually by

$$\hat{\sigma}_e^2 = [\mathbf{y}'\mathbf{y} - \text{Red } (\boldsymbol{\beta}, \mathbf{u}_1, ..., \mathbf{u}_b)]/[n - \text{rank}(\mathbf{W})] \tag{56}$$

where $\mathbf{W} = (\mathbf{X}\ \ \mathbf{Z})$, and the solution to $\boldsymbol{\beta}^o$, $\mathbf{u}^o$ is OLS.

In some cases it is easier to compute σ_e^2 by expanding the model to include all possible interactions. Then if there is no covariate, $\hat{\sigma}_e^2$ is the within "smallest subclass" mean square. Then $\hat{\sigma}_e^2$ and the $b+1$ reductions are equated to their expectations. Method 3 has the unfortunate property that there are often more than $b+1$ reductions like (55) possible. Consequently more than one Method 3 estimator exists, and in unbalanced designs the estimates will not be invariant to the choice. One would like to select the set that will give smallest sampling variance, but this is unknown. Consequently it is tempting to select the easiest subset. This usually is

Red $(\boldsymbol{\beta}, \mathbf{u}_1)$, Red $(\boldsymbol{\beta}, \mathbf{u}_2)$, ..., Red $(\boldsymbol{\beta}, \mathbf{u}_b)$, Red $(\boldsymbol{\beta})$. For example: Red $(\boldsymbol{\beta}, \mathbf{u}_2)$ is computed as follows. Solve

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_2 \\ \mathbf{Z}_2'\mathbf{X} & \mathbf{Z}_2'\mathbf{Z}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}_2^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}_2{}'\mathbf{y} \end{pmatrix}.$$

Then reduction $= (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{y} + (\mathbf{u}_2^o)'\mathbf{Z}_2'\mathbf{y}$. To find the expectation of a reduction let a g-inverse of the coefficient matrix of the i^{th} reduction, $(\mathbf{W}_i'\mathbf{W}_i)$, be $\mathbf{C}_i$. Then

$$E(i^{th} \text{ reduction}) = \text{rank } (\mathbf{C}_i)\sigma_e^2 + \textstyle\sum_{j=1}^{s} tr\mathbf{C}_i\mathbf{W}_i'\mathbf{Z}_j\mathbf{Z}_j'\mathbf{W}_i\sigma_j^2 \; + \; \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{57}$$

$\mathbf{W}_i = [\mathbf{X} \;\; \mathbf{Z}_i$ for any included $\mathbf{u}_i]$, and $\mathbf{C}_i$ is the g-inverse. For example, in Red $(\boldsymbol{\beta}, \; \mathbf{u}_1, \; \mathbf{u}_3)$, $\mathbf{W} = [\mathbf{X} \;\; \mathbf{Z}_1 \;\; \mathbf{Z}_3]$.

Certain of the coefficients in (57) are n. These are all σ_j^2 included in the reduction and also any σ_k^2 for which

$$\mathbf{Z}_k \; = \; \mathbf{W}_j\mathbf{L}.$$

A serious computational problem with Method 3 is that it may be impossible with existing computers to find a g-inverse of some of the $\mathbf{W}_i'\mathbf{W}_i$. Partitioned matrix methods can sometimes be used to advantage. Partition

$$\mathbf{W}_i'\mathbf{W}_i = \begin{pmatrix} \mathbf{W}_1'\mathbf{W}_1 & \mathbf{W}_1'\mathbf{W}_2 \\ \mathbf{W}_2'\mathbf{W}_1 & \mathbf{W}_2'\mathbf{W}_2 \end{pmatrix},$$

and

$$\mathbf{W}_i'\mathbf{y} = \begin{pmatrix} \mathbf{W}_1'\mathbf{y} \\ \mathbf{W}_2'\mathbf{y} \end{pmatrix}.$$

It is advantageous to have $\mathbf{W}_1'\mathbf{W}_1$ be diagonal or at least of some form that is easy to invert. Define $\boldsymbol{\beta}$ and included $\mathbf{u}_i$ as $\boldsymbol{\alpha}$ and partition as $\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix}$. Then the equations to solve are

$$\begin{pmatrix} \mathbf{W}_1'\mathbf{W}_1 & \mathbf{W}_1'\mathbf{W}_2 \\ \mathbf{W}_2'\mathbf{W}_1 & \mathbf{W}_2'\mathbf{W}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{W}_1'\mathbf{y} \\ \mathbf{W}_2'\mathbf{y} \end{pmatrix}.$$

Absorb $\boldsymbol{\alpha}_1$ by writing equations

$$\mathbf{W}_2'\mathbf{P}\mathbf{W}_2\boldsymbol{\alpha}_2 = \mathbf{W}_2'\mathbf{P}\mathbf{y} \tag{58}$$

where $\mathbf{P} = \mathbf{I} - \mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'$. Solve for $\boldsymbol{\alpha}_2$ in (58). Then

$$\text{reduction} \; = \; \mathbf{y}'\mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{y} + \boldsymbol{\alpha}_2'\mathbf{W}_2'\mathbf{P}\mathbf{y}. \tag{59}$$

To find the coefficient of σ_j^2 in the expectation of this reduction, define

$$(\mathbf{W}_2'\mathbf{P}\mathbf{W}_2)^- \ = \ \mathbf{C}.$$

The coefficient of σ_j^2 is

$$tr(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{Z}_j\mathbf{Z}_j'\mathbf{W}_1 \ + \ tr\mathbf{C}\mathbf{W}_2'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{W}_2. \tag{60}$$

Of course if $\mathbf{u}_j$ is included in the reduction, the coefficient is n.

Let us illustrate Method 3 by the same example used in Method 1 except now we regard $\mathbf{t}$ as fixed. Consequently the σ_i^2 are σ_{ts}^2, σ_s^2, and we need 3 reductions, each including μ, $\mathbf{t}$. The only possible reductions are Red (μ,$\mathbf{t}$,$\mathbf{ts}$), Red (μ,$\mathbf{t}$,$\mathbf{s}$), and Red (μ,$\mathbf{t}$). Consequently in this special case Method 3 is unique. To find the first of these reductions we can simply take the last 10 rows and columns of the least squares equations. That is, dg [8, 3, 2, 5, 7, 4, 1, 6, 2, 1] $\hat{\mathbf{st}}$ = [54, 21, 13, 25, 55, 33, 8, 44, 17, 9]′. The resulting reduction is 2037.56 with expectation,

$$10\sigma_e^2 \ + \ 39(\sigma_{ts}^2 + \sigma_s^2) \ + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}.$$

For the reduction due to (μ, t, s) we can take the subset of OLS equations represented by rows (and columns) 2-7 inclusive. This gives equations to solve as follows.

$$\begin{pmatrix} 18 & 0 & 0 & 8 & 3 & 2 \\ & 12 & 0 & 7 & 4 & 1 \\ & & 9 & 6 & 2 & 0 \\ & & & 21 & 0 & 0 \\ & & & & 9 & 0 \\ & & & & & 3 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} 113 \\ 96 \\ 70 \\ 153 \\ 71 \\ 21 \end{pmatrix} \tag{61}$$

We can delete μ and s_4 because the above is a full rank subset of the coefficient matrix that includes μ and s_4. The inverse of the above matrix is

$$\begin{pmatrix} .1717 & .1602 & .1417 & -.1593 & -.1599 & -.1678 \\ & .3074 & .1989 & -.2203 & -.2342 & -.2093 \\ & & .2913 & -.2035 & -.2004 & -.1608 \\ & & & .2399 & .1963 & .1796 \\ & & & & .3131 & .1847 \\ & & & & & .5150 \end{pmatrix}, \tag{62}$$

and this gives a solution vector [5.448, 6.802, 6.760, 1.011, 1.547, 1.100]. The reduction is 2029.57. The coefficient of σ_s^2 in the expectation is 39 since $\mathbf{s}$ is included. To find the coefficient of σ_{ts}^2 define as $\mathbf{T}$ the submatrix of (51) formed by taking columns and rows (2-7). Then the coefficient of σ_{ts}^2 = trace [matrix (62)] $\mathbf{T}$ = 26.7638. The coefficient of σ_e^2 is 6. The reduction due to $\mathbf{t}$ and its expectation has already been done for Method 1.

Another way of formulating a reduction and corresponding expectations is to compute i^{th} reduction as follows. Solve

$$\begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_i'\mathbf{W}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \gamma_1^o \\ \gamma_2^o \\ \gamma_3^o \end{pmatrix} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{pmatrix} = \mathbf{r}. \tag{63}$$

$$\begin{aligned} \mathbf{r} &= \mathbf{W}'\mathbf{y}, \text{where } \mathbf{W} = (\mathbf{X} \;\; \mathbf{Z}) \\ \text{Red} &= \mathbf{r}'\mathbf{Q}_i\mathbf{r}, \end{aligned}$$

where $\mathbf{Q}_i$ is some g-inverse of the coefficient matrix, (63). Then the coefficient of σ_e^2 in the expectation is

$$\text{rank } (\mathbf{Q}_i) = \text{rank } (\mathbf{W}_i'\mathbf{W}_i). \tag{64}$$

Coefficient of σ_j^2 is

$$tr \; \mathbf{Q}_i\mathbf{W}'\mathbf{Z}_j\mathbf{Z}_j'\mathbf{W}. \tag{65}$$

Let the entire vector of expectations be

$$E \begin{pmatrix} \hat{\sigma}_e^2 \\ \text{Red } (1) \\ \vdots \\ \text{Red } (b+1) \end{pmatrix} = \mathbf{P} \begin{pmatrix} \sigma_e^2 \\ \sigma_1^2 \\ \vdots \\ \sigma_b^2 \\ \beta'\mathbf{X}'\mathbf{X}\beta \end{pmatrix}.$$

Then the unbiased estimators are

$$\begin{pmatrix} \hat{\sigma}_e^2 \\ \hat{\sigma}_1^2 \\ \vdots \\ \hat{\sigma}_b^2 \\ \widehat{\beta'\mathbf{X}'\mathbf{X}\beta} \end{pmatrix} = \mathbf{P}^{-1} \begin{pmatrix} \hat{\sigma}_e^2 \\ \text{Red } (1) \\ \vdots \\ \text{Red } (\text{b}+1) \end{pmatrix} \tag{66}$$

provided $\mathbf{P}^{-1}$ exists. If it does not, Method 3 estimators, at least with the chosen $b+1$ reductions, do not exist. In our example

$$E \begin{pmatrix} \hat{\sigma}_e^2 \\ \text{Red} \;\; (ts) \\ \text{Red} \;\; (ts) \\ \text{Red} \;\; (t) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 10 & 39 & 39 & 1 \\ 6 & 26.7638 & 39 & 1 \\ 3 & 15.7222 & 15.7222 & 1 \end{pmatrix} \begin{pmatrix} \sigma_e^2 \\ \sigma_{ts}^2 \\ \sigma_s^2 \\ \beta'\mathbf{X}'\mathbf{X}\beta \end{pmatrix}.$$

$$\begin{pmatrix} \hat{\sigma}_e^2 \\ \hat{\sigma}_{ts}^2 \\ \hat{\sigma}_s^2 \\ \widehat{\beta'\mathbf{X}'\mathbf{X}\beta} \end{pmatrix} = \begin{pmatrix} .3945 \\ .5240 \\ .0331 \\ 2011.89 \end{pmatrix},$$

$$= \begin{pmatrix} 1. & 0 & 0 & 0 \\ -.32690 & .08172 & -.08172 & 0 \\ .02618 & -.03877 & .08172 & -.04296 \\ 1.72791 & -.67542 & 0 & 1.67542 \end{pmatrix} \begin{pmatrix} .3945 \\ 2037.56 \\ 2029.57 \\ 2021.83 \end{pmatrix}.$$

These are different from the Method 1 estimates.

10 A Simple Method for General $\mathbf{X}\beta$

We now present a very simple method for the general $\mathbf{X}\boldsymbol{\beta}$ model provided an easy g-inverse of $\mathbf{X}'\mathbf{X}$ can be obtained. Write the following equations.

$$\begin{pmatrix} \mathbf{Z}_1'\mathbf{P}\mathbf{Z}_1 & \mathbf{Z}_1'\mathbf{P}\mathbf{Z}_2 & \cdots & \mathbf{Z}_1'\mathbf{P}\mathbf{Z}_b \\ \mathbf{Z}_2'\mathbf{P}\mathbf{Z}_1 & \mathbf{Z}_2'\mathbf{P}\mathbf{Z}_2 & \cdots & \mathbf{Z}_2'\mathbf{P}\mathbf{Z}_b \\ \vdots & & & \\ \mathbf{Z}_b'\mathbf{P}\mathbf{Z}_1 & \mathbf{Z}_b'\mathbf{P}\mathbf{Z}_2 & \cdots & \mathbf{Z}_b'\mathbf{P}\mathbf{Z}_b \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_b \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_1'\mathbf{P}\mathbf{y} \\ \mathbf{Z}_2'\mathbf{P}\mathbf{y} \\ \vdots \\ \mathbf{Z}_b'\mathbf{P}\mathbf{y} \end{pmatrix} \tag{67}$$

$\mathbf{P} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$. $\boldsymbol{\beta}^o$ is absorbed from the least squares equations to obtain (67). We could then compute b reductions from (67) and this would be Method 3. An easier method, however, is described next.

Let $\mathbf{D}_i$ be a diagonal matrix formed from the diagonals of $\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_i$. Then compute the following b quadratics,

$$\mathbf{y}'\mathbf{P}\mathbf{Z}_i\mathbf{D}_i^{-1}\mathbf{Z}_i'\mathbf{P}\mathbf{y}. \tag{68}$$

This computation is simple because $\mathbf{D}_i^{-1}$ is diagonal. It is simply the sum of squares of elements of $\mathbf{Z}_i'\mathbf{P}\mathbf{y}$ divided by the corresponding element of $\mathbf{D}_i$. The expectation is also easy. It is

$$q_i\sigma_e^2 \ + \ \textstyle\sum_{j=1}^{s} \ tr \ \mathbf{D}_i^{-1}\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\sigma_j^2. \tag{69}$$

Because $\mathbf{D}_i^{-1}$ is diagonal we need to compute only the diagonals of $\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i$ to find the last term of (69). Then as in Methods 1 and 3 we find some estimate of σ_e^2 and equate $\hat{\sigma}_e^2$ and the s quadratics of (68) to their expectations.

Let us illustrate the method with our same example, regarding $\mathbf{t}$ as fixed.

$$\mathbf{X}'\mathbf{X} \ = \begin{pmatrix} 39 & 18 & 12 & 9 \\ & 18 & 0 & 0 \\ & & 12 & 0 \\ & & & 9 \end{pmatrix},$$

and a g-inverse is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ & 18^{-1} & 0 & 0 \\ & & 12^{-1} & 0 \\ & & & 9^{-1} \end{pmatrix}.$$

The coefficient matrix of equations like (67) is in (70), (71) and (72) and the right hand side is (.1111, 4.6111, .4444, -5.1667, 3.7778, 2.1667, .4444, -6.3889, -1, 1, 0, -2.6667, 1.4444, 1.2222)′.

Upper left 7×7

$$\begin{pmatrix} 9.3611 & -5.0000 & -1.4722 & -2.8889 & 4.4444 & -1.3333 & -.8889 \\ & 6.7222 & -.6667 & -1.0556 & -1.3333 & 2.5 & -.3333 \\ & & 2.6944 & -.5556 & -.8889 & -.3333 & 1.7778 \\ & & & 4.5 & -2.2222 & -.8333 & -.5556 \\ & & & & 4.4444 & -1.3333 & -.8889 \\ & & & & & 2.5 & -.3333 \\ & & & & & & 1.7778 \end{pmatrix} \quad (70)$$

Upper right 7×7 and (lower left 7×7)'

$$\begin{pmatrix} -2.2222 & 2.9167 & -2.3333 & -.5833 & 2.0 & -1.3333 & -.6667 \\ -.8333 & -2.3333 & 2.6667 & -.3333 & -1.3333 & 1.5556 & -.2222 \\ -.5556 & -.5833 & -.3333 & .9167 & 0 & 0 & 0 \\ 3.6111 & 0 & 0 & 0 & -.6667 & -.2222 & .8889 \\ -2.2222 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.8333 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.5556 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (71)$$

Lower right 7×7

$$\begin{pmatrix} 3.6111 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 2.9167 & -2.3333 & -.5833 & 0 & 0 & 0 \\ & & 2.6667 & -.3333 & 0 & 0 & 0 \\ & & & .9167 & 0 & 0 & 0 \\ & & & & 2.0 & -1.3333 & -.6667 \\ & & & & & 1.5556 & -.2222 \\ & & & & & & .8889 \end{pmatrix} \quad (72)$$

The diagonals of the variance of the reduced right hand sides are needed in this method and other elements are needed for approximate MIVQUE in Chapter 11. The coefficients of σ_e^2 in this variance are in (70), ..., (72). The coefficients of σ_{ts}^2 are in (73), (74) and (75). These are computed by (Cols. 5-14 of 10.70) (same)$'$.

Upper left 7×7

$$\begin{pmatrix} 47.77 & -24.54 & -5.31 & -17.93 & 27.26 & -7.11 & -3.85 \\ & 25.75 & .39 & -1.60 & -7.11 & 8.83 & .22 \\ & & 5.66 & -.74 & -3.85 & .22 & 4.37 \\ & & & 20.27 & -16.30 & -1.94 & -.74 \\ & & & & 27.26 & -7.11 & -3.85 \\ & & & & & 8.83 & .22 \\ & & & & & & 4.37 \end{pmatrix} \quad (73)$$

Upper right 7×7 and (lower left 7×7)'

$$\begin{pmatrix} -16.30 & 14.29 & -12.83 & -1.46 & 6.22 & -4.59 & -1.63 \\ -1.94 & -12.83 & 12.67 & .17 & -4.59 & 4.25 & .35 \\ -.74 & -1.46 & .17 & 1.29 & 0 & 0 & 0 \\ 18.98 & 0 & 0 & 0 & -1.63 & .35 & 1.28 \\ -16.30 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1.94 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.74 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (74)$$

Lower right 7×7

$$\begin{pmatrix} 18.98 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 14.29 & -12.83 & -1.46 & 0 & 0 & 0 \\ & & 12.67 & .17 & 0 & 0 & 0 \\ & & & 1.29 & 0 & 0 & 0 \\ & & & & 6.22 & -4.59 & -1.63 \\ & & & & & 4.25 & .35 \\ & & & & & & 1.28 \end{pmatrix} \quad (75)$$

The coefficients of σ_s^2 are in (76), (77), and (78). These are computed by (Cols 1-4 of 10.70) (same)$'$.

Upper left 7×7

$$\begin{pmatrix} 123.14 & -76.39 & -12.81 & -33.95 & 56.00 & -22.08 & -7.67 \\ & 71.75 & 1.67 & 2.97 & -28.25 & 24.57 & 1.60 \\ & & 10.18 & .96 & -6.81 & -.14 & 6.63 \\ & & & 30.02 & -20.94 & -2.35 & -.57 \\ & & & & 27.26 & -7.11 & -3.85 \\ & & & & & 8.83 & .22 \\ & & & & & & 4.37 \end{pmatrix} \quad (76)$$

Upper right 7×7 and (lower left 7×7)'

$$\begin{pmatrix} -26.25 & 39.83 & -34.69 & -5.14 & 27.31 & -19.62 & -7.70 \\ 2.07 & -29.88 & 29.81 & .06 & -18.26 & 17.36 & .90 \\ .32 & -4.31 & .76 & 3.55 & -1.69 & 1.05 & .64 \\ 23.86 & -5.64 & 4.11 & 1.53 & -7.37 & 1.21 & 6.16 \\ -16.30 & 16.59 & -13.63 & -2.96 & 12.15 & -7.51 & -4.64 \\ -1.94 & -9.53 & 9.89 & -.36 & -5.44 & 5.85 & -.41 \\ -.74 & -2.85 & .59 & 2.26 & -.96 & .79 & .17 \end{pmatrix} \quad (77)$$

Lower right 7×7

$$\begin{pmatrix} 18.98 & -4.21 & 3.15 & 1.06 & -5.74 & .86 & 4.88 \\ & 14.29 & -12.83 & -1.46 & 8.94 & -7.52 & -1.43 \\ & & 12.67 & .17 & -8.22 & 7.26 & .96 \\ & & & 1.29 & -.72 & .26 & .46 \\ & & & & 6.22 & -4.59 & -1.63 \\ & & & & & 4.25 & .35 \\ & & & & & & 1.28 \end{pmatrix} \tag{78}$$

The reduction for **ts** is

$$\frac{3.778^2}{4.444} + \cdots + \frac{1.222^2}{.889} = 23.799.$$

The expectation is $10\ \sigma_e^2 + 35.7262\ (\sigma_{ts}^2 + \sigma_s^2)$, where 10 is the number of elements in the **ts** vector and

$$35.7262 = \frac{27.259}{4.444} + \cdots + \frac{1.284}{.889}.$$

The reduction for **s** is

$$\frac{.111^2}{9.361} + \cdots + \frac{(-5.167)^2}{4.5} = 9.170.$$

The expectation is $4\ \sigma_e^2 + 15.5383\ \sigma_{ts}^2 + 34.2770\ \sigma_s^2$, where

$$15.5383 = \frac{47.773}{9.361} + \cdots + \frac{20.265}{4.5}.$$
$$34.2770 = \frac{123.144}{9.361} + \cdots + \frac{30.019}{4.5}.$$

Thus

$$E \begin{pmatrix} \hat{\sigma}_e^2 \\ \text{Red}(\mathbf{ts}) \\ \text{Red}(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 10 & 35.7262 & 35.7262 \\ 4 & 15.5383 & 34.2770 \end{pmatrix} \begin{pmatrix} \sigma_e^2 \\ \sigma_{ts}^2 \\ \sigma_s^2 \end{pmatrix}.$$

Then

$$\begin{pmatrix} \hat{\sigma}_e^2 \\ \hat{\sigma}_{ts}^2 \\ \hat{\sigma}_s^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -.29855 & .05120 & -.05337 \\ .01864 & -.02321 & .05337 \end{pmatrix} \begin{pmatrix} .3945 \\ 23.799 \\ 9.170 \end{pmatrix} = \begin{pmatrix} .3945 \\ .6114 \\ -.0557 \end{pmatrix}.$$

11 Henderson's Method 2

Henderson's Method 2 (1953) is probably of interest from an historical viewpoint only. It has the disadvantage that random by fixed interactions and random within fixed nesting are not permitted. It is a relatively easy method, but usually no easier than the method described in Sect. 10.10, absorption of $\boldsymbol{\beta}$, and little if any easier than an approximate MIVQUE procedure described in Chapter 11.

Method 2 involves correction of the data by a least squares solution to $\boldsymbol{\beta}$ excluding μ. Then a Method 1 analysis is carried out under the assumption of a model

$$\mathbf{y} \ = \ \mathbf{1}\alpha \ + \ \sum_i \mathbf{Z}_i \mathbf{u}_i + \mathbf{e}.$$

If the solution to $\boldsymbol{\beta}^o$ is done as described below, the expectations of the Method 1 reductions are identical to those for a truly random model except for an increase in the coefficients of σ_e^2 . Partition

$$\mathbf{Z} \ = \ [\mathbf{Z}_a \ \ \mathbf{Z}_b]$$

such that rank

$$(\mathbf{Z}_a) \ = \ \text{rank } (\mathbf{Z}).$$

Then partition

$$\mathbf{X} \ = \ (\mathbf{X}_a \ \ \mathbf{X}_b)$$

such that

$$\text{rank } (\mathbf{X}_a \ \ \mathbf{Z}_a) \ = \ \text{rank } (\mathbf{X} \ \ \mathbf{Z}).$$

See Henderson, Searle, and Schaeffer (1974). Solve equations (79) for $\boldsymbol{\beta}_a$.

$$\begin{pmatrix} \mathbf{X}_a'\mathbf{X}_a & \mathbf{X}_a'\mathbf{Z}_a \\ \mathbf{Z}_a'\mathbf{X}_a & \mathbf{Z}_a'\mathbf{Z}_a \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_a \\ \mathbf{u}_a \end{pmatrix} = \begin{pmatrix} \mathbf{X}_a'\mathbf{y} \\ \mathbf{Z}_a'\mathbf{y} \end{pmatrix} \tag{79}$$

Let the upper submatrix (pertaining to $\boldsymbol{\beta}_a$) of the inverse of the matrix of (79) be denoted by $\mathbf{P}$. This can be computed as

$$\mathbf{P} = [\mathbf{X}_a'\mathbf{X}_a - \mathbf{X}_a'\mathbf{Z}_a \ (\mathbf{Z}_a'\mathbf{Z}_a)^{-1}\mathbf{Z}_a'\mathbf{X}_a]^{-1}. \tag{80}$$

Now compute

$$\mathbf{1}'\mathbf{y}^* \ = \ \mathbf{1}'\mathbf{y} - \mathbf{1}'\mathrm{X}_a\boldsymbol{\beta}_a. \tag{81}$$

$$\mathbf{Z}_i'\mathbf{y}^* \ = \ \mathbf{Z}_i'\mathbf{y} - \mathbf{Z}_i'\mathrm{X}_a\boldsymbol{\beta}_a \ i = 1, \ldots, b. \tag{82}$$

Then compute the following quadratics

$$(\mathbf{1}'\mathbf{y}^*)^2/n, \text{ and} \tag{83}$$

$$(\mathbf{Z}_i'\mathbf{y}^*)'(\mathbf{Z}_i'\mathbf{Z}_i)^{-1}(\mathbf{Z}_i'\mathbf{y}^*) \text{ for } i = 1, \ldots, b. \tag{84}$$

The expectations of these quadratics are identical to those with $\mathbf{y}$ in place of $\mathbf{y}^*$ except for an increase in the coefficient of σ_e^2 computed as follows. Increase in coefficient of σ_e^2 in expectation of (83) by

$$tr\mathbf{P}(\mathbf{X}_a'\mathbf{1}\mathbf{1}'\mathbf{X}_a)/n. \tag{85}$$

Increase in the coefficient of σ_e^2 in expectation of (84) is

$$tr\mathbf{P}(\mathbf{X}_a'\mathbf{Z}_i(\mathbf{Z}_i'\mathbf{Z}_i)^{-1}\mathbf{Z}_i'\mathbf{X}_a). \tag{86}$$

Note that $\mathbf{X}_a'\mathbf{Z}_i(\mathbf{Z}_i'\mathbf{Z}_i)^{-1}\mathbf{Z}_i'\mathbf{X}_a$ is the quantity that would be subtracted from $\mathbf{X}_a'\mathbf{X}_a$ if we were to "absorb" $\mathbf{u}_i$. σ_e^2 can be estimated in a number of ways but usually by the conventional residual

$$[\mathbf{y}'\mathbf{y} - (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{y} - (\mathbf{u}^o)'\mathbf{Z}'\mathbf{y}]/[n - \text{rank }(\mathbf{X} \quad \mathbf{Z})].$$

Sampling variances for Method 2 can be computed by the same procedure as for Method l except that the variance of adjusted right hand sides of μ and $\mathbf{u}$ equations is increased by $\begin{pmatrix} \mathbf{1}'\mathbf{X}_a \\ \mathbf{Z}'\mathbf{X}_a \end{pmatrix} \mathbf{P}(\mathbf{X}_a'\mathbf{1} \quad \mathbf{X}_a'\mathbf{Z})\ \sigma_e^2$ over the unadjusted. As is true for other quadratic estimators, quadratics in the adjusted right hand sides are uncorrelated with σ_e^2, the OLS residual mean square.

We illustrate Method 2 with our same data, but now we assume that σ_{ts}^2 does not exist. This 2 way mixed model could be done just as easily by Method 3 as by Method 2, but it suffices to illustrate the latter. Delete μ and t_3 and include all 4 levels of $\mathbf{s}$. First solve for β_a in these equations.

$$\begin{pmatrix} 18 & 0 & 8 & 3 & 2 & 5 \\ & 12 & 7 & 4 & 1 & 0 \\ & & 21 & 0 & 0 & 0 \\ & & & 9 & 0 & 0 \\ & & & & 3 & 0 \\ & & & & & 6 \end{pmatrix} \begin{pmatrix} \beta_a \\ u_a \end{pmatrix} = \begin{pmatrix} 113 \\ 96 \\ 153 \\ 71 \\ 21 \\ 34 \end{pmatrix}$$

The solution is $\boldsymbol{\beta}_a$ = [-1.31154, .04287]$'$, $\mathbf{u}_a$ = (7.7106, 8.30702, 7.86007, 6.75962)$'$. The adjusted right hand sides are

$$\begin{pmatrix} 279 \\ 153 \\ 71 \\ 21 \\ 34 \end{pmatrix} - \begin{pmatrix} 18 & 12 \\ 8 & 7 \\ 3 & 4 \\ 2 & 1 \\ 5 & 0 \end{pmatrix} \begin{pmatrix} -1.31154 \\ .04287 \end{pmatrix} = \begin{pmatrix} 302.093 \\ 163.192 \\ 74.763 \\ 23.580 \\ 40.558 \end{pmatrix}.$$

Then the sum of squares of adjusted right hand sides for sires is

$$\frac{(163.192)^2}{12} + \cdots + \frac{(40.558)^2}{6} = 2348.732.$$

The adjusted C.F. is $(302.093)^2/39 = 2340.0095$. $\mathbf{P}$ is the upper 2x2 of the inverse of the coefficient matrix (79) is

$$\mathbf{P} = \begin{pmatrix} .179532 & .110888 \\ & .200842 \end{pmatrix}.$$

$$\mathbf{X}_a'\mathbf{Z}_i(\mathbf{Z}_i'\mathbf{Z}_i)^{-1}\mathbf{Z}_i'\mathbf{X}_a = \begin{pmatrix} 8 & 3 & 2 & 5 \\ 7 & 4 & 1 & 0 \end{pmatrix} \begin{pmatrix} 21 & 0 & 0 & 0 \\ & 9 & 0 & 0 \\ & & 3 & 0 \\ & & & 6 \end{pmatrix}^{-1} \begin{pmatrix} 8 & 7 \\ 3 & 4 \\ 2 & 1 \\ 5 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 9.547619 & 4.666667 \\ & 4.444444 \end{pmatrix}.$$

The trace of this (86) is 3.642 to be added to the coefficient of σ_e^2 in E (sires S.S). The trace of $\mathbf{P}$ times the following matrix

$$\begin{pmatrix} 18 \\ 12 \end{pmatrix} \begin{pmatrix} 39 \end{pmatrix}^{-1} (18 \;\; 12) = \begin{pmatrix} 8.307692 & 5.538462 \\ & 3.692308 \end{pmatrix}.$$

gives 3.461 to be added to the coefficient of σ_e^2 in E(CF). Then

$$E(Sire\ SS) = 7.642\ \sigma_e^2 + 39\ \sigma_s^2 + \text{a quadratic.}$$
$$E(C.F.) = 4.461\ \sigma_e^2 + 14.538\ \sigma_s^2 + \text{the same quadratic.}$$

Then taking some estimate of σ_e^2 one equates these expectations to the computed sums of squares.

12 An Unweighted Means ANOVA

A simple method for testing hypotheses approximately is the unweighted means analysis described in Yates (1934). This method is appropriate for the mixed model described in Section 4 provided that every subclass is filled and there are no covariates. The "smallest" subclass means are taken as the observations as in Section 6 in Chapter 1. Then a conventional analysis of variance for equal subclass numbers (in this case 1) is performed. The expectations of these mean squares, except for the coefficients of σ_e^2 are exactly the same as they would be had there actually been only one observation per subclass. An algorithm for finding such expectations is given in Henderson (1959).

The coefficient of σ_e^2 is the same in every mean square. To compute this let s = the number of "smallest" subclasses, and let n_i be the number of observations in the i^{th} subclass. Then the coefficient of σ_e^2 is

$$\textstyle\sum_{i=1}^{s} \quad n_i^{-1}/s. \tag{87}$$

Estimate σ_e^2 by

$$[y'y - \sum_{i=1}^{s} \; y_i^2/n_i]/(n-s), \tag{88}$$

where y_i is the sum of observations in the i^{th} subclass. Henderson (1978a) described a simple algorithm for computing sampling variances for the unweighted means method.

We illustrate estimation by a two way mixed model,

$y_{ijk} \; = \; a_i \; + \; b_j \; + \; \gamma_{ij} \; + \; e_{ijk}.$

b_j is fixed.

$Var(\mathbf{a}) \; = \; \mathbf{I}\sigma_a^2, \; Var(\boldsymbol{\gamma}) \; = \; \mathbf{I}\sigma_\gamma^2, \; Var(\mathbf{e}) \; = \; \mathbf{I}\sigma_e^2.$

Let the data be

	n_{ij} B			$\bar{y}_{ij.}$ B		
A	1	2	3	1	2	3
1	5	4	1	8	10	5
2	2	10	5	7	8	4
3	1	4	2	6	9	3
4	2	1	5	10	12	8

The mean squares and their expectation in the unweighted means analysis are

	df	MS	E(ms)
A	3	9.8889	$.475 \; \sigma_e^2 \; + \; \sigma_\gamma^2 \; + \; 3\sigma_a^2$
B	2	22.75	$.475 \; \sigma_e^2 \; + \; \sigma_\gamma^2 \; + \; Q(b)$
AB	6	.3056	$.475 \; \sigma_e^2 \; + \; \sigma_\gamma^2$

Suppose $\hat{\sigma}_e^2$ estimated as described above is .2132. Then

$$\hat{\sigma}_\gamma^2 \; = \; .3056 - .475(.2132) = .2043,$$

and

$$\hat{\sigma}_a^2 \; = \; (9.8889 - .3056)/3 = 3.1944.$$

The coefficient of σ_e^2 is $(5^{-1} + 4^{-1} + ... + 5^{-1})/12 = .475$.

13 Mean Squares For Testing $\mathbf{K}'\mathbf{u}$

Section 2.c in Chapter 4 described a general method for testing the hypothesis, $\mathbf{K}'\boldsymbol{\beta} = \mathbf{0}$ against the unrestricted hypothesis. The mean square for this test is

$$(\boldsymbol{\beta}^o)'\mathbf{K}(\mathbf{K}'\mathbf{C}\mathbf{K})^{-1}\mathbf{K}'\boldsymbol{\beta}^o/\text{f}.$$

$\mathbf{C}$ is a symmetric g-inverse of the GLS equations or is the corresponding partition of a g-inverse of the mixed model equations and f is the number of rows in $\mathbf{K}'$ chosen to have full row rank. Now as in other ANOVA based methods of estimation of variances we can compute as though $\mathbf{u}$ is fixed and then take expectations of the resulting mean squares to estimate variances. The following precaution must be observed. $\mathbf{K}'\mathbf{u}$ must be estimable under a fixed $\mathbf{u}$ model. Then we compute

$$(\mathbf{u}^o)'\mathbf{K}(\mathbf{K}'\mathbf{C}\mathbf{K})^{-1}\mathbf{K}'\mathbf{u}^o/\text{f}, \tag{89}$$

where $\mathbf{u}^o$ is some solution to (90) and f = number of rows in $\mathbf{K}'$.

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}. \tag{90}$$

The assumption is that $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$. $\mathbf{C}$ is the lower $q \times q$ submatrix of a g-inverse of the coefficient matrix in (90). Then the expectation of (89) is

$$f^{-1}\ tr\ \mathbf{K}(\mathbf{K}'\mathbf{C}\mathbf{K})^{-1}\mathbf{K}'\ Var(\mathbf{u}) + \sigma_e^2. \tag{91}$$

This method seems particularly appropriate in the filled subclass case for then with interactions it is relatively easy to find estimable functions of $\mathbf{u}$. To illustrate, consider the two way mixed model of Section 11. Functions for estimating σ_a^2 are

$$\begin{pmatrix} a_1 + \bar{\gamma}_{1.} - a_4 - \bar{\gamma}_{4.} \\ a_2 + \bar{\gamma}_{2.} - a_4 - \bar{\gamma}_{4.} \\ a_3 + \bar{\gamma}_{3.} - a_4 - \bar{\gamma}_{4.} \end{pmatrix} /3.$$

Functions for estimating σ_γ^2 are

$$[\gamma_{ij} \ - \ \gamma_{i3} \ - \ \gamma_{4j} \ + \ \gamma_{34}]/6; \quad i = 1,2,3; \quad j = 1,2.$$

This is an example of a weighted square of means analysis.

The easiest solution to the OLS equations for the 2 way case is $\mathbf{a}^o$, $\mathbf{b}^o$ = null and $\gamma_{ij}^o = \bar{y}_{ij}$. Then the first set of functions can be estimated as $\bar{\gamma}_{i.}^o - \bar{\gamma}_{4.}^o$ $(i = 1,2,3)$. Reduce $\mathbf{K}'$ to this same dimension and take $\mathbf{C}$ as a 12×12 diagonal matrix with diagonal elements $= n_{ij}^{-1}$.

Chapter 11
MIVQUE of Variances and Covariances

C. R. Henderson

1984 - Guelph

The methods described in Chapter 10 for estimation of variances are quadratic, translation invariant, and unbiased. For the balanced design where there are equal numbers of observations in all subclasses and no covariates, equating the ANOVA mean squares to their expectations yields translation invariant, quadratic, unbiased estimators with minimum sampling variance regardless of the form of distribution, Albert (1976), see also Graybill and Wirtham (1956). Unfortunately, such an estimator cannot be derived in the unbalanced case unless **G** and **R** are known at least to proportionality. It is possible, however, to derive locally best, translation invariant, quadratic, unbiased estimators under the assumption of multivariate normality. This method is sometimes called MIVQUE and is due to C.R. Rao (1971). Additional pioneering work in this field was done by La Motte (1970,1971) and by Townsend and Searle (1971). By "locally best" is meant that if $\tilde{\mathbf{G}} = \mathbf{G}$ and $\tilde{\mathbf{R}} = \mathbf{R}$, the MIVQUE estimator has minimum sampling variance in the class of quadratic, unbiased, translation invariant estimators. $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ are prior values of **G** and **R** that are used in computing the estimators. For the models which we have described in this book MIVQUE based on the mixed model equations is computationally advantageous. A result due to La Motte (1970) and a suggestion given to me by Harville have been used in deriving this type of MIVQUE algorithm. The equations to be solved are in (1).

$$\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} + \tilde{\mathbf{G}}^{-1} \end{pmatrix} \begin{pmatrix} \beta^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \end{pmatrix} \quad (1)$$

These are mixed model equations based on the model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}. \quad (2)$$

We define $Var(\mathbf{u})$, $Var(\mathbf{e})$ and $Var(\mathbf{y})$ as in (2, 3, 4, 5, 6, 7, 8) of Chapter 10.

1 La Motte Result For MIVQUE

La Motte defined

$$Var(\mathbf{y}) = \mathbf{V} = \textstyle\sum_{i=1}^{k} \mathbf{V}_i\theta_i. \quad (3)$$

Then

$$\tilde{\mathbf{V}} = \sum_{i=1}^{k} \mathbf{V}_i \tilde{\theta}_i, \tag{4}$$

where $\tilde{\theta}_i$ are prior values of θ_i. The θ_i are unknown parameters and the $\mathbf{V}_i$ are known matrices of order $n \times n$. He proved that MIVQUE of $\boldsymbol{\theta}$ is obtained by computing

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)'\tilde{\mathbf{V}}^{-1}\mathbf{V}_i\tilde{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o),\ i = 1, \ldots,\ k, \tag{5}$$

equating these k quadratics to their expectations, and then solving for $\boldsymbol{\theta}$. $\boldsymbol{\beta}^o$ is any solution to equations

$$\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X}\boldsymbol{\beta}^o \ = \ \mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{y}. \tag{6}$$

These are GLS equations under the assumption that $\mathbf{V} \ = \ \tilde{\mathbf{V}}$.

2 Alternatives To La Motte Quadratics

In this section we show that other quadratics in $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o$ exist which yield the same estimates as the La Motte formulation. This is important because there may be quadratics easier to compute than those of (5), and their expectations may be easier to compute.

Let the k quadratics of (5) be denoted by $\mathbf{q}$. Let $E(\mathbf{q}) = \mathbf{B}\boldsymbol{\theta}$, where $\mathbf{B}$ is $k \times k$. Then provided $\mathbf{B}$ is nonsingular, MIVQUE of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}} = \mathbf{B}^{-1}\mathbf{q}. \tag{7}$$

Let $\mathbf{H}$ be any $k \times k$ nonsingular matrix. Compute a set of quadratics $\mathbf{Hq}$ and equate to their expectations.

$$E(\mathbf{Hq}) \ = \ \mathbf{H}E(\mathbf{q}) \ = \ \mathbf{HB}\theta. \tag{8}$$

Then an unbiased estimator is

$$\begin{aligned} \boldsymbol{\theta}^o &= (\mathbf{HB})^{-1}\mathbf{Hq} \\ &= \mathbf{B}^{-1}\mathbf{q} = \hat{\boldsymbol{\theta}}, \end{aligned} \tag{9}$$

the MIVQUE estimator of La Motte. Therefore, if we derive the La Motte quadratics, $\mathbf{q}$, for MIVQUE, we can find another set of quadratics which are also MIVQUE, and these are represented by $\mathbf{Hq}$, where $\mathbf{H}$ is nonsingular.

3 Quadratics Equal To La Motte's

The relationship between La Motte's model and ours is as follows

$$\mathbf{V} \text{ of LaMotte} = \mathbf{Z} \begin{pmatrix} \mathbf{G}_{11}g_{11} & \mathbf{G}_{12}g_{12} & \mathbf{G}_{13}g_{13} & \cdots \\ \mathbf{G}_{12}'g_{12} & \mathbf{G}_{22}g_{22} & \mathbf{G}_{23}g_{23} & \cdots \\ \mathbf{G}_{13}'g_{13} & \mathbf{G}_{23}'g_{23} & \mathbf{G}_{33}g_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \mathbf{Z}'$$

$$+ \begin{pmatrix} \mathbf{R}_{11}r_{11} & \mathbf{R}_{12}r_{12} & \mathbf{R}_{13}r_{13} & \cdots \\ \mathbf{R}'_{12}r_{12} & \mathbf{R}_{22}r_{22} & \mathbf{R}_{23}r_{23} & \cdots \\ \mathbf{R}'_{13}r_{13} & \mathbf{R}'_{13}r_{23} & \mathbf{R}_{33}r_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \tag{10}$$

$$= \mathbf{ZGZ}' + \mathbf{R}. \tag{11}$$

$$\text{or } \mathbf{V}_1\theta_1 = \mathbf{Z} \begin{pmatrix} \mathbf{G}_{11} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \mathbf{Z}' g_{11},$$

$$\mathbf{V}_2\theta_2 = \mathbf{Z} \begin{pmatrix} \mathbf{0} & \mathbf{G}_{12} & \mathbf{0} & \cdots \\ \mathbf{G}'_{12} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \mathbf{Z}' g_{12}, \tag{12}$$

etc., and

$$\mathbf{V}_{b+1}\theta_{b+1} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} r_{11},$$

$$\mathbf{V}_{b+2}\theta_{b+2} = \begin{pmatrix} \mathbf{0} & \mathbf{R}_{12} & \mathbf{0} & \cdots \\ \mathbf{R}'_{12} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} r_{12}, \tag{13}$$

etc. Define the first $b(b+1)/2$ of (12) as $\mathbf{ZG}^*_{ij}\mathbf{Z}'$ and the last $c(c+1)/2$ of (13) as $\mathbf{R}^*_{ij}$. Then for one of the first $b(b+1)/2$ of La Motte's quadratic we have

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)'\tilde{\mathbf{V}}^{-1}\mathbf{ZG}^*_{ij}\mathbf{Z}'\tilde{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o). \tag{14}$$

Write this as

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)'\tilde{\mathbf{V}}^{-1}\mathbf{Z}\tilde{\mathbf{G}}\tilde{\mathbf{G}}^{-1}\mathbf{G}^*_{ij}\tilde{\mathbf{G}}^{-1}\tilde{\mathbf{G}}\mathbf{Z}'\tilde{\mathbf{V}}^{-1}(\mathbf{y}\mathbf{X}\boldsymbol{\beta}^o). \tag{15}$$

This can be done because $\tilde{\mathbf{G}}\tilde{\mathbf{G}}^{-1} = \mathbf{I}$. Now note that $\tilde{\mathbf{G}}\mathbf{Z}'\tilde{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) = \hat{\mathbf{u}} =$ BLUP of $\mathbf{u}$ given $\mathbf{G} = \tilde{\mathbf{G}}$ and $\mathbf{R} = \tilde{\mathbf{R}}$. Consequently (15) can be written as

$$\hat{\mathbf{u}}'\tilde{\mathbf{G}}^{-1}\mathbf{G}^*_{ij}\tilde{\mathbf{G}}^{-1}\hat{\mathbf{u}}. \tag{16}$$

By the same type of argument the last c quadratics are

$$\hat{\mathbf{e}}'\tilde{\mathbf{R}}^{-1}\mathbf{R}^*_{ij}\tilde{\mathbf{R}}^{-1}\hat{\mathbf{e}}, \tag{17}$$

where $\hat{\mathbf{e}}$ is BLUP of $\mathbf{e}$ given that $\mathbf{G} = \tilde{\mathbf{G}}$ and $\mathbf{R} = \tilde{\mathbf{R}}$. Taking into account that

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{11}g_{11} & \mathbf{G}_{12}g_{12} & \cdots \\ \mathbf{G}_{12}'g_{12} & \mathbf{G}_{22}g_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix}$$

the matrices of the quadratics in $\hat{\mathbf{u}}$ can be computed easily. Let

$$\tilde{\mathbf{G}}^{-1} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} & \cdots \\ \mathbf{C}_{12}' & \mathbf{C}_{22} & \mathbf{C}_{23} & \cdots \\ \mathbf{C}_{13}' & \mathbf{C}_{23}' & \mathbf{C}_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} = [\mathbf{C}_1 \ \mathbf{C}_2 \ \mathbf{C}_3 \ \ldots].$$

For example,

$$\mathbf{C}_2 = \begin{pmatrix} \mathbf{C}_{12} \\ \mathbf{C}_{22} \\ \mathbf{C}_{23}' \\ \vdots \end{pmatrix}.$$

Then

$$\tilde{\mathbf{G}}^{-1}\mathbf{G}_{ii}^*\tilde{\mathbf{G}}^{-1} = \mathbf{C}_i\mathbf{G}_{ii}\mathbf{C}_i'. \tag{18}$$

$$\tilde{\mathbf{G}}^{-1}\mathbf{G}_{ij}^*\tilde{\mathbf{G}}^{-1} = \mathbf{C}_i\mathbf{G}_{ij}\mathbf{C}_j' + \mathbf{C}_j\mathbf{G}_{ij}'\mathbf{C}_i \text{ for } i \neq j. \tag{19}$$

The quadratics in $\hat{\mathbf{e}}$ are like (18) and (19) with $\tilde{\mathbf{R}}^{-1}$, $\mathbf{R}_{ij}$ substituted for $\tilde{\mathbf{G}}^{-1}$, $\mathbf{G}_{ij}$ and with $\tilde{\mathbf{R}}^{-1} = [\mathbf{C}_1 \ \mathbf{C}_2 \ \ldots]$. For special cases these quadratics simplify considerably. First consider the case in which all $g_{ij} = 0$. Then

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{11}g_{11} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{G}_{22}g_{22} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_{33}g_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix},$$

and

$$\mathbf{G}^{-1} = \begin{pmatrix} \mathbf{G}_{11}^{-1}g_{11}^{-1} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{G}_{22}^{-1}g_{22}^{-1} & \cdots \\ \vdots & \vdots & \end{pmatrix}.$$

Then the quadratics in $\hat{\mathbf{u}}$ become

$$\hat{\mathbf{u}}_i'\mathbf{G}_{ii}^{-1}g_{ii}^{-2}\hat{\mathbf{u}},$$

or an alternative is obviously

$$\hat{\mathbf{u}}_i'\mathbf{G}_{ii}^{-1}\hat{\mathbf{u}}_i,$$

obtained by multiplying these quadratics by

$$\text{dg}(g_{11}^2, g_{22}^2, \ldots). \tag{20}$$

Similarly if all $r_{ij} = 0$, the quadratics in $\hat{\mathbf{e}}$ can be converted to

$$\hat{\mathbf{e}}_i^{'} \mathbf{R}_{ii}^{-1} \hat{\mathbf{e}}_i. \tag{21}$$

The traditional mixed model for variance components reduces to a particularly simple form. Because all $g_{ij} = 0$, for $i \neq j$, all $\mathbf{G}_{ii} = \mathbf{I}$, and $\mathbf{R} = \mathbf{I}$, the quadratics can be written as

$$\hat{\mathbf{u}}_i^{'} \hat{\mathbf{u}}_i \ \ \text{i} = 1, \ldots, \ \text{b}, \quad \text{and} \quad \hat{\mathbf{e}}'\hat{\mathbf{e}}.$$

Pre-multiplying these quadratics by

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \\ \sigma_e^2/\sigma_1^2 & \sigma_e^2/\sigma_2^2 & \cdots & 1 \end{pmatrix}$$

we obtain

$$\begin{pmatrix} \hat{\mathbf{u}}_1^{'} \hat{\mathbf{u}}_1 \\ \hat{\mathbf{u}}_2^{'} \hat{\mathbf{u}}_2 \\ \vdots \\ \hat{\mathbf{e}}'\hat{\mathbf{e}} \ + \ \sum_i \frac{\sigma_e^2}{\sigma_i^2} \ \hat{\mathbf{u}}_i^{'} \hat{\mathbf{u}}_i \end{pmatrix}.$$

But the last of these quadratics is $\mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta}^o - \mathbf{y}'\mathbf{Z}\hat{\mathbf{u}}$, or a quantity corresponding to the least squares residual. This is the algorithm described in Henderson (1973).

One might wish in this model to estimate σ_e^2 by the OLS residual mean square, that is,

$$\hat{\sigma}_e^2 \ = \ [\mathbf{y}'\mathbf{y} - (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{y} - (\mathbf{u}^o)'\mathbf{Z}'\mathbf{y}]/[n - \text{rank}(\mathbf{X} \quad \mathbf{Z})],$$

where $\boldsymbol{\beta}^o$, $\mathbf{u}^o$ are some solution to OLS equations. If this is done, $\hat{\sigma}_e^2$ is not MIVQUE and neither are $\hat{\sigma}_i^2$, but they are probably good approximations to MIVQUE.

Another special case is the multiple trait model with additive genetic assumptions and elements of $\mathbf{u}$ for missing observations included. Ordering animals within traits,

$$\begin{aligned} \mathbf{y}' \ &= \ [\mathbf{y}_1^{'} \ \ \mathbf{y}_2^{'} \ldots]. \\ \mathbf{y}_i^{'} \ & \quad \text{is the vector of records on the } i^{th} \text{ trait.} \\ \mathbf{u}' \ &= \ (\mathbf{u}_1^{'} \ \ \mathbf{u}_2^{'} \ldots). \\ \mathbf{u}_i \ & \quad \text{is the vector of breeding values for the } i^{th} \text{ trait.} \\ \mathbf{e}' \ &= \ (\mathbf{e}_1^{'} \ \ \mathbf{e}_2^{'} \ldots). \end{aligned}$$

Every $\mathbf{u}_i$ vector has the same number of elements by including missing $\mathbf{u}$. Then

$$\mathbf{G} = \begin{pmatrix} \mathbf{A}g_{11} & \mathbf{A}g_{12} & \cdots \\ \mathbf{A}g_{12} & \mathbf{A}g_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix} = \mathbf{A} * \mathbf{G}_0, \tag{22}$$

where

$$\mathbf{G}_0 = \begin{pmatrix} g_{11} & g_{12} & \cdots \\ g_{12} & g_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix}$$

is the additive genetic variance- covariance matrix for a non-inbred population, and $*$ denotes the direct product operation.

$$\mathbf{G}^{-1} = \begin{pmatrix} \mathbf{A}^{-1}g^{11} & \mathbf{A}^{-1}g^{12} & \cdots \\ \mathbf{A}^{-1}g^{12} & \mathbf{A}^{-1}g^{22} & \cdots \\ \vdots & \vdots & \end{pmatrix} = \mathbf{A}^{-1} * \mathbf{G}_0^{-1}. \tag{23}$$

Applying the methods of (16) to (18) and (19) the quadratics in $\hat{\mathbf{u}}$ illustrated for a 3 trait model are

$$\begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{12}' & \mathbf{B}_{22} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{u}}_1'\mathbf{A}^{-1}\hat{\mathbf{u}}_1 \\ \hat{\mathbf{u}}_1'\mathbf{A}^{-1}\hat{\mathbf{u}}_2 \\ \hat{\mathbf{u}}_1'\mathbf{A}^{-1}\hat{\mathbf{u}}_3 \\ \hat{\mathbf{u}}_2'\mathbf{A}^{-1}\hat{\mathbf{u}}_2 \\ \hat{\mathbf{u}}_2'\mathbf{A}^{-1}\hat{\mathbf{u}}_3 \\ \hat{\mathbf{u}}_3'\mathbf{A}^{-1}\hat{\mathbf{u}}_3 \end{pmatrix}.$$

$$\mathbf{B}_{11} = \begin{pmatrix} g^{11}g^{11} & 2g^{11}g^{12} & 2g^{11}g^{13} \\ & 2g^{11}g^{22} + 2g^{12}g^{12} & 2g^{11}g^{23} + 2g^{12}g^{13} \\ & & 2g^{11}g^{33} + 2g^{13}g^{13} \end{pmatrix}.$$

$$\mathbf{B}_{12} = \begin{pmatrix} g^{12}g^{12} & 2g^{12}g^{13} & g^{13}g^{13} \\ 2g^{12}g^{22} & 2g^{12}g^{23} + 2g^{13}g^{22} & 2g^{13}g^{23} \\ 2g^{12}g^{23} & 2g^{12}g^{33} + 2g^{13}g^{23} & 2g^{13}g^{33} \end{pmatrix}.$$

$$\mathbf{B}_{22} = \begin{pmatrix} g^{22}g^{22} & 2g^{22}g^{23} & g^{23}g^{23} \\ & 2g^{22}g^{33} + 2g^{23}g^{23} & 2g^{23}g^{33} \\ & & g^{33}g^{33} \end{pmatrix}.$$

Premultiplying these quadratics in $\hat{\mathbf{u}}_i$ by the inverse of $\begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{12}' & \mathbf{B}_{22} \end{pmatrix}$ we obtain an equivalent set of quadratics that are

$$\hat{\mathbf{u}}_i'\mathbf{A}^{-1}\hat{\mathbf{u}}_j \text{ for i} = 1, \ldots, 3;\ \text{j} = \text{i}, \ldots, 3. \tag{24}$$

Similarly if there are no missing observations,

$$\mathbf{R} = \begin{pmatrix} \mathbf{I}\, r_{11} & \mathbf{I}\, r_{12} & \cdots \\ \mathbf{I}\, r_{12} & \mathbf{I}\, r_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix}$$

Then quadratics in $\hat{\mathbf{e}}$ are $\hat{\mathbf{e}}_i'\hat{\mathbf{e}}_j$ for i = 1, ..., t; j = i, ..., t.

4 Computation Of Missing $\hat{\mathbf{u}}$

In most problems, MIVQUE is easier to compute if missing $\mathbf{u}$ are included in $\hat{\mathbf{u}}$ rather than ignoring them. Section 3 illustrates this with $\mathbf{A}^{-1}$ being the matrix of all quadratics in $\hat{\mathbf{u}}$.

Three methods for prediction of elements of $\mathbf{u}$ not in the model for $\mathbf{y}$ were described in Chapter 5. Any of these can be used for MIVQUE. Probably the easiest is to include the missing ones in the mixed model equations.

5 Quadratics In $\hat{\mathbf{e}}$ With Missing Observations

When there are missing observations the quadratics in $\hat{\mathbf{e}}$ are easier to envision if we order the data by traits in animals rather than by animals in traits. Then $\mathbf{R}$ is block diagonal with the order of the i^{th} diagonal block being the number of traits recorded for the i^{th} animal. Now we do not need to store $\tilde{\mathbf{R}}^{-1}$ nor even all of the diagonal blocks. Rather we need to store only one block for each of the combinations of traits observed. For example, with 3 traits the possible combinations are

	Traits		
Combinations	1	2	3
1	X	X	X
2	X	X	-
3	X	-	X
4	-	X	X
5	X	-	-
6	-	X	-
7	-	-	X

There are $2^t - 1$ possible combinations for t traits. In the case of sequential culling the possible types are

Combinations	Traits		
	1	2	3
1	X	X	X
2	X	X	-
3	X	-	-

There are t possible combinations for t traits.

The block of $\tilde{\mathbf{R}}^{-1}$ for animals with the same traits measured will be identical. Thus if 50 animals have traits 1 and 2 recorded, there will be 50 identical 2×2 blocks in $\tilde{\mathbf{R}}^{-1}$, and only one of these needs to be stored.

The same principle applies to the matrices of quadratics in $\hat{\mathbf{e}}$. All of the quadratics are of the form $\hat{\mathbf{e}}_i^{'}\mathbf{Q}\hat{\mathbf{e}}_i$, where $\hat{\mathbf{e}}_i$ refers to the subvector of $\hat{\mathbf{e}}$ pertaining to the i^{th} animal. But animals with the same record combinations, will have identical matrices of quadratics for estimation of a particular variance or covariance. The computation of these matrices is simple. For a particular set of records let the block in $\tilde{\mathbf{R}}^{-1}$ be $\mathbf{P}$, which is symmetric and with order equal to the number of traits recorded. Label rows and columns by trait number. For example, suppose traits 1, 3, 7 are recorded. Then the rows (and columns) of $\mathbf{P}$ are understood to have labels 1, 3, 7. Let

$$\mathbf{P} \;\equiv\; (\mathbf{p}_1 \; \mathbf{p}_2 \; \mathbf{p}_3 \ldots),$$

where $\mathbf{p}_i$ is the i^{th} column vector of $\mathbf{P}$. Then the matrix of the quadratic for estimating r_{ii} is

$$\mathbf{p}_i\mathbf{p}_i^{'}. \tag{25}$$

The matrix for estimating r_{ij} $(i \neq j)$ is

$$\mathbf{p}_i\mathbf{p}_j^{'} \;+\; \mathbf{p}_i\mathbf{p}_j^{'}. \tag{26}$$

Let us illustrate with an animal having records on traits 2, 4, 7. The block of $\mathbf{R}$ corresponding to this type of information is

$$\begin{pmatrix} 6 & 4 & 3 \\ & 8 & 5 \\ & & 7 \end{pmatrix}.$$

Then the block corresponding to $\mathbf{R}^{-1}$ is the inverse of this, which is

$$\begin{pmatrix} .25410 & -.10656 & -.03279 \\ & .27049 & -.14754 \\ & & .26230 \end{pmatrix}.$$

Then the matrix for estimation of r_{22} is

$$\begin{pmatrix} .25410 \\ -.10656 \\ -.03279 \end{pmatrix} (.25410 \;-.10656 \;-.03279).$$

The matrix for computing r_{27} is

$$\begin{pmatrix} .25410 \\ -.10656 \\ -.03279 \end{pmatrix} (-.03279 \;\; -.14754 \;\; -.26230)$$

+ the transpose of this product.

6 Expectations Of Quadratics In $\hat{\mathbf{u}}$ And $\hat{\mathbf{e}}$

MIVQUE can be computed by equating certain quadratics in $\hat{\mathbf{u}}$ and in $\hat{\mathbf{e}}$ to their expectations. To find the expectations we need a g-inverse of the mixed model coefficient matrix with $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$, prior values, substituted for $\mathbf{G}$ and $\mathbf{R}$. The formulas for these expectations are in Section 6 of Chapter 10. It is obvious from these descriptions of expectations that extensive matrix products are required. However, some of the matrices have special forms such as diagonality, block diagonality, and symmetry. It is essential that these features be exploited. Also note that the trace of the products of several matrices can be expressed as the trace of the product of two matrices, say trace ($\mathbf{AB}$). Because only the sum of the diagonals of the product $\mathbf{AB}$ is required, it would be foolish to compute the off-diagonal elements. Some special computing algorithms are

$$\text{trace } (\mathbf{AB}) = \textstyle\sum_i \sum_j a_{ij} b_{ji} \tag{27}$$

when $\mathbf{A}$ and $\mathbf{B}$ are nonsymmetric.

$$\text{trace } (\mathbf{AB}) = \textstyle\sum_i a_{ii} b_{ii} \; + \; 2\sum_i \sum_{j>i} a_{ij} b_{ij} \tag{28}$$

when $\mathbf{A}$ and $\mathbf{B}$ are both symmetric.

$$\text{tr } (\mathbf{AB}) = \textstyle\sum_i a_{ii} b_{ii} \tag{29}$$

when either $\mathbf{A}$ or $\mathbf{B}$ or both are diagonal.

It is particularly important to take advantage of the form of matrices of quadratics in $\hat{\mathbf{e}}$ in animal models. When the data are ordered by traits within animals the necessary quadratics have the form $\sum_i \hat{\mathbf{e}}_i' \mathbf{Q} \hat{\mathbf{e}}_i$, where $\mathbf{Q}_i$ is a block of order equal to the number of traits observed in the i^{th} animal. Then the expectation of this quadratic is tr $\mathbf{Q}\; Var(\hat{\mathbf{e}}_i)$. Consequently we do not need to compute all elements of $Var(\hat{\mathbf{e}})$, but rather only the elements in blocks down the diagonal corresponding to the various $\mathbf{Q}$. In some cases, depending upon the form of $\mathbf{X}\boldsymbol{\beta}$, these blocks may be identical for animals with the same traits observed.

Many problems are such that the coefficient matrix is too large for computation of a g-inverse with present computers. Consequently we present in Section 7 an approximate MIVQUE based on computing an approximate g-inverse.

7 Approximate MIVQUE

MIVQUE for large data sets is prohibitively expensive with 1983 computers because a g-inverse of a very large coefficient matrix is required. Why not use an approximate g-inverse that is computationally feasible? This was the idea presented by Henderson(1980). The method is called "Diagonal MIVQUE" by some animal breeders. The feasibility of this method and the more general one presented in this section requires that an approximate g-inverse of $\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X}$ can be computed easily. First "absorb" $\boldsymbol{\beta}^o$ from the mixed model equations.

$$\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} + \tilde{\mathbf{G}}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \end{pmatrix}. \tag{30}$$

This gives

$$[\mathbf{Z}'\mathbf{P}\mathbf{Z} + \tilde{\mathbf{G}}^{-1}]\ \hat{\mathbf{u}} = \mathbf{Z}'\mathbf{P}\mathbf{y} \tag{31}$$

where $\mathbf{P} = \tilde{\mathbf{R}}^{-1} - \tilde{\mathbf{R}}^{-1}\mathbf{X}(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-}\mathbf{X}'\tilde{\mathbf{R}}^{-1}$, and $(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-}$ is chosen to be symmetric. From the coefficient matrix of (31) one may see some simple approximate solution to $\hat{\mathbf{u}}$, say $\tilde{\mathbf{u}}$. Corresponding to this solution is a matrix $\tilde{\mathbf{C}}_{11}$ such that

$$\tilde{\mathbf{u}} = \tilde{\mathbf{C}}_{11}\mathbf{Z}'\mathbf{P}\mathbf{y} \tag{32}$$

Interpret $\tilde{\mathbf{C}}_{11}$ as an approximation to

$$\mathbf{C}_{11} = [\mathbf{Z}'\mathbf{P}\mathbf{Z} + \tilde{\mathbf{G}}^{-1}]^{-1}$$

Then given $\tilde{\mathbf{u}}$,

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-}\ (\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} - \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}\tilde{\mathbf{u}}).$$

Thus an approximate g-inverse to the coefficient matrix is

$$\tilde{\mathbf{C}} = \begin{pmatrix} \tilde{\mathbf{C}}_{00} & \tilde{\mathbf{C}}_{01} \\ \tilde{\mathbf{C}}_{10} & \tilde{\mathbf{C}}_{11} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{C}}_0 \\ \tilde{\mathbf{C}}_1 \end{pmatrix}. \tag{33}$$

$$\begin{aligned} \tilde{\mathbf{C}}_{00} &= (\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-} + (\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-}\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}\tilde{\mathbf{C}}_{11}\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X}(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-}. \\ \tilde{\mathbf{C}}_{01} &= (\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-}\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}\tilde{\mathbf{C}}_{11}. \\ \tilde{\mathbf{C}}_{10} &= \tilde{\mathbf{C}}_{11}\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X}(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-}. \end{aligned}$$

This matrix post-multiplied by $\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \end{pmatrix}$ equals $\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\mathbf{u}} \end{pmatrix}$. Note that $\tilde{\mathbf{C}}_{11}$ may be non-symmetric.

What are some possibilities for finding an approximate easy solution to $\mathbf{u}$ and consequently for writing $\tilde{\mathbf{C}}_{11}$? The key to this decision is the pattern of elements of the matrix of (31). If the diagonal is large relative to off-diagonal elements of the same row for every row, setting $\tilde{\mathbf{C}}_{11}$ to the inverse of a diagonal matrix formed from the diagonals of the coefficient matrix is a logical choice. Harville suggested that for the two way mixed variance components model one might solve for the main effect elements of $\mathbf{u}$ by using only the diagonals, but the interaction terms would be solved by adjusting the right hand side for the previously estimated associated main effects and then dividing by the diagonal. This would result in a lower triangular $\tilde{\mathbf{C}}_{11}$.

The multi-trait equations would tend to exhibit block diagonal dominance if the elements of $\mathbf{u}$ are ordered traits within animals. Then $\tilde{\mathbf{C}}_{11}$ might well take the form

$$\begin{pmatrix} \mathbf{B}_1^{-1} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{B}_2^{-1} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

where $\mathbf{B}_i^{-1}$ is the inverse of the i^{th} diagonal block, $\mathbf{B}_i$. Having solved for $\tilde{\mathbf{u}}$ and $\tilde{\boldsymbol{\beta}}$ and having derived $\tilde{\mathbf{C}}$ one would then proceed to compute quadratics in $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{e}}$ as in regular MIVQUE. Their expectations can be found as described in Section 7 of Chapter 10 except that $\tilde{\mathbf{C}}$ is substituted for $\mathbf{C}$.

8 MIVQUE (0)

MIVQUE simplifies greatly in the conventional variance components model if the priors are

$$\tilde{g}_{ii}/\tilde{r}_{11} = \hat{\sigma}_i^2/\hat{\sigma}_e^2 = 0 \text{ for all i} = 1, \ldots, \text{ b}.$$

Now

$$\tilde{\mathbf{V}} = \mathbf{I}\hat{\sigma}_e^2, \ \boldsymbol{\beta}^o = (\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{y},$$

and

$$\begin{aligned} &(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)' \ \tilde{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{G}_{ii}\mathbf{Z}_i'\tilde{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) \\ &= \mathbf{y}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}')'\mathbf{Z}_i\mathbf{Z}_i'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}')\mathbf{y}/\hat{\sigma}_e^4. \end{aligned} \quad (34)$$

Note that this, except for the constant, $\hat{\sigma}_e^4$, is simply the sum of squares of right hand sides of the OLS equations pertaining to $\mathbf{u}_i^o$ after absorbing $\boldsymbol{\beta}^o$. This is easy to compute, and the expectations are simple. Further, for estimation of σ_e^2 we derive the quadratic,

$$\mathbf{y}'\mathbf{y} - (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{y}, \quad (35)$$

and the expectation of this is simple.

This method although simple to compute has been found to have large sampling variances when σ_i^2/σ_e^2 departs very much from 0, Quaas and Bolgiano(1979). Approximate MIVQUE involving diagonal $\tilde{\mathbf{C}}_{11}$ is not much more difficult and gives substantially smaller variances when $\sigma_i^2/\sigma_e^2 \ > \ 0$.

For the general model with $\tilde{\mathbf{G}} = \mathbf{0}$ the MIVQUE computations are effected as follows. This is an extension of MIVQUE(0) with $\mathbf{R} \neq \mathbf{I}\sigma_e^2$, and Var $(\mathbf{u}_i) \neq \mathbf{I}\sigma_i^2$. Absorb $\boldsymbol{\beta}^o$ from equations

$$\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \end{pmatrix}. \tag{36}$$

Then compute

$$\mathbf{y}'\tilde{\mathbf{R}}^{-1}\mathbf{y} - (\mathbf{y}'\tilde{\mathbf{R}}^{-1}\mathbf{X})(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-}\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \tag{37}$$

and $\mathbf{r}_i\mathbf{G}_{ij}\mathbf{r}_j$ i=1, . . . , b; j=i, . . . , b, where $\mathbf{r}_i$ = absorbed right hand side for $\mathbf{u}_i^o$ equations. Estimate r_{ij} from following quadratics

$$\hat{\mathbf{e}}_i\tilde{\mathbf{R}}_{ij}\hat{\mathbf{e}}_j$$

where

$$\hat{\mathbf{e}} \ = \ [\mathbf{I} - \mathbf{X}(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^{-}\mathbf{X}'\tilde{\mathbf{R}}^{-1}]\mathbf{y}. \tag{38}$$

9 MIVQUE For Singular G

The formulation of (16) cannot be used if $\mathbf{G}$ is singular, neither can (18) if $\mathbf{G}_{ii}$ is singular, nor (25) if $\mathbf{A}$ is singular. A simple modification gets around this problem. Solve for $\hat{\boldsymbol{\alpha}}$ in (51) of Chapter 5. Then for (16) substitute

$$\hat{\boldsymbol{\alpha}}'\mathbf{G}_{ij}^*\hat{\boldsymbol{\alpha}}, \quad \text{where} \quad \hat{\mathbf{u}} = \tilde{\mathbf{G}}\hat{\boldsymbol{\alpha}} \tag{39}$$

For (20) substitute

$$\hat{\boldsymbol{\alpha}}_i'\mathbf{G}_{ii}\hat{\boldsymbol{\alpha}}_i. \tag{40}$$

For (25) substitute

$$\hat{\boldsymbol{\alpha}}_i'\mathbf{A}\hat{\boldsymbol{\alpha}}_j \tag{41}$$

See Section 16 for expectations of quadratics in $\hat{\boldsymbol{\alpha}}$.

10 MIVQUE For The Case $\mathbf{R} = \mathbf{I}\sigma_e^2$

When $\mathbf{R} \ = \mathbf{I}\sigma_e^2$ the mixed model equations can be written as

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \sigma_e^2\tilde{\mathbf{G}}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} \ = \ \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}. \tag{42}$$

If $\boldsymbol{\beta}^o$ is absorbed, the equations in $\hat{\mathbf{u}}$ are

$$(\mathbf{Z}'\mathbf{P}\mathbf{Z} + \sigma_e^2 \tilde{\mathbf{G}}^{-1})\hat{\mathbf{u}} = \mathbf{Z}'\mathbf{P}\mathbf{y}, \tag{43}$$

where

$$\mathbf{P} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'.$$

Let

$$(\mathbf{Z}'\mathbf{P}\mathbf{Z} + \sigma_e^2 \tilde{\mathbf{G}}^{-1})^{-1} = \mathbf{C}. \tag{44}$$

Then

$$\begin{aligned}
\hat{\mathbf{u}} &= \mathbf{C}\mathbf{Z}'\mathbf{P}\mathbf{y}. && (45)\\
\hat{\mathbf{u}}'\mathbf{Q}_i\hat{\mathbf{u}} &= \mathbf{y}'\mathbf{P}\mathbf{Z}\mathbf{C}'\mathbf{Q}_i\mathbf{C}\mathbf{Z}'\mathbf{P}\mathbf{y} &&\\
&= \text{tr } \mathbf{C}'\mathbf{Q}_i\mathbf{C}\mathbf{Z}'\mathbf{P}\mathbf{y}\mathbf{y}'\mathbf{P}\mathbf{Z}. && (46)\\
E(\hat{\mathbf{u}}'\mathbf{Q}_i\hat{\mathbf{u}}) &= \text{tr } \mathbf{C}'\mathbf{Q}_i\mathbf{C}\ Var(\mathbf{Z}'\mathbf{P}\mathbf{y}). && (47)\\
Var(\mathbf{Z}'\mathbf{P}\mathbf{y}) &= \textstyle\sum_{i=1}^{b}\sum_{j=1}^{b} \mathbf{Z}'\mathbf{P}\mathbf{Z}_i\mathbf{G}_{ij}\mathbf{Z}_j'\mathbf{P}\mathbf{Z}g_{ij} &&\\
&\quad + \mathbf{Z}'\mathbf{P}\mathbf{P}\mathbf{Z}\sigma_e^2. && (48)
\end{aligned}$$

One might wish to obtain an approximate MIVQUE by estimating σ_e^2 from the OLS residual. When this is done, the expectation of the residual is $[n - r(\mathbf{X}\ \mathbf{Z})]\ \sigma_e^2$ regardless of the value of $\tilde{\mathbf{G}}$. This method is easier than true MIVQUE and has advantages in computation of sampling variances because the estimator of σ_e^2 is uncorrelated with the various $\hat{\mathbf{u}}'\mathbf{Q}\hat{\mathbf{u}}$. This method also is computable with absorption of $\boldsymbol{\beta}^o$.

A further simplification based on the ideas of Section 7, would be to look for some simple approximate solution to $\hat{\mathbf{u}}$ in (44). Call this solution $\tilde{\mathbf{u}}$ and the corresponding approximate g-inverse of the matrix of (44) $\tilde{\mathbf{C}}$. Then proceed as in (46) ... (48) except substitute $\tilde{\mathbf{u}}$ for $\hat{\mathbf{u}}$ and $\tilde{\mathbf{C}}$ for $\mathbf{C}$.

11 Sampling Variances

MIVQUE consists of computing $\hat{\mathbf{u}}'\mathbf{Q}_i\hat{\mathbf{u}}$, $i = 1, \ldots, b$, where b = number of elements of $\mathbf{g}$ to be estimated, and $\hat{\mathbf{e}}'\mathbf{Q}_j\hat{\mathbf{e}}$, $j = 1, \ldots, t$, where t = number of elements of $\mathbf{r}$ to be estimated. Let a g-inverse of the mixed model matrix be $\mathbf{C} \equiv \begin{pmatrix} \mathbf{C}_\beta \\ \mathbf{C}_u \end{pmatrix}$, and let $\mathbf{W} = (\mathbf{X}\quad \mathbf{Z})$.

Then

$$\begin{aligned}
\hat{\mathbf{u}} &= \mathbf{C}_u\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{y},\\
\hat{\mathbf{u}}'\mathbf{Q}\hat{\mathbf{u}} &= \mathbf{y}'\tilde{\mathbf{R}}^{-1}\mathbf{W}\mathbf{C}_u'\mathbf{Q}_i\mathbf{C}_u\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \equiv \mathbf{y}'\mathbf{B}_i\mathbf{y}, \qquad (49)\\
\hat{\mathbf{e}} &= (\mathbf{I} - \mathbf{W}\mathbf{C}\mathbf{W}'\tilde{\mathbf{R}}^{-1})\mathbf{y},
\end{aligned}$$

and

$$\hat{\mathbf{e}}'\mathbf{Q}_j\hat{\mathbf{e}} = \mathbf{y}'[\mathbf{I} - \mathbf{WCW}'\tilde{\mathbf{R}}^{-1}]'\mathbf{Q}_j[\mathbf{I} - \mathbf{WCW}'\tilde{\mathbf{R}}^{-1}]\mathbf{y} \equiv \mathbf{y}'\mathbf{F}_j\mathbf{y}. \quad (50)$$

Let

$$E\begin{pmatrix} \mathbf{y}'\mathbf{B}_1\mathbf{y} \\ \vdots \\ \mathbf{y}'\mathbf{F}_1\mathbf{y} \\ \vdots \end{pmatrix} = \mathbf{P}\begin{pmatrix} \mathbf{g} \\ \mathbf{r} \end{pmatrix} = \mathbf{P}\boldsymbol{\theta}, \quad \text{where } \boldsymbol{\theta} = \begin{pmatrix} \mathbf{g} \\ \mathbf{r} \end{pmatrix}.$$

Then MIVQUE of $\boldsymbol{\theta}$ is

$$\mathbf{P}^{-1}\begin{pmatrix} \mathbf{y}'\mathbf{B}_1\mathbf{y} \\ \vdots \\ \mathbf{y}'\mathbf{F}_1\mathbf{y} \\ \vdots \end{pmatrix} = \begin{pmatrix} \mathbf{y}'\mathbf{H}_1\mathbf{y} \\ \mathbf{y}'\mathbf{H}_2\mathbf{y} \\ \vdots \end{pmatrix}. \quad (51)$$

Then

$$Var(\hat{\theta}_i) = 2\ tr[\mathbf{H}_i\ Var(\mathbf{y})]^2. \quad (52)$$

$$Cov(\hat{\theta}_i, \hat{\theta}_j) = 2\ tr\mathbf{H}_i\ [Var(\mathbf{y})]\ \mathbf{H}_j\ [Var(\mathbf{y})]. \quad (53)$$

These are of course quadratics in unknown elements of $\mathbf{g}$ and $\mathbf{r}$. A numerical solution is easier. Let $\tilde{\mathbf{V}} = Var(\mathbf{y})$ for some assumed values of $\mathbf{g}$ and $\mathbf{r}$. Then

$$Var(\hat{\theta}_i) = 2\ tr(\mathbf{H}_i\tilde{\mathbf{V}})^2. \quad (54)$$

$$Cov(\hat{\theta}_i, \hat{\theta}_j) = 2\ tr(\mathbf{H}_i\tilde{\mathbf{V}}\mathbf{H}_j\tilde{\mathbf{V}}). \quad (55)$$

If approximate MIVQUE is computed using $\tilde{\mathbf{C}}$ an approximation to $\mathbf{C}$, the computations are the same except that $\tilde{\mathbf{C}}$, $\tilde{\mathbf{u}}$, $\tilde{\mathbf{e}}$ are used in place of $\mathbf{C}$, $\hat{\mathbf{u}}$, $\hat{\mathbf{e}}$.

11.1 Result when σ_e^2 estimated from OLS residual

When $\mathbf{R} = \mathbf{I}\sigma_e^2$, one can estimate σ_e^2 by the residual mean square of OLS and an approximate MIVQUE obtained. The quadratics to be computed in addition to $\hat{\sigma}_e^2$ are only $\hat{\mathbf{u}}'\mathbf{Q}_i\hat{\mathbf{u}}$. Let

$$E\begin{pmatrix} \hat{\mathbf{u}}'\mathbf{Q}_i\hat{\mathbf{u}} \\ \vdots \\ \hat{\sigma}_e^2 \end{pmatrix} = \begin{pmatrix} \mathbf{P} & \mathbf{f} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathbf{g} \\ \sigma_e^2 \end{pmatrix}.$$

Then

$$\begin{pmatrix} \hat{\mathbf{g}} \\ \hat{\sigma}_e^2 \end{pmatrix} = \begin{pmatrix} \mathbf{P} & \mathbf{f} \\ 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} \hat{\mathbf{u}}'\mathbf{Q}_i\hat{\mathbf{u}} \\ \vdots \\ \hat{\sigma}_e^2 \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{u}}'\mathbf{H}_1\hat{\mathbf{u}} \\ \hat{\mathbf{u}}'\mathbf{H}_2\hat{\mathbf{u}} \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} s_1\hat{\sigma}_e^2 \\ s_2\hat{\sigma}_e^2 \\ \vdots \\ \hat{\sigma}_e^2 \end{pmatrix}. \quad (56)$$

Then

$$Var(\hat{g}_i) = 2\, tr[\mathbf{H}_i\, Var(\hat{\mathbf{u}})]^2 + s_i^2\, Var(\hat{\sigma}_e^2). \quad (57)$$

$$Cov(\hat{\mathbf{g}}_i, \hat{\mathbf{g}}_j) = 2\, tr[\mathbf{H}_i\, Var(\hat{\mathbf{u}})\, \mathbf{H}_j\, Var(\hat{\mathbf{u}})] + s_i s_j\, Var(\hat{\sigma}_e^2). \quad (58)$$

where

$$Var(\hat{\mathbf{u}}) = \mathbf{C}_u[Var(\mathbf{r})]\mathbf{C}_u',$$

and $\mathbf{r}$ equals the right hand sides of mixed model equations.

$$Var(\hat{\mathbf{u}}) = \mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}\mathbf{G}\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{W} + \mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{R}\tilde{\mathbf{R}}^{-1}\mathbf{W}. \quad (59)$$

If $Var(\mathbf{r})$ is evaluated with the same values of $\mathbf{G}$ and $\mathbf{R}$ used in the mixed model equations, namely $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$, then

$$Var(\mathbf{r}) = \mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}\tilde{\mathbf{G}}\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{W} + \mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{W}. \quad (60)$$

$$Var(\hat{\sigma}_e^2) = 2\sigma_e^4/[n - \text{rank}\ (\mathbf{W})], \quad (61)$$

where $\hat{\sigma}_e^2$ is the OLS residual mean square. This would presumably be evaluated for $\sigma_e^2 = \hat{\sigma}_e^2$.

12 Illustrations Of Approximate MIVQUE

12.1 MIVQUE with $\hat{\sigma}_e^2$ = OLS residual

We next illustrate several approximate MIVQUE using as $\hat{\sigma}_e^2$ the OLS residual. The same numerical example of treatments by sires in Chapter 10 is employed. In all of these we absorb $\boldsymbol{\beta}^o$ to obtain the equations already presented in (70) to (72) in Chapter 10. We use prior $\sigma_e^2/\sigma_s^2 = 10$, $\sigma_e^2/\sigma_{ts}^2 = 5$ as in Chapter 10. Then the equations in $\hat{\mathbf{s}}$ and $\hat{\mathbf{ts}}$ are those of (70) to (72) with 10 added to the first 4 diagonals and 5 to the last 10 diagonals. The inverse of this matrix is in (62), (63) and (64). This gives the solution

$$\hat{\mathbf{s}}' = [-.02966,\ .17793,\ .02693,\ -.17520].$$

$$\hat{\mathbf{ts}} = [.30280,\ .20042,\ .05299,\ -.55621,\ -.04723,\ .04635,\ .00088,\ -.31489,\ .10908,\ .20582].$$

$$(\hat{\mathbf{ts}})'\hat{\mathbf{ts}} = .60183,\ \hat{\mathbf{s}}'\hat{\mathbf{s}} = .06396.$$

Upper left 7×7

$$\begin{pmatrix} .0713 & .0137 & .0064 & .0086 & -.0248 & .0065 & .0070 \\ & .0750 & .0051 & .0062 & .0058 & -.0195 & .0052 \\ & & .0848 & .0037 & .0071 & .0048 & -.0191 \\ & & & .0815 & .0118 & .0081 & .0069 \\ & & & & .1331 & .0227 & .0169 \\ & & & & & .1486 & .0110 \\ & & & & & & .1582 \end{pmatrix} \quad (62)$$

Upper right 7×7 and (lower left $7 \times 7)'$

$$\begin{pmatrix} .0112 & -.0178 & .0121 & .0057 & -.0147 & .0088 & .0060 \\ .0085 & .0126 & -.0177 & .0050 & .0089 & -.0129 & .0040 \\ .0071 & .0061 & .0052 & -.0113 & -.0004 & .0001 & .0003 \\ -.0268 & -.0009 & .0004 & .0006 & .0063 & .0040 & -.0102 \\ .0273 & .0092 & -.0066 & -.0027 & .0081 & -.0045 & -.0036 \\ .0177 & -.0058 & .0071 & -.0014 & -.0039 & .0054 & -.0015 \\ .0138 & -.0025 & -.0011 & .0036 & -.0003 & .0004 & -.0001 \end{pmatrix} \quad (63)$$

Lower right 7×7

$$\begin{pmatrix} .1412 & -.0009 & .0005 & .0004 & -.0039 & -.0013 & .0052 \\ & .1489 & .0364 & .0147 & .0063 & -.0054 & -.0009 \\ & & .1521 & .0115 & -.0057 & .0055 & .0002 \\ & & & .1737 & -.0006 & -.0001 & .0007 \\ & & & & .1561 & .0274 & .0164 \\ & & & & & .1634 & .0092 \\ & & & & & & .1744 \end{pmatrix} \quad (64)$$

The expectation of $(\hat{\mathbf{ts}})'\hat{\mathbf{ts}}$ is

$$E[\mathbf{r}'\mathbf{C}_2'\mathbf{C}_2\mathbf{r}] = tr\mathbf{C}_2'\mathbf{C}_2 \; Var(\mathbf{r}),$$

where $\mathbf{r}$ = right hand sides of the absorbed equations, and $\mathbf{C}_2$ is the last 10 rows of the inverse above.

$$\begin{aligned} Var(\mathbf{r}) \;&=\; \text{matrix of (10.73) to (10.75) } \sigma_{ts}^2 + \text{matrix of (10.76)} \\ &\quad \text{to (10.78) } \sigma_s^2 + \text{matrix of (10.70) to (10.72) } \sigma_e^2. \end{aligned}$$

$\mathbf{C}_2'\mathbf{C}_2$ is in (65), (66), and (67). Similarly $\mathbf{C}_1'\mathbf{C}_1$ is in (68), (69), and (70) where $\mathbf{C}_2, \mathbf{C}_1$ refer to last 10 rows and last 4 rows of (62) to (64) respectively. This leads to expectations as follows

$$\begin{aligned} E(\hat{\mathbf{ts}}'\hat{\mathbf{ts}}) &= .23851 \; \sigma_e^2 + .82246 \; \sigma_{ts}^2 + .47406 \; \sigma_s^2. \\ E(\hat{\mathbf{s}}'\hat{\mathbf{s}}) &= .03587 \; \sigma_e^2 + .11852 \; \sigma_{ts}^2 + .27803 \; \sigma_s^2. \end{aligned}$$

Using $\hat{\sigma}_e^2 = .3945$ leads then to estimates,

$$\hat{\sigma}_{ts}^2 \;=\; .6815, \quad \hat{\sigma}_s^2 \;= -.1114.$$

Upper left 7×7

$$.01 \begin{pmatrix} .1657 & -.0769 & -.0301 & -.0588 & -.3165 & .0958 & .0982 \\ & .1267 & -.0167 & -.0331 & .0987 & -.2867 & .0815 \\ & & .0682 & -.0215 & .0977 & .0815 & -.2809 \\ & & & .1134 & .1201 & .1094 & .1012 \\ & & & & 1.9485 & .6920 & .5532 \\ & & & & & 2.3159 & .4018 \\ & & & & & & 2.5645 \end{pmatrix} \tag{65}$$

Upper right 7×7 and (lower left 7×7)'

$$.01 \begin{pmatrix} .1224 & -.2567 & .1594 & .0973 & -.2418 & .1380 & .1038 \\ .1065 & .1577 & -.2466 & .0889 & .1368 & -.2127 & .0758 \\ .1018 & .1007 & .0901 & -.1909 & -.0029 & .0012 & .0041 \\ -.3307 & -.0017 & -.0029 & .0046 & .1078 & .0759 & -.1837 \\ .8063 & .2207 & -.1480 & -.0726 & .2062 & -.1135 & -.0927 \\ .5902 & -.1364 & .1821 & -.0457 & -.1033 & .1505 & -.0473 \\ .4805 & -.0689 & -.0411 & .1101 & -.0059 & .0085 & -.0026 \end{pmatrix} \tag{66}$$

Lower right 7×7

$$.01 \begin{pmatrix} 2.1231 & -.0153 & .0071 & .0082 & -.0970 & -.0456 & .1426 \\ & 2.3897 & 1.0959 & .5144 & .1641 & -.1395 & -.0247 \\ & & 2.4738 & .4303 & -.1465 & .1450 & .0016 \\ & & & 3.0553 & -.0176 & -.0055 & .0231 \\ & & & & 2.5572 & .8795 & .5633 \\ & & & & & 2.7646 & .3559 \\ & & & & & & 3.0808 \end{pmatrix} \tag{67}$$

Upper right 7×7

$$.01 \begin{pmatrix} .5391 & .2087 & .1100 & .1422 & -.1542 & .0299 & .0510 \\ & .5879 & .0925 & .1109 & .0208 & -.1295 & .0429 \\ & & .7271 & .0705 & .0520 & .0382 & -.1522 \\ & & & .6764 & .0814 & .0613 & .0583 \\ & & & & .0840 & -.0145 & -.0199 \\ & & & & & .0510 & -.0091 \\ & & & & & & .0488 \end{pmatrix} \tag{68}$$

Upper right 7×7 and (lower left 7×7)'

$$.01 \begin{pmatrix} .0732 & -.1066 & .0656 & .0411 & -.0878 & .0484 & .0393 \\ .0658 & .0728 & -.1130 & .0402 & .0503 & -.0820 & .0317 \\ .0620 & .0464 & .0431 & -.0896 & -.0063 & .0017 & .0046 \\ -.2010 & -.0126 & .0043 & .0083 & .0438 & .0319 & -.0757 \\ -.0496 & .0548 & -.0360 & -.0187 & .0488 & -.0244 & -.0244 \\ -.0274 & -.0339 & .0450 & -.0111 & -.0220 & .0340 & -.0120 \\ -.0198 & -.0183 & -.0103 & .0286 & -.0006 & .0020 & -.0014 \end{pmatrix} \tag{69}$$

Lower right 7×7

$$.01 \begin{pmatrix} .0968 & -.0025 & .0014 & .0012 & -.0261 & -.0116 & .0377 \\ & .0514 & -.0406 & -.0108 & .0366 & -.0321 & -.0045 \\ & & .0485 & -.0079 & -.0335 & .0335 & 0 \\ & & & .0187 & -.0031 & -.0014 & .0045 \\ & & & & .0335 & -.0219 & -.0117 \\ & & & & & .0258 & -.0039 \\ & & & & & & .0156 \end{pmatrix} \tag{70}$$

12.2 Approximate MIVQUE using a diagonal g-inverse

An easy approximate MIVQUE involves solving for $\hat{\mathbf{u}}$ in the reduced equations by dividing the right hand sides by the corresponding diagonal coefficient. Thus the approximate $\mathbf{C}$, denoted by $\tilde{\mathbf{C}}$ is diagonal with diagonal elements the reciprocal of the diagonals of (10.70) to (10.72). This gives

$\tilde{\mathbf{C}}$ = dg (.0516, .0598, .0788, .0690, .1059, .1333, .1475, .1161, .1263, .1304, .1690, .1429, .1525, .1698)

and an approximate solution,

$\tilde{\mathbf{u}}'$ = (.0057, .2758, .0350, -.3563, .4000, .2889, .0656, -.7419, -.1263, .1304, 0, -.3810, .2203, .2076).

Then $(\hat{\mathbf{ts}})'\hat{\mathbf{ts}} = 1.06794$ with expectation,

$$.4024\ \sigma_e^2\ +\ 1.5570\ (\sigma_{ts}^2\ +\ \sigma_s^2).$$

Also $\hat{\mathbf{s}}'\hat{\mathbf{s}} = .20426$ with expectation,

$$.0871\ \sigma_e^2\ +\ .3510\ \sigma_{ts}^2\ +\ .7910\ \sigma_s^2.$$

Now

$\mathbf{C}_2\mathbf{C}_2' = $ dg (0, 0, 0, 0, 1.1211, 1.7778, 2.1768, 1.3486, 1.5959, 1.7013, 2.8566, 2.0408, 2.3269, 2.8836)/100,

and

$\mathbf{C}_1\mathbf{C}_1' = $ dg (.2668, .3576, .6205, .4756, 0, 0, 0, 0, 0, 0, 0, 0,, 0, 0, 0, 0, 0, 0)/100.

Consequently one would need to compute only the diagonals of (10.70) to (10.72), if one were to use this method of estimation.

12.3 Approximate MIVQUE using a block diagonal approximate g-inverse

Examination of (10.70) to (10.72) shows that a subset of coefficients, namely [s_j, ts_{1j}, $ts_{2j}\ldots$] tends to be dominant. Consequently one might wish to exploit this structure. If the $\hat{ts}_{ij}$ were reordered by i within j and the interactions associated with s_i placed adjacent to s_i, the matrix would exhibit block diagonal dominance. Consequently we solve for $\tilde{\mathbf{u}}$ in equations with the coefficient matrix zeroed except for coefficients of s_j and associated ts_{ij}, etc. blocks. This matrix is in (71, 72, and 73) below.

Upper left 7 × 7

$$\begin{pmatrix} 19.361 & 0 & 0 & 0 & 4.444 & 0 & 0 \\ & 16.722 & 0 & 0 & 0 & 2.5 & 0 \\ & & 12.694 & 0 & 0 & 0 & 1.778 \\ & & & 14.5 & 0 & 0 & 0 \\ & & & & 9.444 & 0 & 0 \\ & & & & & 7.5 & 0 \\ & & & & & & 6.778 \end{pmatrix} \quad (71)$$

Upper right 7 × 7 and (lower left 7 × 7)'

$$\begin{pmatrix} 0 & 2.917 & 0 & 0 & 2. & 0 & 0 \\ 0 & 0 & 2.667 & 0 & 0 & 1.556 & 0 \\ 0 & 2 & 0 & .917 & 0 & 0 & 0 \\ 3.611 & 0 & 0 & 0 & 0 & 0 & .889 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (72)$$

Lower right 7 × 7

$$\text{dg}\ (8.611,\ 7.9167,\ 7.6667,\ 5.9167,\ 7.0,\ 6.556,\ 5.889) \quad (73)$$

A matrix like (71) to (73) is easy to invert if we visualize the diagonal blocks with re-ordering. For example,

$$\begin{pmatrix} 19.361 & 4.444 & 2.917 & 2.000 \\ & 9.444 & 0 & 0 \\ & & 7.917 & 0 \\ & & & 7.000 \end{pmatrix}^{-1} = \begin{pmatrix} .0640 & -.0301 & -.0236 & -.0183 \\ & .1201 & .0111 & .0086 \\ & & .1350 & .0067 \\ & & & .1481 \end{pmatrix}.$$

This illustrates that only 4^2 or 3^2 order matrices need to be inverted. Also, each of those has a diagonal submatrix of order either 3 or 2. The resulting solution vector is

(-.0343, .2192, .0271, -.2079, .4162, .2158, .0585, -.6547, -.1137, .0542, -.0042, -.3711, .1683, .2389).

This gives $(\tilde{\mathbf{ts}})'\tilde{\mathbf{ts}} = .8909$ with expectation

$$.32515\ \sigma_e^2 + 1.22797\ \sigma_{ts}^2 + .79344\ \sigma_s^2,$$

and $\tilde{\mathbf{s}}'\tilde{\mathbf{s}} = .0932$ with expectation

$$.05120\ \sigma_e^2 + .18995\ \sigma_{ts}^2 + .45675\ \sigma_s^2.$$

12.4 Approximate MIVQUE using a triangular block diagonal approximate g-inverse

Another possibility for finding an approximate solution is to compute $\tilde{\mathbf{s}}$ by dividing the right hand side by the corresponding diagonal. Then $\tilde{\mathbf{ts}}$ are solved by adjusting the right hand side for the associated $\tilde{\mathbf{s}}$ and dividing by the diagonal coefficient. This leads to a block triangular coefficient matrix when **ts** are placed adjacent to **s**. Without such re-ordering the matrix is as shown in (74), (75), and (76).

Upper left 7×7

$$\begin{pmatrix} 19.361 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 16.722 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 12.694 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 14.5 & 0 & 0 & 0 \\ 4.444 & 0 & 0 & 0 & 9.444 & 0 & 0 \\ 0 & 2.5 & 0 & 0 & 0 & 7.5 & 0 \\ 0 & 0 & 1.778 & 0 & 0 & 0 & 6.778 \end{pmatrix} \quad (74)$$

Upper right 7×7 = null matrix

Lower left 7×7

$$\begin{pmatrix} 0 & 0 & 0 & 3.611 & 0 & 0 & 0 \\ 2.917 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2.667 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .917 & 0 & 0 & 0 & 0 \\ 2.0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.556 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .889 & 0 & 0 & 0 \end{pmatrix} \tag{75}$$

Lower right 7×7

$$\text{dg}\ (8.611,\ 7.917,\ 7.667,\ 5.917,\ 7.0,\ 6.556,\ 5.889) \tag{76}$$

This matrix is particularly easy to invert. The inverse has the zero elements in exactly the same position as the original matrix and one can obtain these by inverting triangular blocks illustrated by

$$\begin{pmatrix} 19.361 & 0 & 0 & 0 \\ 4.444 & 9.4444 & 0 & 0 \\ 2.917 & 0 & 7.917 & 0 \\ 2.000 & 0 & 0 & 7.000 \end{pmatrix}^{-1} = \begin{pmatrix} .0516 & 0 & 0 & 0 \\ -.0243 & .1059 & 0 & 0 \\ -.0190 & 0 & .1263 & 0 \\ -.0148 & 0 & 0 & .1429 \end{pmatrix}.$$

This results in the solution

(.0057, .2758, .0350, -.3563, .3973, .1970, .0564, -.5925, -.1284, .0345, -.0054, -.3826, .1549, .2613).

This gives $(\tilde{\mathbf{ts}})'\tilde{\mathbf{ts}} = .80728$ with expectation

$$.30426\ \sigma_e^2\ +\ 1.12858\sigma_{ts}^2\ +\ .60987\ \sigma_s^2,$$

and $\tilde{\mathbf{s}}'\tilde{\mathbf{s}} = .20426$ with expectation

$$.08714\ \sigma_e^2\ +\ .35104\sigma_{ts}^2\ +\ .79104\ \sigma_s^2.$$

13 An Algorithm for $\mathbf{R} = \mathbf{R}_* \sigma_e^2$ and Cov $(\mathbf{u}_i, \mathbf{u}_j') = \mathbf{0}$

Simplification of MIVQUE computations result if

$$\mathbf{R} = \mathbf{R}_* \sigma_e^2,\ \text{Var}\ (\mathbf{u}_i)\ =\ \mathbf{G}_{*i}\sigma_e^2\ ;\ \text{and Cov}\ (\mathbf{u}_i, \mathbf{u}_j')\ =\ \mathbf{0}.$$

$\mathbf{R}_*$ and the $\mathbf{G}_{*i}$ are known, and we wish to estimate σ_e^2 and the σ_i^2 . The mixed model equations can be written as

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}_*^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}_*^{-1}\mathbf{Z}_1 & \mathbf{X}'\mathbf{R}_*^{-1}\mathbf{Z}_2 & \ldots \\ \mathbf{Z}_1'\mathbf{R}_*^{-1}\mathbf{X} & \mathbf{Z}_1'\mathbf{R}_*^{-1}\mathbf{Z}_1 + \mathbf{G}_{*1}^{-1}\alpha_1 & \mathbf{Z}_1'\mathbf{R}_*^{-1}\mathbf{Z}_2 & \ldots \\ \mathbf{Z}_2'\mathbf{R}_*^{-1}\mathbf{X} & \mathbf{Z}_2'\mathbf{R}_*^{-1}\mathbf{Z}_1 & \mathbf{Z}_2'\mathbf{R}_*^{-1}\mathbf{Z}_2 + \mathbf{G}_{*2}^{-1}\alpha_2 & \ldots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}}_1 \\ \hat{\mathbf{u}}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}_*^{-1}\mathbf{y} \\ \mathbf{Z}_1'\mathbf{R}_*^{-1}\mathbf{y} \\ \mathbf{Z}_2'\mathbf{R}_*^{-1}\mathbf{y} \\ \vdots \end{pmatrix}. \quad (77)$$

α_i = prior values of σ_e^2/σ_i^2. A set of quadratics equivalent to La Motte's are

$$\hat{\mathbf{e}}'\mathbf{R}_*^{-1}\hat{\mathbf{e}},\ \hat{\mathbf{u}}_i'\mathbf{G}_{*i}^{-1}\hat{\mathbf{u}}_i \quad (\text{i} = 1, 2, \ldots).$$

But because $\hat{\mathbf{e}}'\mathbf{R}_*^{-1}\hat{\mathbf{e}} = \mathbf{y}'\mathbf{R}_*^{-1}\mathbf{y}$ - (soln. vector)' (r.h.s. vector)

$$- \textstyle\sum_i \boldsymbol{\alpha}_i \hat{\mathbf{u}}_i'\mathbf{G}_{*i}^{-1}\hat{\mathbf{u}}_i,$$

an equivalent set of quadratics is

$$\mathbf{y}'\mathbf{R}_*^{-1}\mathbf{y} \ - \ (\text{soln. vector})' \ (\text{r.h.s. vector})$$

and

$$\hat{\mathbf{u}}_i'\mathbf{G}_{*i}^{-1}\hat{\mathbf{u}}_i \quad (\text{i} = 1, 2, \ldots).$$

14 Illustration Of MIVQUE In Multivariate Model

We illustrate several of the principles regarding MIVQUE with the following design

No. of Progeny			
	Sires		
Treatment	1	2	3
1	1	2	0
2	2	2	2

We assume treatments fixed with means t_1, t_2 respectively. The three sires are a random sample of unrelated sires from some population. Sire 1 had one progeny on treatment 1, and 2 different progeny on treatment 2, etc. for the other 2 sires. The sire and error variances are different for the 2 treatments. Further there is a non-zero error covariance between treatments. Thus we have to estimate g_{11} = sire variance for treatment 1, g_{22} = sire variance for treatment 2, g_{12} = sire covariance, r_{11} = error variance for treatment 1, and r_{22} = error variance for treatment 2. We would expect no error covariance if the progeny are from unrelated dams as we shall assume. The record vector ordered by sires in treatments is [2, 3, 5, 7, 5, 9, 6, 8, 3].

We first use the basic La Motte method.

$$\mathbf{V_1} \text{ pertaining to } g_{11} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 & 0 & 0 \\ & & & & & 0 & 0 & 0 & 0 \\ & & & & & & 0 & 0 & 0 \\ & & & & & & & 0 & 0 \\ & & & & & & & & 0 \end{pmatrix}.$$

$$\mathbf{V_2} \text{ pertaining to } g_{12} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ & & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 & 0 & 0 \\ & & & & & 0 & 0 & 0 & 0 \\ & & & & & & 0 & 0 & 0 \\ & & & & & & & 0 & 0 \\ & & & & & & & & 0 \end{pmatrix}.$$

$$\mathbf{V_3} \text{ pertaining to } g_{22} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 1 & 1 & 0 & 0 & 0 & 0 \\ & & & & 1 & 0 & 0 & 0 & 0 \\ & & & & & 1 & 1 & 0 & 0 \\ & & & & & & 1 & 0 & 0 \\ & & & & & & & 1 & 1 \\ & & & & & & & & 1 \end{pmatrix}.$$

$$\begin{aligned} \mathbf{V_4} \text{ pertaining to } r_{11} &= \text{dg } [1,\ 1,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]. \\ \mathbf{V_5} \text{ pertaining to } r_{22} &= \text{dg } [0,\ 0,\ 0,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1]. \end{aligned}$$

Use prior values of $g_{11} = 3$, $g_{12} = 2$, $g_{22} = 4$, $r_{11} = 30$, $r_{22} = 35$. Only the proportionality of these is of concern. Using these values

$$\tilde{\mathbf{V}} = \begin{pmatrix} 33 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ & 33 & 3 & 0 & 0 & 2 & 2 & 0 & 0 \\ & & 33 & 0 & 0 & 2 & 2 & 0 & 0 \\ & & & 39 & 4 & 0 & 0 & 0 & 0 \\ & & & & 39 & 0 & 0 & 0 & 0 \\ & & & & & 39 & 4 & 0 & 0 \\ & & & & & & 39 & 0 & 0 \\ & & & & & & & 39 & 4 \\ & & & & & & & & 39 \end{pmatrix}.$$

Computing $\tilde{\mathbf{V}}^{-1}\mathbf{V}_i\tilde{\mathbf{V}}^{-1}$ we obtain the following values for $\mathbf{Q}_i$ (i=1, …,5). These are in the following table (times .001), only non-zero elements are shown.

Element	$\mathbf{Q}_1$	$\mathbf{Q}_2$	$\mathbf{Q}_3$	$\mathbf{Q}_4$	$\mathbf{Q}_5$
(1,1)	.92872	-.17278	.00804	.92872	.00402
(1,4),(1,5)	-.04320	.71675	-.06630	-.04320	-.03315
(2,2),(2,3)	.78781	-.14657	.00682	.94946	.00341
(2,6),(2,7)	-.07328	.66638	-.06135	-.03664	-.03068
(3,3)	.78781	-.14657	.00682	.94946	.00341
(3,6),(3,7)	-.07328	.66638	-.06135	-.03664	-.03068
(4,4)	.00201	-.06630	.54698	.00201	.68165
(4,5)	.00201	-.06630	.54698	.00201	-.13467
(5,5)	.00201	-.06630	.54698	.00201	.68165
(6,6)	.00682	-.12271	.55219	.00341	.68426
(6,7)	.00682	-.12271	.55219	.00341	-.13207
(7,7)	.00682	-.12271	.55219	.00341	.67858
(8,8),(9,9)	0	0	.54083	0	.67858
(8,9)	0	0	.54083	0	-.13775

We need $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o$, $\boldsymbol{\beta}^o$ being a GLS solution. The GLS equations are

$$\begin{pmatrix} .08661 & -.008057 \\ -.008057 & .140289 \end{pmatrix} \boldsymbol{\beta}^o = \begin{pmatrix} .229319 \\ .862389 \end{pmatrix}.$$

The solution is [3.2368, 6.3333]. Then

$$
\begin{aligned}
\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o &= [-1.2368, -.2368,\ 1.7632,\ .6667,\ -1.3333,\ 2.6667, \\
&\quad -.3333,\ 1.6667,\ -3.3333]' \\
&= [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}]\mathbf{y} \equiv \mathbf{T}'\mathbf{y}.
\end{aligned}
$$

Next we need $\mathbf{T}'\mathbf{V}_i\mathbf{T}$ (i=1, …,5) for the variance of $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o$. These are

Element	$\mathbf{T}'\mathbf{V}_1\mathbf{T}$	$\mathbf{T}'\mathbf{V}_2\mathbf{T}$	$\mathbf{T}'\mathbf{V}_3\mathbf{T}$	$\mathbf{T}'\mathbf{V}_4\mathbf{T}$	$\mathbf{T}'\mathbf{V}_5\mathbf{T}$
(1,1)	.84017	-.03573	.00182	.63013	.00091
(1,2),(1,3)	-.45611	-.00817	.00182	-.34208	.00091
(1,4),(1,5)	0	.64814	.00172	0	.00086
(1,6),(1,7)	0	-.64814	.02928	0	.01464
(1,8),(1,9) (2,8),(2,9) (3,8),(3,9)	0	0	-.03101	0	-.01550
(2,2),(3,3)	.24761	.01940	.00182	.68571	.00091
(2,3)	.24761	.01940	.00182	-.31429	.00091
(2,4),(2,5) (3,4),(3,5)	0	-.35186	.00172	0	.00086
(2,6),(2,7) (3,6),(3,7)	0	.35186	.02928	0	.01464
(4,4),(5,5),(6,6) (7,7),(8,8),(9,9)	0	0	.66667	0	.83333
(4,5),(6,7),(8,9)	0	0	.66667	0	-.16667
(4,6),(4,7),(4,8) (4,9),(5,6),(5,7) (5,8),(5,9),(6,8) (6,9),(7,8),(7,9)	0	0	-.33333	0	-.16667

Taking all combinations of tr $\mathbf{Q}_i\mathbf{T}'\mathbf{V}_j\mathbf{T}$ for the expectation matrix and equating to $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)'\mathbf{Q}_i(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)$ we have these equations to solve.

$$
\begin{pmatrix}
.00156056 & -.00029034 & .00001350 & .00117042 & .000006752 \\
 & .00372880 & -.00034435 & -.00021775 & -.00017218 \\
 & & .00435858 & .00001012 & .00217929 \\
 & & & .00198893 & .00000506 \\
 & & & & .00353862
\end{pmatrix}
\begin{pmatrix} g_{11} \\ g_{12} \\ g_{22} \\ r_{11} \\ r_{22} \end{pmatrix}
=
\begin{pmatrix} .00270080 \\ .00462513 \\ .00423360 \\ .00424783 \\ .01762701 \end{pmatrix}.
$$

This gives the solution [.500, 1.496, -2.083, 2.000, 6.333]. Note that the $\hat{g}_{ij}$ do not fall in the parameter space, but this is not surprising with such a small set of data.

Next we illustrate with quadratics in $\hat{u}_1, \ldots, \hat{u}_5$ and $\hat{e}_1, \ldots, \hat{e}_9$ using the same priors as before.

$$\begin{aligned}
\mathbf{G}^*_{11} &= \text{dg } (1,\ 1,\ 0,\ 0,\ 0), \\
\mathbf{G}^*_{12} &= \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ & 0 & 0 & 1 & 0 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix}, \\
\mathbf{G}^*_{22} &= \text{dg } (0,\ 0,\ 1,\ 1,\ 1), \quad \tilde{\mathbf{R}} = \begin{pmatrix} 30\ \mathbf{I} & \mathbf{0} \\ \mathbf{0} & 35\ \mathbf{I} \end{pmatrix}. \\
\mathbf{R}^*_{11} &= \text{dg } (1,1,1,0,0,0,0,0,0), \\
\mathbf{R}^*_{22} &= \text{dg } (0,0,0,1,1,1,1,1,1), \\
\tilde{\mathbf{G}} &= \begin{pmatrix} 3 & 0 & 2 & 0 & 0 \\ & 3 & 0 & 2 & 0 \\ & & 4 & 0 & 0 \\ & & & 4 & 0 \\ & & & & 4 \end{pmatrix}.
\end{aligned}$$

From these, the 3 matrices of quadratics in $\hat{\mathbf{u}}$ are

$$\begin{pmatrix} .25 & 0 & -.125 & 0 & 0 \\ & .25 & 0 & -.125 & 0 \\ & & .0625 & 0 & 0 \\ & & & .0625 & 0 \\ & & & & 0 \end{pmatrix}, \begin{pmatrix} -.25 & 0 & .25 & 0 & 0 \\ & -.25 & 0 & .25 & 0 \\ & & -.1875 & 0 & 0 \\ & & & -.1875 & 0 \\ & & & & 0 \end{pmatrix},$$

$$\text{and} \begin{pmatrix} .0625 & 0 & -.09375 & 0 & 0 \\ & .0625 & -.09375 & 0 & 0 \\ & & .140625 & 0 & 0 \\ & & & .140625 & 0 \\ & & & & .0625 \end{pmatrix}.$$

Similarly matrices of quadratics in $\hat{\mathbf{e}}$ are

dg (.00111111, .00111111, .00111111, 0, 0, 0, 0, 0, 0),

and

dg (0, 0, 0, 1, 1, 1, 1, 1, 1)*.00081633.

The mixed model coefficient matrix is

$$\begin{pmatrix} .1 & 0 & .03333 & .06667 & 0 & 0 & 0 \\ & .17143 & 0 & 0 & .05714 & .05714 & .05714 \\ & & .53333 & 0 & -.25 & 0 & 0 \\ & & & .56667 & 0 & -.25 & 0 \\ & & & & .43214 & 0 & 0 \\ & & & & & .43214 & 0 \\ & & & & & & .30714 \end{pmatrix}.$$

The right hand side vector is

[.33333, 1.08571, .06667, .26667, .34286, .42857, .31429]′.

The solution is

[3.2368, 6.3333, -.1344, .2119, -.1218, .2769, -.1550].

Let the last 5 rows of the inverse of the matrix above = $\mathbf{C}_u$. Then

$$\begin{aligned} Var(\hat{\mathbf{u}}) &= \mathbf{C}_u\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}(\mathbf{G}^*_{11}g_{11} + \mathbf{G}^*_{12}g_{12} + \mathbf{G}^*_{22}g_{22})\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{W}\mathbf{C}_u' \\ &\quad + \mathbf{C}_u\mathbf{W}'\tilde{\mathbf{R}}^{-1}(\mathbf{R}^*_{11}r_{11} + \mathbf{R}^*_{22}r_{22})\tilde{\mathbf{R}}^{-1}\mathbf{W}\mathbf{C}_u' \end{aligned}$$

$$= \begin{pmatrix} .006179 & -.006179 & .003574 & -.003574 & 0 \\ & .006179 & -.003574 & .003574 & 0 \\ & & .002068 & -.002068 & 0 \\ & & & .002068 & 0 \\ & & & & 0 \end{pmatrix} g_{11}$$

$$+ \begin{pmatrix} .009191 & -.009191 & .012667 & -.012667 & 0 \\ & .009191 & -.012667 & .012667 & 0 \\ & & .011580 & -.011580 & 0 \\ & & & .011580 & 0 \\ & & & & 0 \end{pmatrix} g_{12}$$

$$+ \begin{pmatrix} .004860 & -.001976 & .010329 & -.004560 & -.005769 \\ & .004860 & -.004560 & .010329 & -.005769 \\ & & .021980 & -.010443 & -.011538 \\ & & & .021980 & -.011538 \\ & & & & .023076 \end{pmatrix} g_{22}$$

$$+ \begin{pmatrix} .004634 & -.004634 & .002681 & -.002681 & 0 \\ & .004634 & -.002681 & .002681 & 0 \\ & & .001551 & -.001551 & 0 \\ & & & .001551 & 0 \\ & & & & 0 \end{pmatrix} r_{11}$$

$$
+ \begin{pmatrix}
.002430 & -.000988 & .005164 & -.002280 & -.002884 \\
 & .002430 & -.002280 & .005164 & -.002884 \\
 & & .010990 & -.005221 & -.005769 \\
 & & & .010990 & -.005769 \\
 & & & & .011538
\end{pmatrix} r_{22}.
$$

$$
\mathbf{e}' = [-1.1024,\ -4488,\ 1.5512,\ .7885,\ -1.2115,\ 2.3898,\ -.6102,\ 1.8217,\ -3.1783].
$$

Let $\mathbf{C}$ be a g-inverse of the mixed model coefficient matrix, and $\mathbf{T} = \mathbf{I} - \mathbf{WCW}'\tilde{\mathbf{R}}^{-1}$. Then

$$
Var(\hat{\mathbf{e}}) = \mathbf{T}(\mathbf{ZG}^*_{11}\mathbf{Z}'g_{11} + \mathbf{ZG}^*_{12}\mathbf{Z}'g_{12} + \mathbf{ZG}_{22}\mathbf{Z}'g_{22} + \mathbf{R}^*_{11}r_{11} + \mathbf{R}^*_{22}r_{22})\mathbf{T}'
$$

$$
= \begin{pmatrix}
.702 & -.351 & -.351 & -.038 & -.038 & .031 & .038 & 0 & 0 \\
 & .176 & .176 & .019 & .019 & -.019 & -.019 & 0 & 0 \\
 & & .176 & .019 & .019 & -.019 & -.019 & 0 & 0 \\
 & & & .002 & .002 & -.002 & -.002 & 0 & 0 \\
 & & & & .002 & -.002 & -.002 & 0 & 0 \\
 & & & & & .002 & .002 & 0 & 0 \\
 & & & & & & .002 & 0 & 0 \\
 & & & & & & & 0 & 0 \\
 & & & & & & & & 0
\end{pmatrix} g_{11}
$$

$$
+ \begin{pmatrix}
-.131 & .065 & .065 & .489 & .489 & -.489 & -.489 & 0 & 0 \\
 & -.033 & -.033 & -.245 & -.245 & .245 & .245 & 0 & 0 \\
 & & -.033 & -.245 & -.245 & .245 & .245 & 0 & 0 \\
 & & & -.053 & -.053 & .053 & .053 & 0 & 0 \\
 & & & & -.053 & .053 & .053 & 0 & 0 \\
 & & & & & -.053 & -.053 & 0 & 0 \\
 & & & & & & -.053 & 0 & 0 \\
 & & & & & & & 0 & 0 \\
 & & & & & & & & 0
\end{pmatrix} g_{12}
$$

$$
+ \begin{pmatrix}
.006 & -.003 & -.003 & -.045 & -.045 & .045 & .045 & 0 & 0 \\
 & .002 & .002 & .023 & .023 & -.023 & -.023 & 0 & 0 \\
 & & .002 & .023 & .023 & -.023 & -.023 & 0 & 0 \\
 & & & .447 & .447 & -.226 & -.226 & -.221 & -.221 \\
 & & & & .447 & -.226 & -.226 & -.221 & -.221 \\
 & & & & & .447 & .447 & -.221 & -.221 \\
 & & & & & & .447 & -.221 & -.221 \\
 & & & & & & & .442 & .442 \\
 & & & & & & & & .442
\end{pmatrix} g_{22}
$$

$$
+ \begin{pmatrix}
.527 & -.263 & -.263 & -.029 & -.029 & .029 & .029 & 0 & 0 \\
 & .632 & -.369 & .014 & .014 & -.014 & -.014 & 0 & 0 \\
 & & .632 & .014 & .014 & -.014 & -.014 & 0 & 0 \\
 & & & .002 & .002 & -.002 & -.002 & 0 & 0 \\
 & & & & .002 & -.002 & -.002 & 0 & 0 \\
 & & & & & .002 & .002 & 0 & 0 \\
 & & & & & & .002 & 0 & 0 \\
 & & & & & & & 0 & 0 \\
 & & & & & & & & 0
\end{pmatrix} r_{11}
$$

$$
+ \begin{pmatrix}
.003 & -.002 & -.002 & -.023 & -.023 & .023 & .023 & 0 & 0 \\
 & .001 & .001 & .011 & .011 & -.011 & -.011 & 0 & 0 \\
 & & .001 & .011 & .011 & -.011 & -.011 & 0 & 0 \\
 & & & .723 & -.277 & -.113 & -.113 & -.110 & -.110 \\
 & & & & .723 & -.113 & -.113 & -.110 & -.110 \\
 & & & & & .723 & -.277 & -.110 & -.110 \\
 & & & & & & .723 & -.110 & -.110 \\
 & & & & & & & .721 & -.279 \\
 & & & & & & & & .721
\end{pmatrix} r_{22}.
$$

Taking the traces of products of $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$ with $Var(\hat{\mathbf{u}})$ and of $\mathbf{Q}_4$, $\mathbf{Q}_5$ with $Var(\hat{\mathbf{e}})$ we get the same expectations as in the La Motte method. Also the quadratics in $\hat{\mathbf{u}}$ and $\hat{\mathbf{e}}$ are the same as the La Motte quadratics in $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)$.

If $\hat{u}_6$ is included, the same quadratics and expectations are obtained. If $\hat{u}_6$ is included and we compute the following quadratics in $\hat{\mathbf{u}}$.

$$
\hat{\mathbf{u}}' \text{ dg } (1\ 1\ 1\ 0\ 0\ 0)\ \hat{\mathbf{u}}, \quad \hat{\mathbf{u}}' \begin{pmatrix}
0 & 0 & 0 & 1 & 0 & 0 \\
 & 0 & 0 & 0 & 1 & 0 \\
 & & 0 & 0 & 0 & 1 \\
 & & & 0 & 0 & 0 \\
 & & & & 0 & 0 \\
 & & & & & 0
\end{pmatrix} \hat{\mathbf{u}},
$$

and $\hat{\mathbf{u}}'$ dg (0, 0, 0, 1, 1, 1) $\hat{\mathbf{u}}$ and equate to expectations we obtain exactly the same estimates as in the other three methods. We also could have computed the following quadratics in $\hat{\mathbf{e}}$ rather than the ones used, namely

$$
\hat{\mathbf{e}}' \text{ dg } (1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0)\hat{\mathbf{e}} \text{ and } \hat{\mathbf{e}}' \text{ dg } (0\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 1)\ \hat{\mathbf{e}}.
$$

Also we could have computed an approximate MIVQUE by estimating r_{11} from within sires in treatment 1 and r_{22} from within sires in treatment 2.

In most problems the "error" variances and covariances contribute markedly to computational labor. If no simplification of this computation can be effected, the La Motte quadratics might be used in place of quadratics in $\mathbf{e}$. Remember, however, that $\tilde{\mathbf{V}}^{-1}$ is usually a large matrix impossible to compute by conventional methods. But if $\tilde{\mathbf{R}}^{-1}$, $\tilde{\mathbf{G}}^{-1}$ and $(\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} + \tilde{\mathbf{G}}^{-1})^{-1}$ are relatively easy to compute one can employ the results,

$$\tilde{\mathbf{V}}^{-1} \;=\; \tilde{\mathbf{R}}^{-1} - \tilde{\mathbf{R}}^{-1}\mathbf{Z}(\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} + \tilde{\mathbf{G}}^{-1})^{-1}\mathbf{Z}\tilde{\mathbf{R}}^{-1}.$$

As already discussed, in most genetic problems simple quadratics in $\hat{\mathbf{u}}$ can be derived usually of the form

$$\hat{\mathbf{u}}_i^{'}\hat{\mathbf{u}}_j \text{ or } \hat{\mathbf{u}}_i\mathbf{A}^{-1}\hat{\mathbf{u}}_j.$$

Then these might be used with the La Motte ones for the r_{ij} rather than quadratics in $\hat{\mathbf{e}}$ for the r_{ij}. The La Motte quadratics are in $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o)$, the variance of $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o$ being

$$[\mathbf{I} - \mathbf{X}(\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{X}'\tilde{\mathbf{V}}^{-1}]\mathbf{V}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{X}'\tilde{\mathbf{V}}^{-1}]'.$$

Remember that $\tilde{\mathbf{V}} \neq \mathbf{V}$ in general, and $\mathbf{V}$ should be written in terms of g_{ij}, r_{ij} for purposes of taking expectations.

15 Other Types Of MIVQUE

The MIVQUE estimators of this chapter are translation invariant and unbiased. La Motte also presented other estimators including not translation invariant biased estimators and translation invariant biased estimators.

15.1 Not translation invariant and biased

The LaMotte estimator of this type is

$$\hat{\theta}_i \;=\; \tilde{\theta}_i \; (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})'\tilde{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})/(\mathrm{n}+2),$$

where $\tilde{\theta}$, $\tilde{\boldsymbol{\beta}}$, and $\tilde{\mathbf{V}}$ are priors. This can also be computed as

$$\hat{\theta}_i \;=\; \tilde{\theta}_i \; [(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})'\tilde{\mathbf{R}}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}) - \hat{\mathbf{u}}'\mathbf{Z}'\tilde{\mathbf{R}}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})]/(n+2),$$

where

$$\hat{\mathbf{u}} = (\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} + \tilde{\mathbf{G}}^{-1})^{-1}\mathbf{Z}'\tilde{\mathbf{R}}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}).$$

The lower bound on MSE of $\hat{\theta}_i$ is $2\tilde{\theta}_i^2/(\mathrm{n}+2)$, when $\tilde{\mathbf{V}}$, $\tilde{\boldsymbol{\beta}}$ are used as priors.

15.2 Translation invariant and biased

An estimator of this type is

$$\hat{\theta}_i \;=\; \tilde{\theta}_i\;(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}^o)'\tilde{\mathbf{V}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}^o)/(n-r+2).$$

This can be written as

$$\tilde{\theta}_i\;[\mathbf{y}'\tilde{\mathbf{R}}^{-1}\mathbf{y}-(\boldsymbol{\beta}^o)'\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y}-\hat{\mathbf{u}}'\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{y}]/(\mathrm{n}-\mathrm{r}+2).$$

$\boldsymbol{\beta}^o$ and $\hat{\mathbf{u}}$ are solution to mixed model equations with $\mathbf{G}=\tilde{\mathbf{G}}$ and $\mathbf{R}=\tilde{\mathbf{R}}$. The lower bound on MSE of $\hat{\theta}_i$ is $2\tilde{\theta}_i^2/(n-r+2)$ when $\tilde{\mathbf{V}}$ is used as the prior for $\mathbf{V}$. The lower bound on $\hat{\theta}_i$ for the translation invariant, unbiased MIVQUE is $2d_i$, when d_i is the i^{th} diagonal of $\mathbf{G}_0^{-1}$ and the ij^{th} element of $\mathbf{G}_0$ is $tr\mathbf{W}_0\mathbf{V}_i^*\mathbf{W}_0\mathbf{V}_i^*$ for

$$\mathbf{W}_0=\tilde{\mathbf{V}}^{-1}-\tilde{\mathbf{V}}^{-1}\mathbf{X}(\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{X}'\tilde{\mathbf{V}}^{-1}.$$

The estimators of sections 15.1 and 15.2 have the peculiar property that $\hat{\theta}_i/\hat{\theta}_j \;=\; \tilde{\theta}_i/\tilde{\theta}_j$. Thus the ratios of estimators are exactly proportional to the ratios of the priors used in the solution.

16 Expectations Of Quadratics In $\hat{\alpha}$

Let some g-inverse of (5.51) with priors on $\mathbf{G}$ and $\mathbf{R}$ be

$$\begin{pmatrix}\mathbf{C}_{11} & \mathbf{C}_{12}\\ \mathbf{C}_{12}' & \mathbf{C}_{22}\end{pmatrix}\equiv\begin{pmatrix}\mathbf{C}_1\\ \mathbf{C}_2\end{pmatrix}$$

Then $\hat{\alpha}=\mathbf{C}_2\mathbf{r}$, where $\mathbf{r}$ is the right hand vector of (5.51), and

$$\begin{aligned}E(\hat{\boldsymbol{\alpha}}'\mathbf{Q}\hat{\boldsymbol{\alpha}}) &= tr\mathbf{Q}\;Var(\hat{\boldsymbol{\alpha}})\\ &= tr\mathbf{Q}\mathbf{C}_2[Var(\mathbf{r})]\mathbf{C}_2'.\\ Var(\mathbf{r}) &= \begin{pmatrix}\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}\\ \tilde{\mathbf{G}}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\end{pmatrix}\mathbf{G}\;(\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X}\quad \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}\tilde{\mathbf{G}})\\ &\quad+\begin{pmatrix}\mathbf{X}'\tilde{\mathbf{R}}^{-1}\\ \tilde{\mathbf{G}}\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\end{pmatrix}\mathbf{R}\;(\tilde{\mathbf{R}}^{-1}\mathbf{X}\quad \tilde{\mathbf{R}}^{-1}\mathbf{Z}\tilde{\mathbf{G}}).\end{aligned} \tag{78}$$

When $\mathbf{R}=\mathbf{I}\sigma_e^2$ and $\mathbf{G}=\mathbf{G}_*\sigma_e^2$, $\hat{\boldsymbol{\alpha}}$ can be obtained from the solution to

$$\begin{pmatrix}\mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}\mathbf{G}_*\\ \mathbf{G}_*\mathbf{Z}'\mathbf{X} & \mathbf{G}_*\mathbf{Z}'\mathbf{Z}\mathbf{G}_*+\mathbf{G}_*\end{pmatrix}\begin{pmatrix}\boldsymbol{\beta}^o\\ \hat{\boldsymbol{\alpha}}\end{pmatrix}=\begin{pmatrix}\mathbf{X}'\mathbf{y}\\ \mathbf{G}_*\mathbf{Z}'\mathbf{y}\end{pmatrix}. \tag{79}$$

In this case $\mathbf{C}_2$ is the last g rows of a g-inverse of (79).

$$\begin{aligned} Var(\mathbf{r}) &= \begin{pmatrix} \mathbf{X}'\mathbf{Z} \\ \tilde{\mathbf{G}}_*\mathbf{Z}'\mathbf{Z} \end{pmatrix} \mathbf{G} \ (\mathbf{Z}'\mathbf{X} \quad \mathbf{Z}'\mathbf{Z}\tilde{\mathbf{G}}_*)\sigma_e^2 \\ &+ \begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}\tilde{\mathbf{G}}_* \\ \tilde{\mathbf{G}}_*\mathbf{Z}'\mathbf{X} & \tilde{\mathbf{G}}_*\mathbf{Z}'\mathbf{Z}\tilde{\mathbf{G}}_* \end{pmatrix} \sigma_e^2. \end{aligned} \tag{80}$$

Chapter 12
REML and ML Estimation

C. R. Henderson

1984 - Guelph

1 Iterative MIVQUE

The restricted maximum likelihood estimator (REML) of Patterson and Thompson (1971) can be obtained by iterating on MIVQUE, Harville (1977). Let the prior value of $\mathbf{g}$ and $\mathbf{r}$ be denoted by $\mathbf{g}_0$ and $\mathbf{r}_0$. Then compute MIVQUE and denote the estimates by $\mathbf{g}_1$ and $\mathbf{r}_1$. Next use these as priors in MIVQUE and denote the estimates $\mathbf{g}_2$ and $\mathbf{r}_2$. Continue this process until $\mathbf{g}_{k+1} = \mathbf{g}_k$ and $\mathbf{r}_{k+1} = \mathbf{r}_k$. Several problems must be recognized.

1. Convergence may be prohibitively slow or may not occur at all.
2. If convergence does occur, it may be to a local rather than a global maximum.
3. If convergence does occur, $\mathbf{g}$ and $\mathbf{r}$ may not fall in the parameter space.

We can check the last by noting that both $\mathbf{G}_k$ and $\mathbf{R}_k$ must be positive definite or positive semidefinite at convergence, where $\mathbf{G}_k$ and $\mathbf{R}_k$ are

$$\begin{pmatrix} \hat{g}_{11} & \hat{g}_{12} & \cdots \\ \hat{g}_{12} & \hat{g}_{22} & \cdots \\ . & . & \\ . & . & \end{pmatrix} \text{ and } \begin{pmatrix} \hat{r}_{11} & \hat{r}_{12} & \cdots \\ \hat{r}_{12} & \hat{r}_{22} & \cdots \\ . & . & \\ . & . & \end{pmatrix}.$$

For positive definitness or positive semidefiniteness all eigenvalues of $\mathbf{G}_k$ and $\mathbf{R}_k$ must be non-negative. Writing a computer program that will guarantee this is not trivial. One possibility is to check at each round, and if the requirement is not met, new starting values are chosen. Another possibility is to alter some elements of $\hat{\mathbf{G}}$ or $\hat{\mathbf{R}}$ at each round in which either $\hat{\mathbf{G}}$ or $\hat{\mathbf{R}}$ is not a valid estimator. (LRS note: Other possibilities are bending in which eigenvalues are modified to be positive and the covariance matrix is reformed using the new eigenvalues with the eigenvectors.)

Quadratic, unbiased estimators may lead to solutions not in the parameter space. This is the price to pay for unbiasedness. If the estimates are modified to force them into the parameter space, unbiasedness no longer can be claimed. What should be done in practice? If the purpose of estimation is to accumulate evidence on parameters with other research, one should report the invalid estimates, for otherwise the average of many estimates will be biased. On the other hand, if the results of the analysis are to be used immediately, for example, in BLUP, the estimate should be required to fall in the parameter space. It would seem illogical for example, to reduce the diagonals of $\hat{\mathbf{u}}$ in mixed model equations because the diagonals of $\bar{\mathbf{G}}^{-1}$ are negative.

2 An Alternative Algorithm For REML

An alternative algorithm for REML that is considerably easier per round of iteration than iterative MIVQUE will now be described. There is, however, some evidence that convergence is slower than in the iterative MIVQUE algorithm. The method is based on the following principle. At each round of iteration find the expectations of the quadratics under the pretense that the current solutions to $\hat{\mathbf{g}}$ and $\hat{\mathbf{r}}$ are equal to $\mathbf{g}$ and $\mathbf{r}$. This leads to much simpler expectations. Note, however, that the first iterate under this algorithm is not MIVQUE. This is the EM (expectation maximization) algorithm, Dempster et al. (1977).

From Henderson (1975a), when $\bar{\mathbf{g}} = \mathbf{g}$ and $\bar{\mathbf{r}} = \mathbf{r}$

$$\begin{aligned} Var(\hat{\mathbf{u}}) &= \mathbf{G} - \mathbf{C}_{11}. & (1) \\ Var(\hat{\mathbf{e}}) &= \mathbf{R} - \mathbf{WCW}' = \mathbf{R} - \mathbf{S}. & (2) \end{aligned}$$

The proof of this is

$$Var(\hat{\mathbf{e}}) = \text{Cov}\ (\hat{\mathbf{e}}, \mathbf{e}') = Cov[(\mathbf{y} - \mathbf{WCW}'\mathbf{R}^{-1}\mathbf{y}), \mathbf{e}'] = \mathbf{R} - \mathbf{WCW}'. \quad (3)$$

A g-inverse of the mixed model coefficient matrix is

$$\begin{pmatrix} \mathbf{C}_{00} & \mathbf{C}_{01} \\ \mathbf{C}_{10} & \mathbf{C}_{11} \end{pmatrix} = \mathbf{C}.$$

Note that if we proceed as in Section 11.5 we will need only diagonal blocks of $\mathbf{WCW}'$ corresponding to the diagonal blocks of $\mathbf{R}$.

$$Var(\hat{\mathbf{u}}) = \sum_i \sum_{j \geq i} \mathbf{G}^*_{ij} g_{ij} - \mathbf{C}_{11} \quad (4)$$

See Chapter 11, Section 3 for definition of $\mathbf{G}^*$. Let $\mathbf{C}$, $\mathbf{S}$, $\hat{\mathbf{u}}$, and $\hat{\mathbf{e}}$ be the values computed for the k^{th} round of iteration. Then solve in the k+l round of iteration for values of $\mathbf{g}, \mathbf{r}$ from the following set of equations.

$$\begin{aligned} tr\mathbf{Q}_1\mathbf{G} &= \hat{\mathbf{u}}'\mathbf{Q}_1\hat{\mathbf{u}} + tr\mathbf{Q}_1\mathbf{C}_{11} \\ &\vdots \end{aligned}$$

$$
\begin{aligned}
tr\mathbf{Q}_b\mathbf{G} &= \hat{\mathbf{u}}'\mathbf{Q}_b\hat{\mathbf{u}}' \; + \; tr\mathbf{Q}_b\mathbf{C}_{11} \\
tr\mathbf{Q}_{b+1}\mathbf{R} &= \hat{\mathbf{e}}'\mathbf{Q}_{b+1}\hat{\mathbf{e}} \; + \; tr\mathbf{Q}_{b+1}\mathbf{S} \\
&\vdots \\
tr\mathbf{Q}_c\mathbf{R} &= \hat{\mathbf{e}}'\mathbf{Q}_c\hat{\mathbf{e}} \; + \; tr\mathbf{Q}_c\mathbf{S}.
\end{aligned} \tag{5}
$$

Note that at each round a set of equations must be solved for all elements of $\mathbf{g}$, and another set of all elements of $\mathbf{r}$. In some cases, however, $\mathbf{Q}$'s can be found such that only one element of g_{ij} (or r_{ij}) appears on the left hand side of each equation of (5). Note also that if $\bar{\mathbf{G}}^{-1}$ appears in a $\mathbf{Q}_i$ the value of $\mathbf{Q}_i$ changes at each round of iteration. The same applies to $\bar{\mathbf{R}}^{-1}$ appearing in $\mathbf{Q}_i$ for $\hat{\mathbf{e}}$. Consequently it is desirable to find $\mathbf{Q}_i$ that isolate a single g_{ij} or r_{ij} in each left hand side of (5) and that are not dependent upon $\bar{\mathbf{G}}$ and $\bar{\mathbf{R}}$. This can be done for the g_{ij} in all genetic problems with which I am familiar.

The second algorithm for REML appears to have the property that if positive definite $\mathbf{G}$ and $\mathbf{R}$ are chosen for starting values, convergence, if it occurs, will always be to positive definite $\hat{\mathbf{G}}$ and $\hat{\mathbf{R}}$. This suggestion has been made by Smith (1982).

3 ML Estimation

A slight change in the second algorithm for REML, presented in Section 2 results in an EM type ML algorithm. In place of $\mathbf{C}_{11}$ substitute $(\mathbf{Z}'\bar{\mathbf{R}}^{-1}\mathbf{Z}+\mathbf{G}^{-1})^{-1}$. In place of $\mathbf{WCW}'$ substitute $\mathbf{Z}(\mathbf{Z}'\bar{\mathbf{R}}^{-1}\mathbf{Z}+\bar{\mathbf{G}}^{-1})^{-1}\mathbf{Z}'$. Using a result reported by Laird and Ware (1982) substituting ML estimates of $\mathbf{G}$ and $\mathbf{R}$ for the corresponding parameters in the mixed model equations yields empirical Bayes estimates of $\mathbf{u}$. As stated in Chapter 8 the $\hat{\mathbf{u}}$ are also ML estimates of the conditional means of $\mathbf{u}$.

If one wishes to use the LaMotte type quadratics for REML and ML, the procedure is as follows. For REML iterate on

$$
tr\mathbf{Q}_j \sum_i \mathbf{V}_i^* \, \theta_i \; = \; (\mathbf{y}-\mathbf{X}\boldsymbol{\beta}^o)'\mathbf{Q}_j(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}^o) \; + \; tr\mathbf{Q}_j \;\; \mathbf{X}(\mathbf{X}'\bar{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{X}'.
$$

$\mathbf{Q}_j$ are the quadratics computed by the LaMotte method described in Chapter 11. Also this chapter describes the $\mathbf{V}_i^*$. Further, $\boldsymbol{\beta}^o$ is a GLS solution.

ML is computed in the same way as REML except that

$$
tr\mathbf{Q}_j \;\; \mathbf{X}(\mathbf{X}'\bar{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{X}' \text{ is deleted.}
$$

The EM type algorithm converges slowly if the maximizing value of one or more parameters is near the boundary of the parameter space, eg. $\hat{\sigma}_i^2 \rightarrow 0$. The result of Hartley and Rao (1967) can be derived by this general EM algorithm.

4 Approximate REML

REML by either iterative MIVQUE or by the method of Section 2 is costly because every round of iteration requires a g-inverse of the mixed model coefficient matrix. The cost could be reduced markedly by iterating on approximate MIVQUE of Section 11.7. Further simplification would result in the $\mathbf{R} = \mathbf{I}\sigma_e^2$ case by using the residual mean square of OLS as the estimate of σ_e^2. Another possibility is to use the method of Section 2, but with an approximate g-inverse and solution at each round of iteration. The properties of such an estimation are unknown.

5 A Simple Result For Expectation Of Residual Sum Of Squares

Section 11.13 shows that in a model with $\mathbf{R} = \mathbf{R}_* \sigma_e^2$, $Var(\mathbf{u}_i) = \mathbf{G}_{*i}\sigma_e^2$, and $Cov(\mathbf{u}_i, \mathbf{u}_j') = \mathbf{0}$, one of the quadratics that can be used is

$$\mathbf{y}'\mathbf{R}_*^{-1}\mathbf{y} \; - \; \text{(soln. vector)}' \text{ (r.h.s. vector)} \tag{6}$$

with equations written as (77) in Chapter 11. $\mathbf{R}_*$ and $\mathbf{G}_{*i}$ are known. Then if $\alpha = \sigma_e^2/\sigma_i^2$, as is defined in taking expectations for the computations of Section 2, the expectation of (6) is

$$[n - \text{rank } (\mathbf{X})]\sigma_e^2. \tag{7}$$

6 Biased Estimation With Few Iterations

What if one has only limited data to estimate a set of variances and covariances, but prior estimates of these parameters have utilized much more data? In that case it might be logical to iterate only a few rounds using the EM type algorithm for REML or ML. Then the estimates would be a compromise between the priors and those that would be obtained by iterating to convergence. This is similar to the consequences of Bayesian estimation. If the priors are good, it is likely that the MSE will be smaller than those for ML or REML. A small simulation trial illustrates this. The model assumed was

$$\begin{aligned} y_{ij} &= \beta \mathbf{X}_{ij} + a_i + e_{ij}. \\ \mathbf{X}' &= (3, 2, 5, 1, 3, 2, 3, 6, 7, 2, 3, 5, 3, 2). \\ n_i &= (3, 2, 4, 5). \\ Var(\mathbf{e}) &= 4\,\mathbf{I}, \\ Var(\mathbf{a}) &= \mathbf{I}, \end{aligned}$$

$$Cov(\mathbf{a}, \mathbf{e}') = \mathbf{0}.$$

5000 samples were generated under this model and EM type REML was carried out with starting values of $\sigma_e^2/\sigma_a^2 = 4, \ .5,$ and 100. Average values and MSE were computed for rounds $1, 2, ..., 9$ of iteration.

	Starting Value $\sigma_e^2/\sigma_a^2 = 4$					
	$\hat{\sigma}_e^2$		$\hat{\sigma}_a^2$		$\hat{\sigma}_e^2/\hat{\sigma}_a^2$	
Rounds	Av.	MSE	Av.	MSE	Av.	MSE
1	3.98	2.37	1.00	.22	4.18	.81
2	3.93	2.31	1.04	.40	4.44	2.97
3	3.88	2.31	1.10	.70	4.75	6.26
4	3.83	2.34	1.16	1.08	5.09	10.60
5	3.79	2.39	1.22	1.48	5.47	15.97
6	3.77	2.44	1.27	1.86	5.86	22.40
7	3.75	2.48	1.31	2.18	6.26	29.90
8	3.74	2.51	1.34	2.43	6.67	38.49
9	3.73	2.53	1.35	2.62	7.09	48.17

In this case only one round appears to be best for estimating σ_e^2/σ_a^2 .

	Starting Value $\sigma_e^2/\sigma_a^2 = .5$					
	$\hat{\sigma}_e^2$		$\hat{\sigma}_a^2$		$\hat{\sigma}_e^2/\hat{\sigma}_a^2$	
Rounds	Av.	MSE	Av.	MSE	Av.	MSE
1	3.14	2.42	3.21	7.27	1.08	8.70
2	3.30	2.37	2.53	4.79	1.66	6.27
3	3.40	2.38	2.20	4.05	2.22	5.11
4	3.46	2.41	2.01	3.77	2.75	5.23
5	3.51	2.43	1.88	3.64	3.28	6.60
6	3.55	2.45	1.79	3.58	3.78	9.20
7	3.57	2.47	1.73	3.54	4.28	12.99
8	3.60	2.49	1.67	3.51	4.76	17.97
9	3.61	2.51	1.63	3.50	5.23	24.11

	Starting Value $\hat{\sigma}_e^2/\hat{\sigma}_a^2 = 100$					
	$\hat{\sigma}_e^2$		$\hat{\sigma}_a^2$		$\hat{\sigma}_e^2/\hat{\sigma}_a^2$	
Rounds	Av.	MSE	Av.	MSE	Av.	MSE
1	4.76	4.40	.05	.91	.99	9011
2	4.76	4.39	.05	.90	.98	8818
3	4.76	4.38	.05	.90	.97	8638
4	4.75	4.37	.05	.90	.96	8470
5	4.75	4.35	.05	.90	.95	8315
6	4.75	4.34	.05	.90	.94	8172
7	4.75	4.32	.05	.90	.92	8042
8	4.74	4.31	.06	.89	.91	7923
9	4.74	4.28	.06	.89	.90	7816

Convergence with this very high starting value of σ_e^2/σ_a^2 relative to the true value of 4 is very slow but the estimates were improving with each round.

7 The Problem Of Finding Permissible Estimates

Statisticians and users of statistics have for many years discussed the problem of "estimates" of variances that are less than zero. Most commonly employed methods of estimation are quadratic, unbiased, and translation invariant, for example ANOVA estimators, Methods 1,2, and 3 of Henderson, and MIVQUE. In all of these methods there is a positive probability that a solution to one or more variances will be negative. Strictly speaking, these are not really estimates if we define, as some do, that an estimate must lie in the parameter space. But, in general, we cannot obtain unbiasedness unless we are prepared to accept such solutions. The argument used is that such "estimates" should be reported because eventually there may be other estimates of the same parameters obtained by unbiased methods, and then these can be averaged to obtain better unbiased estimates.

Other workers obtain truncated estimates. That is, given estimates $\hat{\sigma}_1^2, ..., \ \hat{\sigma}_q^2$, with say $\hat{\sigma}_q^2 < 0$, the estimates are taken as $\hat{\sigma}_1^2, ..., \ \hat{\sigma}_{q-1}^2, \ 0$. Still others revise the model so that the offending variable is deleted from the model, and new estimates are then obtained of the remaining variances. If these all turn out to be non-negative, the process stops. If some new estimate turns negative, then that variance is dropped from the model and a new set of estimates obtained.

These truncated estimators can no longer be defined as unbiased. Verdooren (1980) in an interesting review of variance component estimation uses the terms "permissible" and "impermissible" to characterize estimators. Permissible estimators are those in which the solution is guaranteed to fall in the parameter space, that is all estimates of variances are non-negative. Impermissible estimators are those in which there is a probability greater than 0 that the solution will be negative.

If one insists on permissible estimators, why not then use some method that guarantees this property while at the same time invoking, if possible, other desirable properties of estimators such as consistency, minimum mean squared error, etc.? Of course unbiasedness cannot, in general, be invoked. For example, an algorithm for ML, Henderson (1973), guarantees a permissible estimator provided convergence occurs. A simple extension of this method due to Harville (1977), yields permissible estimators by REML. The problem of permissible estimators is especially acute in multiple trait models. For example, in a two trait phenotypic model say

$$\begin{aligned} y_{ij} &= \mu_i + e_{1j} \\ y_{2j} &= \mu_2 + e_{2j} \end{aligned}$$

we need to estimate

$$Var \begin{pmatrix} e_{1j} \\ e_{2j} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix}. \ \ c_{11} \geq 0, \ c_{22} \geq 0, \ c_{11}c_{22} \geq c_{12}^2.$$

The last of these criteria insures that the estimated correlation between e_{1j} and e_{2j} falls in the range -1 to 1. The literature reporting genetic correlation estimates contains many cases in which the criteria are not met, this in spite of probable lack of reporting of many other sets of computations with such results. The problem is particularly difficult when there are more than 2 variates. Now it is not sufficient for all estimates of variances to be non- negative and all pairs of estimated correlations to fall in the proper range. The requirement rather is that the estimated variance-covariance matrix be either positive definite or at worst positive semi-definite. A condition guaranteeing this is that all latent roots (eigenvalues) be positive for positive definiteness or be non-negative for positive semidefiteness. Most computing centers have available a good subroutine for computing eigenvalues. We illustrate with a 3×3 matrix in which all correlations are permissible, but the matrix is negative definite.

$$\begin{pmatrix} 3 & -3 & 4 \\ -3 & 4 & 4 \\ 4 & 4 & 6 \end{pmatrix}$$

The eigenvalues for this matrix are (9.563, 6.496, -3.059), proving that the matrix is negative definite. If this matrix represented an estimated $\mathbf{G}$ for use in mixed model equations, one would add $\mathbf{G}^{-1}$ to an appropriate submatrix, of OLS equations, but

$$\mathbf{G}^{-1} = \begin{pmatrix} -.042 & -.139 & .147 \\ & -.011 & .126 \\ & & -.016 \end{pmatrix},$$

so one would add negative quantities to the diagonal elements, and this would make no sense. If the purpose of variance-covariance estimation is to use the estimates in setting up mixed model equations, it is essential that permissible estimators be used.

Another difficult problem arises when variance estimates are to be used in estimating h^2. For example, in a sire model, an estimate of h^2 often used is

$$\hat{h}^2 = 4 \, \hat{\sigma}_s^2/(\hat{\sigma}_s^2 + \hat{\sigma}_e^2).$$

By definition $0 < h^2 < 1$, the requirement that $\hat{\sigma}_s^2 > 0$ and $\hat{\sigma}_e^2 > 0$ does not insure that $\hat{h}^2$ is permissible. For this to be true the permissible range of $\hat{\sigma}_s^2/\hat{\sigma}_e^2$ is 0 to 3^{-1}. This would suggest using an estimation method that guarantees that the estimated ratio falls in the appropriate range.

In the multivariate case a method might be derived along these lines. Let some translation invariant unbiased estimator be the solution to

$$\mathbf{C}\hat{\mathbf{v}} = \mathbf{q},$$

where $\mathbf{q}$ is a set of quadratics and $\mathbf{Cv}$ is $E(\mathbf{q})$. Then solve these equations subject to a set of inequalities that forces $\hat{\mathbf{v}}$ to fall in the parameter space, as a minimum, all eigenvalues ≥ 0 where $\hat{\mathbf{v}}$ comprises the elements of the variance-covariance matrix.

8 Method For Singular G

When $\mathbf{G}$ is singular we can use a method for EM type REML that is similar to MIVQUE in Section 11.16. We iterate on $\hat{\boldsymbol{\alpha}}'\mathbf{G}_i^*\hat{\boldsymbol{\alpha}}$, and the expectation is $tr\mathbf{G}_i^* \, Var(\hat{\boldsymbol{\alpha}})$. Under the pretense that $\tilde{\mathbf{G}} = \mathbf{G}$ and $\tilde{\mathbf{R}} = \mathbf{R}$

$$Var(\hat{\boldsymbol{\alpha}}) = \tilde{\mathbf{G}}^-\mathbf{G}\tilde{\mathbf{G}}^- - \mathbf{C}_{22}.$$

$\mathbf{C}_{22}$ is the lower q^2 submatrix of a g-inverse of the coefficient matrix of (5.51), which has rank, $r(\mathbf{X}) + r(\mathbf{G})$. Use a g-inverse with $q-$rank($\mathbf{G}$) rows (and cols.) zeroed in the last q rows (and cols.). Let $\mathbf{G}^-$ be a g-inverse of $\mathbf{G}$ with the same $q-$rank($\mathbf{G}$) rows (and cols.) zeroed as in $\mathbf{C}_{22}$. For ML substitute $(\mathbf{GZ}'\mathbf{R}^{-1}\mathbf{ZG})^-$ for $\mathbf{C}_{22}$.

Chapter 13
Effects of Selection

C. R. Henderson

1984 - Guelph

1 Introduction

The models and the estimation and prediction methods of the preceding chapters have not addressed the problem of data arising from a selection program. Note that the assumption has been that the expected value of every element of $\mathbf{u}$ is 0. What if $\mathbf{u}$ represents breeding values of animals that have been produced by a long-time, effective, selection program? In that case we would expect the breeding values in later generations to be higher than in the earlier ones. Consequently the expected value of $\mathbf{u}$ is not really $\mathbf{0}$ as assumed in the methods presented earlier. Also it should be noted that, in an additive genetic model, $\mathbf{A}\sigma_a^2$ is a correct statement of the covariance matrix of breeding values if no selection has taken place and $\sigma_a^2 =$ additive genetic variance in an unrelated, non-inbred, unselected population. Following selection this no longer is true. Generally variances are reduced and the covariances are altered. In fact, there can be non-zero covariances for pairs of unrelated animals. Further, we often assume for one trait that $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$. Following selection this is no longer true. Variances are reduced and non-zero covariances are generated. Another potentially serious consequence of selection is that previously uncorrelated elements of $\mathbf{u}$ and $\mathbf{e}$ become correlated with selection. If we know the new first and second moments of $(\mathbf{y}, \mathbf{u})$ we can then derive BLUE and BLUP for that model. This is exceedingly difficult for two reasons. First, because selection intensity varies from one herd to another, a different set of parameters would be needed for each herd, but usually with too few records for good estimates to be obtained. Second, correlation of $\mathbf{u}$ with $\mathbf{e}$ complicates the computations. Fortunately, as we shall see later in this chapter, computations that ignore selection and then use the parameters existing prior to selection sometimes result in BLUE and BLUP under the selection model. Unfortunately, comparable results have not been obtained for variance and covariance estimation, although there does seem to be some evidence that MIVQUE with good priors, REML, and ML may have considerable ability to control bias due to selection, Rothschild et al. (1979).

2 An Example of Selection

We illustrate some effects of selection and the properties of BLUE, BLUP, and OLS by a progeny test example. The progeny numbers were distributed as follows

	Treatments	
Sires	1	2
1	10	500
2	10	100
3	10	0
4	10	0

We assume that the sires were ranked from highest to lowest on their progeny averages in Period 1. If that were true in repeated sampling and if we assume normal distributions, one can write the expected first and second moments. Assume unrelated sires, $\sigma_e^2 = 15$, $\sigma_s^2 = 1$ under a model,

$$y_{ijk} \;=\; s_i \;+\; p_j \;+\; e_{ijk}.$$

With no selection

$$E\begin{pmatrix} \bar{y}_{11} \\ \bar{y}_{21} \\ \bar{y}_{31} \\ \bar{y}_{41} \\ \bar{y}_{12} \\ \bar{y}_{22} \end{pmatrix} = \begin{pmatrix} p_1 \\ p_1 \\ p_1 \\ p_1 \\ p_2 \\ p_2 \end{pmatrix}, \; Var = \begin{pmatrix} 2.5 & 0 & 0 & 0 & 1 & 0 \\ & 2.5 & 0 & 0 & 0 & 1 \\ & & 2.5 & 0 & 0 & 0 \\ & & & 2.5 & 0 & 0 \\ & & & & 1.03 & 0 \\ & & & & & 1.15 \end{pmatrix}.$$

With ordering of sires according to first records the corresponding moments are

$$\begin{pmatrix} 1.628 & + & p_1 \\ .460 & + & p_1 \\ -.460 & + & p_1 \\ -1.628 & + & p_1 \\ .651 & + & p_2 \\ .184 & + & p_2 \end{pmatrix} \text{ and } \begin{pmatrix} 1.229 & .614 & .395 & .262 & .492 & .246 \\ & .901 & .590 & .395 & .246 & .360 \\ & & .901 & .614 & .158 & .236 \\ & & & 1.229 & .105 & .158 \\ & & & & .827 & .098 \\ & & & & & .894 \end{pmatrix}.$$

Further, with no ordering $E(\mathbf{s}) \;=\; \mathbf{0}, Var(\mathbf{s}) \;=\; \mathbf{I}$. With ordering these become

$$\begin{pmatrix} .651 \\ .184 \\ -.184 \\ -.651 \end{pmatrix} \text{ and } \begin{pmatrix} .797 & .098 & .063 & .042 \\ & .744 & .094 & .063 \\ & & .744 & .098 \\ & & & .797 \end{pmatrix}.$$

These results are derived from Teicheroew (1956), Sarhan and Greenberg (1956), and Pearson (1903).

Suppose $p_1 = 10$, $p_2 = 12$. Then in repeated sampling the expected values of the 6 subclass means would be

$$\begin{pmatrix} 11.628 & 12.651 \\ 10.460 & 12.184 \\ 9.540 & -- \\ 8.372 & -- \end{pmatrix}$$

Applying BLUE and BLUP, ignoring selection, to these expected data the mixed model equations are

$$\begin{pmatrix} 40 & 0 & 10 & 10 & 10 & 10 \\ & 600 & 500 & 100 & 0 & 0 \\ & & 525 & 0 & 0 & 0 \\ & & & 125 & 0 & 0 \\ & & & & 25 & 0 \\ & & & & & 25 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} = \begin{pmatrix} 400.00 \\ 7543.90 \\ 6441.78 \\ 1323.00 \\ 95.40 \\ 83.72 \end{pmatrix}$$

The solution is $[10.000, \quad 12.000, \quad .651, \quad .184, \quad -.184, \quad -.651]$, thereby demonstrating unbiasedness of $\hat{\mathbf{p}}$ and $\hat{\mathbf{s}}$. The reason for this is discussed in Section 13.5.1.

In contrast the OLS solution gives biased estimators and predictors. Forcing $\sum \hat{s}_i = 0$ as in the BLUP solution we obtain as the solution

$$[10.000, \quad 11.361, \quad 1.297, \quad .790, \quad -.460, -1.628].$$

Except for $\hat{p}_1$ these are biased. If OLS is applied to only the data of period 2, $s_1^o - s_2^o$ is an unbiased predictor of $s_1 - s_2$. The equations in this case are

$$\begin{pmatrix} 600 & 500 & 100 \\ 500 & 500 & 0 \\ 100 & 0 & 100 \end{pmatrix} \begin{pmatrix} p_2^o \\ s_1^o \\ s_2^o \end{pmatrix} = \begin{pmatrix} 7543.90 \\ 6325.50 \\ 1218.40 \end{pmatrix}.$$

A solution is $[0, \quad 12.651, \quad 12.184]$. Then $s_1^o - s_2^o = .467 = E(s_1 - s_2)$ under the selection model. This result is equivalent to a situation in which the observations on the first period are not observable and we define selection at that stage as selection on $\mathbf{u}$, in which case treating $\mathbf{u}$ as fixed in the computations leads to unbiased estimators and predictors. Note, however, that we obtain invariant solutions only for functions that are estimable under a fixed $\mathbf{u}$ model. Consequently p_2 is not estimable and we can predict only the difference between s_1 and s_2.

3 Conditional Means And Variances

Pearson (1903) derived results for the multivariate normal distribution that are extremely useful for studying the selection problem. These are the results that were used in the example in Section 13.2. We shall employ the notation of Henderson (1975a), similar to that of Lawley (1943), rather than Pearson's, which was not in matrix notation. With no selection $[\mathbf{v}_1' \ \ \mathbf{v}_2']$ have a multivariate normal distribution with means,

$$\begin{pmatrix} \mu_1' & \mu_2' \end{pmatrix}, \quad \text{and} \quad Var\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix}. \tag{1}$$

Suppose now in conceptual repeated sampling $\mathbf{v}_2$ is selected in such a way that it has mean $= \mu_2 + \mathbf{k}$ and variance $= \mathbf{C}_s$. Then Pearson's result is

$$E_s\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} \mu_1 + \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{k} \\ \mu_2 + \mathbf{k} \end{pmatrix}. \tag{2}$$

$$Var_s\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_0\mathbf{C}_{12}' & \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_s \\ \mathbf{C}_s\mathbf{C}_{22}^{-1}\mathbf{C}_{12}' & \mathbf{C}_s \end{pmatrix}, \tag{3}$$

where $\mathbf{C}_0 = \mathbf{C}_{22}^{-1}(\mathbf{C}_{22} - \mathbf{C}_s)\mathbf{C}_{22}^{-1}$. Henderson (1975) used this result to derive BLUP and BLUE under a selection model with a multivariate normal distribution of $(\mathbf{y}, \mathbf{u}, \mathbf{e})$ assumed. Let $\mathbf{w}$ be some vector correlated with $(\mathbf{y}, \mathbf{u})$. With no selection

$$E\begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{e} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{X}\boldsymbol{\beta} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{d} \end{pmatrix}, \tag{4}$$

$$Var\begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{e} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{V} & \mathbf{ZG} & \mathbf{R} & \mathbf{B} \\ \mathbf{GZ}' & \mathbf{G} & \mathbf{0} & \mathbf{B}_u \\ \mathbf{R} & \mathbf{0} & \mathbf{R} & \mathbf{B}_e \\ \mathbf{B}' & \mathbf{B}_u' & \mathbf{B}_e' & \mathbf{H} \end{pmatrix}, \tag{5}$$

and

$$\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}, \ \mathbf{B} = \mathbf{ZB}_u + \mathbf{B}_e.$$

Now suppose that in repeated sampling $\mathbf{w}$ is selected such that $E(\mathbf{w}) = \mathbf{s} \neq \mathbf{d}$, and $Var(\mathbf{w}) = \mathbf{H}_s$. Then the conditional moments are as follows.

$$E\begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{X}\boldsymbol{\beta} + \mathbf{Bt} \\ \mathbf{B}_u\mathbf{t} \\ \mathbf{s} \end{pmatrix}, \tag{6}$$

where $\mathbf{t} = \mathbf{H}^{-1}(\mathbf{s} - \mathbf{d})$.

$$Var\begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{V} - \mathbf{BH}_0\mathbf{B}' & \mathbf{ZG} - \mathbf{BH}_0\mathbf{B}' & \mathbf{BH}^{-1}\mathbf{H}_s \\ \mathbf{GZ}' - \mathbf{BH}_0\mathbf{B}' & \mathbf{G} - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u' & \mathbf{B}_u\mathbf{H}^{-1}\mathbf{H}_s \\ \mathbf{H}_s\mathbf{H}^{-1}\mathbf{B}' & \mathbf{H}_s\mathbf{H}^{-1}\mathbf{B}_u' & \mathbf{H}_s \end{pmatrix}, \tag{7}$$

where $\mathbf{H}_0 = \mathbf{H}^{-1}(\mathbf{H} - \mathbf{H}_s)\mathbf{H}^{-1}$.

4 BLUE And BLUP Under Selection Model

To find BLUE of $\mathbf{K}'\boldsymbol{\beta}$ and BLUP of $\mathbf{u}$ under this conditional model, find linear functions that minimize diagonals of $Var(\mathbf{K}'\boldsymbol{\beta})$ and variance of diagonals of $(\hat{\mathbf{u}} - \mathbf{u})$ subject to

$$E(\mathbf{K}'\boldsymbol{\beta}^o) = \mathbf{K}'\boldsymbol{\beta} \text{ and } E(\hat{\mathbf{u}}) = \mathbf{B}_u\mathbf{t}.$$

This is accomplished by modifying GLS and mixed model equations as follows.

$$\begin{pmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{V}^{-1}\mathbf{B} \\ \mathbf{B}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{B}'\mathbf{V}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{t}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ \mathbf{B}'\mathbf{V}^{-1}\mathbf{y} \end{pmatrix} \tag{8}$$

BLUP of $\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}$ is

$$\mathbf{k}'\boldsymbol{\beta}^o + \mathbf{m}'\boldsymbol{\beta}_u\mathbf{t}^o + \mathbf{m}'\mathbf{G}\mathbf{Z}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o - \mathbf{B}\mathbf{t}^o).$$

Modified mixed model equations are

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1`}\mathbf{Z} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{B}_e \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{B}_e - \mathbf{G}^{-1}\mathbf{B}_u \\ \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{X} & \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{Z}^{-}\mathbf{B}_u'\mathbf{G}^{-1} & \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{B}_e + \mathbf{B}_u'\mathbf{G}^{-1}\mathbf{B}_u \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \\ \mathbf{t}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} & \mathbf{Z}'\mathbf{R}^{-1`}\mathbf{y} & \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}'. \tag{9}$$

BLUP of $\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}$ is $\mathbf{k}'\boldsymbol{\beta}^o + \mathbf{m}'\mathbf{u}^o$. In equations (8) and (9) we use $\mathbf{u}^o$ rather than $\hat{\mathbf{u}}$ because the solution is not always invariant. It is necessary therefore to examine whether the function is predictable. The sampling and prediction error variances come from a g-inverse of (8) or (9). Let a g-inverse of the matrix of (8) be

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix},$$

then

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) = \mathbf{K}'\mathbf{C}_{11}\mathbf{K}. \tag{10}$$

Let a g-inverse of the matrix of (9) be

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{13}' & \mathbf{C}_{23}' & \mathbf{C}_{33} \end{pmatrix},$$

Then

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) = \mathbf{K}'\mathbf{C}_{11}\mathbf{K}. \tag{11}$$

$$Cov(\mathbf{K}'\boldsymbol{\beta}^o,\ \hat{\mathbf{u}} - \mathbf{u}) = \mathbf{K}'\mathbf{C}_{12}. \tag{12}$$

$$Var(\hat{\mathbf{u}} - \mathbf{u}) = \mathbf{C}_{22}. \tag{13}$$

$$Cov(\mathbf{K}'\boldsymbol{\beta}^o,\ \hat{\mathbf{u}}') = \mathbf{K}'\mathbf{C}_{13}\mathbf{B}_u'. \tag{14}$$

$$Var(\hat{\mathbf{u}}) = \mathbf{G} - \mathbf{C}_{22} + \mathbf{C}_{23}\mathbf{B}_u' + \mathbf{B}_u\mathbf{C}_{23}' - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u'. \tag{15}$$

Note that (10), ..., (13) are analogous to the results for the no selection model, but (14) and (15) are more complicated. The problems with the methods of this section are that $\mathbf{w}$ may be difficult to define and the values of $\mathbf{B}_u$ and $\mathbf{B}_e$ may not be known. Special cases exist that simplify the problem. This is true particularly if selection is on a subvector of $\mathbf{y}$, and if estimators and predictors can be found that are invariant to the value of $\boldsymbol{\beta}$ associated with the selection functions.

5 Selection On Linear Functions Of y

Suppose that whatever selection has occurred has been a consequence of use of the record vector or some subvector of $\mathbf{y}$. Let the type of selection be described in terms of a set of linear functions, say $\mathbf{L}'\mathbf{y}$, such that

$$E(\mathbf{L}'\mathbf{y}) \;=\; \mathbf{L}'\mathbf{X}\boldsymbol{\beta} + \mathbf{t},$$

where $\mathbf{t} \neq \mathbf{0}$. $\mathbf{t}$ would be $\mathbf{0}$ if there were no selection.

$$Var(\mathbf{L}'\mathbf{y}) \;=\; \mathbf{H}_s.$$

Let us see how this relates to (9).

$$\mathbf{B}_u \;=\; \mathbf{G}\mathbf{Z}'\mathbf{L},\; \mathbf{B}_e \;=\; \mathbf{R}\mathbf{L},\; \mathbf{H} \;=\; \mathbf{L}'\mathbf{V}\mathbf{L}.$$

Substituting these values in (9) we obtain

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{L} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{L}'\mathbf{X} & \mathbf{0} & \mathbf{L}'\mathbf{V}\mathbf{L} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{L}'\mathbf{y} \end{pmatrix}. \tag{16}$$

5.1 Selection with $\mathbf{L}'\mathbf{X} \;=\; \mathbf{0}$

An important property of (16) is that if $\mathbf{L}'\mathbf{X} \;=\; \mathbf{0}$, then $\hat{\mathbf{u}}$ is a solution to the mixed model equations assuming no selection. Thus we have the extremely important result that whenever $\mathbf{L}'\mathbf{X} \;=\; \mathbf{0}$, BLUE and BLUP in the selection model can be computed by using the mixed model equations ignoring selection. Our example in Section 2 can be formulated as a problem with $\mathbf{L}'\mathbf{X} \;=\; \mathbf{0}$. Order the observations by sires within periods. Let

$$\mathbf{y}' \;=\; [\bar{y}_{11.}^{'},\; \bar{y}_{21.}^{'},\; \bar{y}_{31.}^{'},\; \bar{y}_{41.}^{'},\; \bar{y}_{12.}^{'},\; \bar{y}_{22.}^{'}].$$

According to our assumptions of the method of selection

$$\bar{y}_{11.} \;> \bar{y}_{21.} \;> \bar{y}_{31.} \;> \bar{y}_{41.}.$$

Based on this we can write

$$\mathbf{L}' = \begin{pmatrix} \mathbf{1}_{10}' & -\mathbf{1}_{10}' & \mathbf{0}_{620}' & \\ \mathbf{0}_{10}' & \mathbf{1}_{10}' & -\mathbf{1}_{10}' & \mathbf{0}_{610}' \\ \mathbf{0}_{20}' & \mathbf{1}_{10}' & -\mathbf{1}_{10}' & \mathbf{0}_{600}' \end{pmatrix}$$

where

$$\mathbf{1}_{10}' \text{ denotes a row vector of 10 one's.}$$
$$\mathbf{0}_{620}' \text{ denotes a null row vector with 620 elements, etc.}$$

It is easy to see that $\mathbf{L}'\mathbf{X} = \mathbf{0}$, and that explains why we obtain unbiased estimators and predictors from the solution to the mixed model equations.

Let us consider a much more general selection method that insures that $\mathbf{L}'\mathbf{X} = \mathbf{0}$. Suppose in the first cycle of selection that data to be used in selection comprise a subvector of $\mathbf{y}$, say $\mathbf{y}_s$. We know that $\mathbf{X}_s\boldsymbol{\beta}$, consisting of such fixed effects as age, sex and season, causes confusion in making selection decisions, so we adjust the data for some estimate of $\mathbf{X}_s\boldsymbol{\beta}$, say $\mathbf{X}_s\boldsymbol{\beta}^o$ so the data for selection become $\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}^o$. Suppose that we then evaluate the i^{th} candidate for selection by the function $\mathbf{a}_i'(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}^o)$. There are c candidates for selection and s of them are to be selected. Let us order the highest s of the selection functions with labels 1 for the highest, 2 for the next highest, etc. Leave the lowest $c - s$ unordered. Then the animals labelled 1, ..., s are selected, and there may, in addition, be differential usage of them subsequently depending upon their rank. Now express these selection criteria as a set of differences, of $\mathbf{a}'(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}^o)$,

$$1 - 2,\ 2 - 3, (s - 1) - s,\ s - (s + 1), ...,\ s - c.$$

Because $\mathbf{X}_s\boldsymbol{\beta}^o$ is presumably a linear function of $\mathbf{y}$ these differences are a set of linear functions of $\mathbf{y}$, say $\mathbf{L}'\mathbf{y}$. Now suppose $\boldsymbol{\beta}^o$ is computed in such a way that $E(\mathbf{X}_s\boldsymbol{\beta}^o) = \mathbf{X}_s\boldsymbol{\beta}$ in a no selection model. (It need not be an unbiased estimator under a selection model, but if it is, that creates no problem). Then $\mathbf{L}'\mathbf{X}$ will be null, and the mixed model equations ignoring selection yield BLUE and BLUP for the selection model. This result is correct if we know $\mathbf{G}$ and $\mathbf{R}$ to proportionality. Errors in $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ will result in biases under a selection model, the magnitude of bias depending upon how seriously $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ depart from $\mathbf{G}$ and $\mathbf{R}$ and upon the intensity of selection. The result also depends upon normality. The consequences of departure from this distribution are not known in general, but depend upon the form of the conditional means.

We can extend this description of selection for succeeding cycles of selection and still have $\mathbf{L}'\mathbf{X} = \mathbf{0}$. The results above depended upon the validity of the Pearson result and normality. Now with continued selection we no longer have the multivariate normal distribution, and consequently the Pearson result may not apply exactly. Nevertheless with traits of relatively low heritability and with a new set of normally distributed errors for each new set of records, the conditional distribution of Pearson may well be a suitable approximation.

6 With Non-Observable Random Factors

The previous section deals with strict truncation selection on a linear function of records. This is not entirely realistic as there certainly are other factors that influence the selection decisions, for example, death, infertility, undesirable traits not recorded as a part of the data vector, $\mathbf{y}$. It even may be the case that the breeder did have available additional records and used them, but these were not available to the person or organization attempting to estimate or predict. For these reasons, let us now consider a different selection model, the functions used for making selection decision now being

$$\mathbf{a}_i^{'} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o) \ + \ \alpha_i$$

where α_i is a random variable not observable by the person performing estimation and prediction, but may be known or partially known by the breeder. This leads to a definition of $\mathbf{w}$ as follows

$$\mathbf{w} \ = \ \mathbf{L}'\mathbf{y} + \boldsymbol{\theta}.$$

$$Cov(\mathbf{y}, \mathbf{w}') \ = \ \mathbf{B} = \mathbf{VL} + \mathbf{C}, \text{ where } \mathbf{C} = Cov(\mathbf{y}, \boldsymbol{\theta}'). \quad (17)$$

$$Cov(\mathbf{u}, \mathbf{w}') \ = \ \mathbf{B}_u = \mathbf{GZ}'\mathbf{L} + \mathbf{C}_u, \text{ where } \mathbf{C}_u = Cov(\mathbf{u}, \boldsymbol{\theta}') \quad (18)$$

$$Cov(\mathbf{e}, \mathbf{w}') \ = \ \mathbf{B}_e = \mathbf{RL} + \mathbf{C}_e, \text{ where } \mathbf{C}_e = Cov(\mathbf{e}, \boldsymbol{\theta}'). \quad (19)$$

$$Var(\mathbf{w}) \ = \ \mathbf{L}'\mathbf{VL} + \mathbf{L}'\mathbf{C} + \mathbf{C}'\mathbf{L} + \mathbf{C}_\theta, \text{ where } \mathbf{C}_\theta = Var(\boldsymbol{\theta}). \quad (20)$$

Applying these results to (9) we obtain the modified mixed model equations below)

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{L} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{C}_e \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{Z}\mathbf{R}^{-1}\mathbf{C}_e - \mathbf{G}^{-1}\mathbf{C}_u \\ \mathbf{L}'\mathbf{X} + \mathbf{C}_e^{'}\mathbf{R}^{-1}\mathbf{X} & \mathbf{C}_e^{'}\mathbf{R}^{-1}\mathbf{Z}' + \mathbf{C}_u^{'}\mathbf{G}^{-1} & \boldsymbol{\psi} \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{L}'\mathbf{y} + \mathbf{C}_e^{'}\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}, \quad (21)$$

where $\boldsymbol{\psi} \ = \ \mathbf{L}'\mathbf{VL} + \mathbf{C}_e^{'}\mathbf{R}^{-1}\mathbf{C}_e + \mathbf{C}_u^{'}\mathbf{G}^{-1}\mathbf{C}_u + \mathbf{L}'\mathbf{C} + \mathbf{C}'\mathbf{L}$.

Now if $\mathbf{L}'\mathbf{X} = \mathbf{0}$ and if $\boldsymbol{\theta}$ is uncorrelated with $\mathbf{u}$ and $\mathbf{e}$, these equations reduce to the regular mixed model equations that ignore selection. Thus the non-observable variable used in selection causes no difficulty when it is uncorrelated with $\mathbf{u}$ and $\mathbf{e}$. If the correlations are non-zero, one needs the magnitudes of $\mathbf{C}_e$, $\mathbf{C}_u$ to obtain BLUE and BLUP. This could be most difficult to determine. The selection models of Sections 5 and 6 are described in Henderson (1982).

7 Selection On A Subvector Of y

Many situations exist in which selection has occurred on $\mathbf{y}_1$, but $\mathbf{y}_2$ is unselected, where the model is

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1\boldsymbol{\beta} \\ \mathbf{X}_2\boldsymbol{\beta} \end{pmatrix} + \begin{pmatrix} \mathbf{Z}_1\mathbf{u} \\ \mathbf{Z}_2\mathbf{u} \end{pmatrix} + \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix},$$

$$Var\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{12} & \mathbf{R}_{22} \end{pmatrix}.$$

Presumably $\mathbf{y}_1$ are data from earlier generations. Suppose that selection which has occurred can be described as

$$\mathbf{L}'\mathbf{y} = (\mathbf{M}' \;\; \mathbf{0})\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}.$$

Then the equations of (16) become

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}_1'\mathbf{M} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{M}'\mathbf{X}_1 & \mathbf{0} & \mathbf{M}'\mathbf{V}_{11}\mathbf{M} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{M}'\mathbf{y}_1 \end{pmatrix} \tag{22}$$

Then if $\mathbf{M}'\mathbf{X}_1 = \mathbf{0}$, unmodified mixed model equations yield unbiased estimators and predictors. Also if selection is on $\mathbf{M}'\mathbf{y}_1$ plus a non-observable variable uncorrelated with $\mathbf{u}$ and $\mathbf{e}$ and $\mathbf{M}'\mathbf{X}_1 = \mathbf{0}$, the unmodified equations are appropriate.

Sometimes $\mathbf{y}_1$ is not available to the person predicting functions of $\boldsymbol{\beta}$ and $\mathbf{u}$. Now if we assume that $\mathbf{R}_{12} = \mathbf{0}$,

$$\begin{aligned} E(\mathbf{y}_2 \mid \mathbf{M}'\mathbf{y}_1) &= \mathbf{Z}_2\mathbf{G}\mathbf{Z}_1'\mathbf{M}\mathbf{k}. \\ E(\mathbf{u} \mid \mathbf{M}'\mathbf{y}_1) &= \mathbf{G}\mathbf{Z}_1'\mathbf{M}\mathbf{k}, \end{aligned}$$

where

$$\mathbf{k} = (\mathbf{M}'\mathbf{V}_{11}\mathbf{M})^{-1}\mathbf{t},$$

$\mathbf{t}$ being the deviation of mean of $\mathbf{M}'\mathbf{y}_1$ from $\mathbf{X}_1'\boldsymbol{\beta}$. If we solve for $\boldsymbol{\beta}^o$ and $\mathbf{u}^o$ in the equations (23) that regard $\mathbf{u}$ as fixed for purposes of computation, then

$$E[\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{T}'\mathbf{u}^o] = \mathbf{K}'\boldsymbol{\beta} + E[\mathbf{T}'\mathbf{u} \mid \mathbf{M}'\mathbf{y}_1]$$

provided that $\mathbf{K}'\boldsymbol{\beta} + \mathbf{T}'\mathbf{u}$ is estimable under a fixed $\mathbf{u}$ model.

$$\begin{pmatrix} \mathbf{X}_2'\mathbf{R}_{22}^{-1}\mathbf{X}_2 & \mathbf{X}_2'\mathbf{R}_{22}^{-1}\mathbf{Z}_2 \\ \mathbf{Z}_2'\mathbf{R}_{22}^{-1}\mathbf{X}_2 & \mathbf{Z}_2'\mathbf{R}_{22}^{-1}\mathbf{Z}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}_2'\mathbf{R}_{22}^{-1}\mathbf{y}_2 \\ \mathbf{Z}_2'\mathbf{R}_{22}^{-1}\mathbf{y}_2 \end{pmatrix}. \tag{23}$$

This of course does not prove that $\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{T}'\mathbf{u}^o$ is BLUP of this function under $\mathbf{M}'\mathbf{y}_1$ selection and utilizing only $\mathbf{y}_2$. Let us examine modified mixed model equations regarding $\mathbf{y}_2$ as the data vector and $\mathbf{M}'\mathbf{y}_1 = \mathbf{w}$. We set up equations like (21).

$$\begin{aligned}\mathbf{B}_e &= Cov[\mathbf{e}_2, \mathbf{y}_1'\mathbf{M}] = \mathbf{0} \text{ if we assume } \mathbf{R}_{12} = \mathbf{0}. \\ \mathbf{B}_u &= Cov(\mathbf{u}, \mathbf{y}_1'\mathbf{M}) = \mathbf{G}\mathbf{Z}_1'\mathbf{M}.\end{aligned}$$

Then the modified mixed model equations become

$$\begin{pmatrix} \mathbf{X}_2'\mathbf{R}_{22}^{-1}\mathbf{X}_2 & \mathbf{X}_2'\mathbf{R}_{22}^{-1}\mathbf{Z}_2 & \mathbf{0} \\ \mathbf{Z}_2'\mathbf{R}_{22}^{-1}\mathbf{X}_2 & \mathbf{Z}_2'\mathbf{R}_{22}^{-1}\mathbf{Z}_2 + \mathbf{G}^{-1} & -\mathbf{Z}_1'\mathbf{M} \\ \mathbf{0} & -\mathbf{M}'\mathbf{Z}_1 & \mathbf{M}'\mathbf{Z}_1\mathbf{G}\mathbf{Z}_1'\mathbf{M} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_2'\mathbf{R}_{22}^{-1}\mathbf{y}_2 \\ \mathbf{Z}_2'\mathbf{R}_{22}^{-1}\mathbf{y}_2 \\ \mathbf{0} \end{pmatrix}. \quad (24)$$

A sufficient set of conditions for the solution to $\boldsymbol{\beta}^o$ and $\mathbf{u}^o$ in these equations being equal to those of (23) is that $\mathbf{M}' = \mathbf{I}$ and $\mathbf{Z}_1$ be non-singular. In that case if we "absorb" $\boldsymbol{\theta}$ we obtain the equations of (23).

Now it seems implausible that $\mathbf{Z}_1$ be non-singular. In fact, it would usually have more rows than columns. A more realistic situation is the following. Let $\bar{\mathbf{y}}_1$ be the mean of smallest subclasses in the $\bar{\mathbf{y}}_1$ vector. Then the model for $\bar{\mathbf{y}}_1$ is

$$\bar{\mathbf{y}}_1 = \bar{\mathbf{X}}_1\boldsymbol{\beta} + \bar{\mathbf{Z}}_1\mathbf{u} + \mathbf{e}_1.$$

See Section 1.6 for a description of such models. Now suppose selection can be described as $\mathbf{I}\bar{\mathbf{y}}_1$. Then

$$\begin{aligned}\mathbf{B}_e &= \mathbf{0} \;\text{ if }\; \mathbf{R}_{12} = \mathbf{0}, \;\text{ and} \\ \mathbf{B}_u &= \bar{\mathbf{Z}}_1.\end{aligned}$$

Then a sufficient condition for GLS using $\mathbf{y}_2$ only and computing as though $\mathbf{u}$ is fixed to be BLUP under the selection model and regarding $\mathbf{y}_2$ as that data vector is that $\bar{\mathbf{Z}}_1$ be non-singular. This might well be the case in some practical situations. This is the selection model in our sire example.

8 Selection On u

Cases exist in animal breeding in which the data represent observations associated with $\mathbf{u}$ that have been subject to prior selection, but with the data that were used for such selection not available. Henderson (1975a) described this as $\mathbf{L}'\mathbf{u}$ selection. If no selection on the observable $\mathbf{y}$ vector has been effected, BLUE and BLUP come from solution to equations (25).

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{0} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & -\mathbf{L} \\ \mathbf{0} & -\mathbf{L}' & \mathbf{L}'\mathbf{G}\mathbf{L} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{0} \end{pmatrix} \quad (25)$$

These reduce to (26) by "absorbing" $\boldsymbol{\theta}$.

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} - \mathbf{L}(\mathbf{L}'\mathbf{G}\mathbf{L})^{-1}\mathbf{L}' \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} \tag{26}$$

The notation $\mathbf{u}^o$ is used rather than $\hat{\mathbf{u}}$ since the solution may not be unique, in which case we need to consider functions of $\mathbf{u}^o$ that are invariant to the solution. It is simple to prove that $\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}^o$ is an unbiased predictor of $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$, where $\boldsymbol{\beta}^o$ and $\mathbf{u}^o$ are some solution to (27) and this is an estimable function under a fixed $\mathbf{u}$ model

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{27}$$

A sufficient condition for this to be BLUP is that $\mathbf{L} = \mathbf{I}$. The proof comes by substituting $\mathbf{I}$ for $\mathbf{L}$ in (26). In sire evaluation $\mathbf{L}'\mathbf{u}$ selection can be accounted for by proper grouping. Henderson (1973) gave an example of this for unrelated sires. Quaas and Pollak (1981) extended this result for related sires. Let $\mathbf{G} = \mathbf{A}\sigma_s^2$. Write the model for progeny as

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{Z}\mathbf{Q}\mathbf{g} + \mathbf{Z}\mathbf{S} + \mathbf{e},$$

where $\mathbf{h}$ refers to fixed herd-year-season and $\mathbf{g}$ to fixed group effects. Then it was shown that such grouping is equivalent to no grouping, defining $\mathbf{L} = \mathbf{G}^{-1}\mathbf{Q}$, and then using (25). We illustrate this method with the following data.

group	sire	n_i	y_i
1	1	2	10
	2	3	12
	3	1	7
2	4	2	6
	5	3	8
3	6	1	3
	7	2	5
	8	1	2
	9	2	8

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & .5 & .5 & 0 & .25 & .25 & 0 & .125 \\ & 1 & 0 & 0 & .5 & 0 & 0 & .25 & 0 \\ & & 1 & .25 & 0 & .5 & .125 & 0 & .25 \\ & & & 1 & 0 & .125 & .5 & 0 & .0625 \\ & & & & 1 & 0 & 0 & .5 & 0 \\ & & & & & 1 & .0625 & 0 & .5 \\ & & & & & & 1 & 0 & .03125 \\ & & & & & & & 1 & 0 \\ & & & & & & & & 1 \end{pmatrix}.$$

Assume a model $y_{ijk} = \mu + g_i + s_{ij} + e_{ijk}$. Let $\sigma^2_e = 1$, $\sigma^2_s = 12^{-1}$, then $\mathbf{G} = 12^{-1}\mathbf{A}$. The solution to the mixed model equations with μ dropped is

$$\hat{\mathbf{g}} = (4.8664,\ 2.8674,\ 2.9467),$$

$$\hat{\mathbf{s}} = (.0946,\ -.1937,\ .1930,\ .0339,\ -.1350,\ .1452,\ -.0346,\ -.1192,\ .1816).$$

The sire evaluations are $\hat{g}_i + \hat{s}_{ij}$ and these are (4.961, 4.673, 5.059, 2.901, 2.732, 3.092, 2.912, 2.827, 3.128).

$$\mathbf{Q}' = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

This gives

$$\mathbf{L}' = \begin{pmatrix} 12 & 16 & 12 & -8 & -8 & -8 & 0 & 0 & 0 \\ -8 & -8 & 0 & 20 & 20 & 0 & -8 & -8 & 0 \\ 0 & 0 & -8 & -8 & -8 & 12 & 16 & 16 & 8 \end{pmatrix} = \mathbf{G}^{-1}\mathbf{Q},$$

and

$$\mathbf{L}'\mathbf{G}\mathbf{L} = \begin{pmatrix} 40 & -16 & -8 \\ & 40 & -16 \\ & & 52 \end{pmatrix}.$$

Then the equations like (25) give a solution

$$\mu^o = 2.9014,$$

$$\mathbf{s}^o = (2.0597,\ 1.7714,\ 2.1581,\ 0, -.1689,\ .1905,\ .0108,\ -.0739,\ .2269),$$

$$\boldsymbol{\theta} = (1.9651,\ -.0339,\ .0453).$$

The sire evaluation is $\mu^o + s^o_i$ and this is the same as when groups were included.

9 Inverse Of Conditional A Matrix

In some applications the base population animals are not a random sample from some population, but rather have been selected. Consequently the additive genetic variance-covariance matrix for these animals is not $\sigma^2_a\mathbf{I}$, where σ^2_a is the additive genetic variance in the population from which these animals were were taken. Rather it is $\mathbf{A}_s\sigma^2_{a*}$, where $\sigma^2_{a*} \neq \sigma^2_a$ in general. If the base population had been a random sample from some population, the entire $\mathbf{A}$ matrix would be

$$\begin{pmatrix} \mathbf{I} & \mathbf{A}_{12} \\ \mathbf{A}'_{12} & \mathbf{A}_{22} \end{pmatrix}. \quad (28)$$

The inverse of this can be found easily by the method described by Henderson (1976). Denote this by

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix}. \tag{29}$$

If the Pearson result holds, the **A** matrix for this conditional population is

$$\begin{pmatrix} \mathbf{A}_s & \mathbf{A}_s\mathbf{A}_{12} \\ \mathbf{A}_{12}'\mathbf{A}_s & \mathbf{A}_{22} - \mathbf{A}_{12}'(\mathbf{I} - \mathbf{A}_s)\mathbf{A}_{12} \end{pmatrix} \tag{30}$$

The inverse of this matrix is

$$\begin{pmatrix} \mathbf{C}_s & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{pmatrix}, \tag{31}$$

$$\text{where } \mathbf{C}_s = \mathbf{A}_s^{-1} - \mathbf{C}_{12}\mathbf{A}_{12}', \tag{32}$$

and $\mathbf{C}_{12}$, $\mathbf{C}_{22}$ are the same as in (29)

Note that most of the elements of the inverse of the conditional matrix (31) are the same as the elements of the inverse of the unconditional matrix (29). Thus the easy method for $\mathbf{A}^{-1}$ can be used, and the only elements of the unconditional **A** needed are those of $\mathbf{A}_{12}$. Of course this method is not appropriate for the situation in which $\mathbf{A}_s$ is singular. We illustrate with

$$\text{unconditional } \mathbf{A} = \begin{pmatrix} 1.0 & 0 & .2 & .3 & .1 \\ & 1.0 & .1 & .2 & .2 \\ & & 1.1 & .3 & .5 \\ & & & 1.2 & .2 \\ & & & & 1.3 \end{pmatrix}.$$

The first 2 animals are selected so that

$$\mathbf{A}_s = \begin{pmatrix} .7 & -.4 \\ -.4 & .8 \end{pmatrix}.$$

Then by (30) the conditional **A** is

$$\begin{pmatrix} .7 & -.4 & .1 & .13 & -.01 \\ & .8 & 0 & .04 & .12 \\ & & 1.07 & .25 & .47 \\ & & & 1.117 & .151 \\ & & & & 1.273 \end{pmatrix}.$$

The inverse of the unconditional **A** is

$$\begin{pmatrix} 1.103168 & .064741 & -.136053 & -.251947 & -.003730 \\ & 1.062405 & .003286 & -.170155 & -.143513 \\ & & 1.172793 & -.191114 & -.411712 \\ & & & .977683 & -.031349 \\ & & & & .945770 \end{pmatrix}.$$

The inverse of the conditional $\mathbf{A}$ is

$$\begin{pmatrix} 2.103168 & 1.064741 & -.136053 & -.251947 & -.003730 \\ & 1.812405 & .003286 & -.170155 & -.143513 \\ & & 1.172793 & -.191114 & -.411712 \\ & & & .977683 & -.031349 \\ & & & & .954770 \end{pmatrix}.$$

$$\mathbf{A}_s^{-1} = \begin{pmatrix} 2.0 & 1.0 \\ 1.0 & 1.75 \end{pmatrix}, \ \mathbf{C}_{12} = \begin{pmatrix} -.136053 & -.251947 & -.003730 \\ .003286 & -.170155 & -.143513 \end{pmatrix},$$

$$\mathbf{A}_{12}' = \begin{pmatrix} .2 & .1 \\ .3 & .2 \\ .1 & .2 \end{pmatrix},$$

and

$$\mathbf{A}_s^{-1} - \mathbf{C}_{12}\mathbf{A}_{12}' = \begin{pmatrix} 2.103168 & 1.064741 \\ & 1.812405 \end{pmatrix},$$

which checks with the upper 2 × 2 submatrix of the inverse of conditional $\mathbf{A}$.

10 Minimum Variance Linear Unbiased Predictors

In all previous discussions of prediction in both the no selection and the selection model we have used as our criteria linear and unbiased with minimum variance of the prediction error. That is, we use $\mathbf{a}'\mathbf{y}$ as the predictor of $\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}$ and find $\mathbf{a}$ that minimizes $E(\mathbf{a}'\mathbf{y} - \mathbf{k}'\boldsymbol{\beta} - \mathbf{m}'\mathbf{u})^2$ subject to the restriction that $E(\mathbf{a}'\mathbf{y}) = \mathbf{k}'\boldsymbol{\beta} + E(\mathbf{m}'\mathbf{u})$. This is a logical criterion for making selection decisions. For other purposes such as estimating genetic trend one might wish to minimize the variance of the predictor rather than the variance of the prediction error. Consequently in this section we shall derive a predictor of $\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}$, say $\mathbf{a}'\mathbf{y}$, such that $E(\mathbf{a}'\mathbf{y}) = \mathbf{k}'\boldsymbol{\beta} + E(\mathbf{m}'\mathbf{u})$ and has minimum variance. For this purpose we use the $\mathbf{L}'\mathbf{y}$ type of selection described in Section 5. Let

$$\begin{aligned} E(\mathbf{L}'\mathbf{y}) &= \mathbf{L}'\mathbf{X}\boldsymbol{\beta} + \mathbf{t}, \ \mathbf{t} \neq \mathbf{0}. \\ Var(\mathbf{L}'\mathbf{y}) &= \mathbf{H}_s \neq \mathbf{L}'\mathbf{V}\mathbf{L}. \end{aligned}$$

Then

$$\begin{aligned} E(\mathbf{y} \mid \mathbf{L}'\mathbf{y}) &= \mathbf{X}\boldsymbol{\beta} + \mathbf{V}\mathbf{L}(\mathbf{L}'\mathbf{V}\mathbf{L})^{-1}\mathbf{t} \equiv \mathbf{X}\boldsymbol{\beta} + \mathbf{V}\mathbf{L}\mathbf{d}. \\ E(\mathbf{u} \mid \mathbf{L}'\mathbf{y}) &= \mathbf{G}\mathbf{Z}'\mathbf{L}(\mathbf{L}'\mathbf{V}\mathbf{L})^{-1}\mathbf{t} \equiv \mathbf{G}\mathbf{Z}'\mathbf{L}\mathbf{d}. \\ Var(\mathbf{y} \mid \mathbf{L}'\mathbf{y}) &= \mathbf{V} - \mathbf{V}\mathbf{L}(\mathbf{L}'\mathbf{V}\mathbf{L})^{-1}(\mathbf{L}'\mathbf{V} - \mathbf{H}_s)(\mathbf{L}'\mathbf{V}\mathbf{L})^{-1}\mathbf{L}'\mathbf{V} \equiv \mathbf{V}_s. \end{aligned}$$

Then we minimize $Var(\mathbf{a}'\mathbf{y})$ subject to $E(\mathbf{a}'\mathbf{y}) = \mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{G}\mathbf{Z}'\mathbf{L}\mathbf{d}$. For this expectation to be true it is required that

$$\mathbf{X}'\mathbf{a} = \mathbf{k} \text{ and } \mathbf{L}'\mathbf{V}\mathbf{a} = \mathbf{L}'\mathbf{Z}\mathbf{G}\mathbf{m}.$$

Therefore we solve equations (33) for $\mathbf{a}$.

$$\begin{pmatrix} \mathbf{V}_s & \mathbf{X} & \mathbf{VL} \\ \mathbf{X}' & \mathbf{0} & \mathbf{0} \\ \mathbf{L}'\mathbf{V} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \boldsymbol{\theta} \\ \boldsymbol{\phi} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{k} \\ \mathbf{L}'\mathbf{ZGm} \end{pmatrix} \tag{33}$$

Let a g-inverse of the matrix of (33) be

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{13}' & \mathbf{C}_{23}' & \mathbf{C}_{33} \end{pmatrix}. \tag{34}$$

Then

$$\mathbf{a}' = \mathbf{k}'\mathbf{C}_{12}' + \mathbf{m}'\mathbf{GZ}'\mathbf{LC}_{13}'.$$

But it can be shown that a g-inverse of the matrix of (35) gives the same values of $\mathbf{C}_{11}, \mathbf{C}_{12}, \mathbf{C}_{13}$. These are subject to $\mathbf{L}'\mathbf{X} = \mathbf{0}$,

$$\begin{aligned} \mathbf{C}_{11} &= \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1} - \mathbf{L}(\mathbf{L}'\mathbf{V1})^{-1}\mathbf{L}'. \\ \mathbf{C}_{12} &= \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}, \ \mathbf{C}_{13} = \mathbf{L}(\mathbf{L}'\mathbf{VL})^{-1}\mathbf{L}'. \end{aligned}$$

Consequently we can solve for $\mathbf{a}$ in (35), a simpler set of equations than (33).

$$\begin{pmatrix} \mathbf{V} & \mathbf{X} & \mathbf{VL} \\ \mathbf{X}' & \mathbf{0} & \mathbf{0} \\ \mathbf{L}'\mathbf{V} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \boldsymbol{\theta} \\ \boldsymbol{\phi} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{k} \\ \mathbf{L}'\mathbf{ZGm} \end{pmatrix}. \tag{35}$$

By techniques described in Henderson (1975) it can be shown that

$$\mathbf{a}'\mathbf{y} = \mathbf{k}'\boldsymbol{\beta}^o + \mathbf{m}'\mathbf{GZ}'\mathbf{Lt}^o$$

where $\boldsymbol{\beta}^o, \ \mathbf{t}^o$ are a solution to (36).

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{0} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{L}'\mathbf{VL} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \mathbf{t}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{L}'\mathbf{y} \end{pmatrix}. \tag{36}$$

Thus $\boldsymbol{\beta}^o$ is a GLS solution ignoring selection, and $\mathbf{t}^o = (\mathbf{L}'\mathbf{VL})^{-1}\mathbf{L}'\mathbf{y}$. It was proved in Henderson (1975a) that

$$\begin{aligned} Var(\mathbf{K}'\boldsymbol{\beta}^o) &= \mathbf{K}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{K} = \mathbf{K}'\mathbf{C}_{11}\mathbf{K}, \\ Cov(\mathbf{K}'\boldsymbol{\beta}^o, \ \mathbf{t}') &= \mathbf{0}, \text{ and} \\ Var(\mathbf{t}) &= (\mathbf{L}'\mathbf{VL})^{-1}\mathbf{H}_s(\mathbf{L}'\mathbf{VL})^{-1}. \end{aligned}$$

Thus the variance of the predictor, $\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{m}'\hat{\mathbf{u}}$, is

$$\mathbf{K}'\mathbf{C}_{11}\mathbf{K} + \mathbf{M}'\mathbf{GZ}'\mathbf{L}(\mathbf{L}'\mathbf{VL})^{-1}\mathbf{H}_s(\mathbf{L}'\mathbf{VL})^{-1}\mathbf{L}'\mathbf{ZGM}. \tag{37}$$

In contrast to BLUP under the $\mathbf{L}'\mathbf{y}$ ($\mathbf{L}'\mathbf{X} = \mathbf{0}$) selection model, minimization of prediction variance is more difficult than minimization of variance of prediction error because the former requires writing a specific $\mathbf{L}$ matrix, and if the variance of the predictor is wanted, an estimate of $Var(\mathbf{L}'\mathbf{y})$ after selection is needed.

We illustrate with the following example with phenotypic observations in two generations under an additively genetic model.

Time	
1	2
y_{11}	y_{24}
y_{12}	y_{25}
y_{13}	y_{26}

The model is

$$y_{ij} = t_i + a_{ij} + e_{ij}.$$

$$Var(\mathbf{a}) = \begin{pmatrix} 1 & 0 & 0 & .5 & .5 & 0 \\ & 1 & 0 & 0 & 0 & .5 \\ & & 1 & 0 & 0 & 0 \\ & & & 1 & .25 & 0 \\ & & & & 1 & 0 \\ & & & & & 1 \end{pmatrix}.$$

This implies that animal 1 is a parent of animals 4 and 5, and animal 2 is a parent of animal 6. Let $Var(\mathbf{e}) = 2\mathbf{I}_6$. Thus $h^2 = 1/3$. We assume that animal 1 was chosen to have 2 progeny because $y_{11} > y_{12}$. Animal 2 was chosen to have 1 progeny and animal 3 none because $y_{12} > y_{13}$. An $\mathbf{L}$ matrix describing this type of selection and resulting in $\mathbf{L}'\mathbf{X} = \mathbf{0}$ is

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \end{pmatrix}.$$

Suppose we want to predict

$$3^{-1} \ (-1 \ \ -1 \ \ -1 \ \ 1 \ \ 1 \ \ 1) \ \mathbf{u}.$$

This would be an estimate of the genetic trend in one generation. The mixed model coefficient matrix modified for $\mathbf{L}'\mathbf{y}$ is

$$\begin{pmatrix} 1.5 & 0 & .5 & .5 & .5 & 0 & 0 & 0 & 0 & 0 \\ & 1.5 & 0 & 0 & 0 & .5 & .5 & .5 & 0 & 0 \\ & & 2.1667 & 0 & 0 & -.6667 & -.6667 & 0 & 0 & 0 \\ & & & 1.8333 & 0 & 0 & 0 & -.6667 & 0 & 0 \\ & & & & 1.5 & 0 & 0 & 0 & 0 & 0 \\ & & & & & 1.8333 & 0 & 0 & 0 & 0 \\ & & & & & & 1.8333 & 0 & 0 & 0 \\ & & & & & & & 1.8333 & 0 & 0 \\ & & & & & & & & 6 & -3 \\ & & & & & & & & & 6 \end{pmatrix}.$$

The right hand sides are $\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} & \mathbf{L}'\mathbf{y} \end{pmatrix}'$. Then solving for functions of $\mathbf{y}$ it is found that BLUP of $\mathbf{m}'\mathbf{u}$ is

$$(.05348\ .00208\ -.05556\ .00623\ .00623\ -.01246)\mathbf{y}.$$

In contrast the predictor with minimum variance is

$$[\ .05556\ \ 0\ \ -.05556\ \ 0\ \ 0\ \ 0\ ]\mathbf{y}.$$

This is a strange result in that only 2 of the 6 records are used. The variances of these two predictors are .01921 and .01852 respectively. The difference between these depends upon $\mathbf{H}_s$ relative to $\mathbf{L}'\mathbf{V}\mathbf{L}$. When $\mathbf{H}_s = \mathbf{L}'\mathbf{V}\mathbf{L}$, the variance is .01852.

As a matter of interest suppose that $\mathbf{t}$ is known and we predict using $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}$. Then the BLUP predictor is

$$(-.02703 - .06667 - .11111 .08108 .08108 .08108)(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

with variance = .09980. Note that the variance is larger than when $\mathbf{X}\beta$ is unknown. This is a consequence of the result that in both BLUP and in selection index the more information available the smaller is prediction error variance and the larger is the variance of the predictor. In fact, with perfect information the variance of $(\mathbf{m}'\hat{\mathbf{u}})$ is equal to $Var(\mathbf{m}'\mathbf{u})$ and the variance of $(\mathbf{m}'\mathbf{u} - \mathbf{m}'\hat{\mathbf{u}})$ is 0. The minimum variance predictor is the same when $\mathbf{t}$ is known as when it is unknown. Now we verify that the predictors are unbiased in the selection model described. By the Pearson result for multivariate normality,

$$E(\mathbf{y}_s) = \frac{1}{18} \begin{pmatrix} 12 & 6 \\ -6 & 6 \\ -6 & -12 \\ 2 & 1 \\ 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{d}_1 \\ \mathrm{d}_2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{t}_1 \\ \mathrm{t}_2 \end{pmatrix},$$

and

$$E(\mathbf{u}_s) = \frac{1}{18} \begin{pmatrix} 4 & 2 \\ -2 & 2 \\ -2 & -4 \\ 2 & 1 \\ 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{d}_1 \\ \mathrm{d}_2 \end{pmatrix}.$$

It is easy to verify that all of the predictors described have this same expectation. If $\mathbf{t} = \boldsymbol{\beta}$ were known, a particularly simple unbiased predictor is

$$3^{-1} \, (-1 - 1 - 1 \; 1 \; 1 \; 1) \; (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

But the variance of this predictor is very much larger than the others. The variance is 1.7222 when $\mathbf{H}_s = \mathbf{L}'\mathbf{V}\mathbf{L}$.

Chapter 14
Restricted Best Linear Prediction

C. R. Henderson

1984 - Guelph

1 Restricted Selection Index

Kempthorne and Nordskog (1959) derived restricted selection index. The model and design assumed was that the record on the j^{th} trait for the i^{th} animal is

$$y_{ij} \;=\; \mathbf{x}_{ij}^{'}\boldsymbol{\beta} + u_{ij} + e_{ij}.$$

Suppose there are n animals and t traits. It is assumed that every animal has observations on all traits. Consequently there are nt records. Further assumptions follow. Let $\mathbf{u}_i$ and $\mathbf{e}_i$ be the vectors of dimension $t \times l$ pertaining to the i^{th} animal. Then it was assumed that

$$\begin{aligned}
Var(\mathbf{u}_i) &= \mathbf{G}_0 \text{ for all } i = 1, \ldots, n, \\
Var(\mathbf{e}_i) &= \mathbf{R}_0 \text{ for all } i = 1, \ldots, n, \\
Cov(\mathbf{u}_i,\ \mathbf{u}_j^{'}) &= \mathbf{0} \text{ for all i} \neq \text{j}, \\
Cov(\mathbf{e}_i,\ \mathbf{e}_j^{'}) &= \mathbf{0} \text{ for all i} \neq \text{j}, \\
Cov(\mathbf{u}_i,\ \mathbf{e}_j^{'}) &= \mathbf{0} \text{ for all i, j}.
\end{aligned}$$

Further $\mathbf{u}, \mathbf{e}$ are assumed to have a multivariate normal distribution and $\boldsymbol{\beta}$ is assumed known. This is the model for which truncation on selection index for $\mathbf{m}'\mathbf{u}_i$ maximizes the expectation of the mean of selected $\mathbf{m}'\mathbf{u}_i$, the selection index being the conditional mean and thus meeting the criteria for Cochran's (1951) result given in Section 5.1.

Kempthorne and Nordskog were interested in maximizing improvement in $\mathbf{m}'\mathbf{u}_i$ but at the same time not altering the expected value of $\mathbf{C}_0^{'}\mathbf{u}_i$ in the selected individuals, $\mathbf{C}_0^{'}$ being of dimension sxt and having s linearly independent rows. They proved that such a restricted selection index is

$$\mathbf{a}'\mathbf{y}^*,$$

where $\mathbf{y}^*$ = the deviations of $\mathbf{y}$ from their known means and $\mathbf{a}$ is the solution to

$$\begin{pmatrix} \mathbf{G}_0 + \mathbf{R}_0 & \mathbf{G}_0\mathbf{C} \\ \mathbf{C}'\mathbf{G}_0 & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_0\mathbf{m} \\ \mathbf{0} \end{pmatrix}. \qquad (1)$$

This is a nice result but it depends upon knowing $\boldsymbol{\beta}$ and having unrelated animals and the same information on each candidate for selection. An extension of this to related animals, to unequal information, and to more general designs including progeny and sib tests is presented in the next section.

2 Restricted BLUP

We now return to the general mixed model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e},$$

where $\boldsymbol{\beta}$ is unknown, $Var(\mathbf{u}) = \mathbf{G}$, $Var(\mathbf{e}) = \mathbf{R}$ and $Cov(\mathbf{u}, \mathbf{e}') = \mathbf{0}$. We want to predict $\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}$ by $\mathbf{a}'\mathbf{y}$ where $\mathbf{a}$ is chosen so that $\mathbf{a}'\mathbf{y}$ is invariant to $\boldsymbol{\beta}$, $Var(\mathbf{a}'\mathbf{y} - \mathbf{k}'\boldsymbol{\beta} - \mathbf{m}'\mathbf{u})$ is minimum, and the expected value of $\mathbf{C}'\mathbf{u}$ given $\mathbf{a}'\mathbf{y} = \mathbf{0}$. This is accomplished by solving mixed model equations modified as in (2) and taking as the prediction $\mathbf{k}'\boldsymbol{\beta}^o + \mathbf{m}'\hat{\mathbf{u}}$.

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{G}\mathbf{C} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{G}\mathbf{C} \\ \mathbf{C}'\mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{C}'\mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{C}'\mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{G}\mathbf{C} \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{C}'\mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{2}$$

It is easy to prove that $\mathbf{C}'\hat{\mathbf{u}} = \mathbf{0}$. Premultiply the second equation by $\mathbf{C}'\mathbf{G}$ and subtract from this the third equation. This gives $\mathbf{C}'\hat{\mathbf{u}} = \mathbf{0}$.

3 Application

Quaas and Henderson (1977) presented computing algorithms for restricted BLUP in an additively genetic model and with observations on a set of correlated animals. The algorithms permit missing data on some or all observations of animals to be evaluated. Two different algorithms are presented, namely records ordered traits within animals and records ordered animals within traits. They found that in this model absorption of $\boldsymbol{\theta}$ results in a set of equations with rank less than $r + q$, the rank of regular mixed model equations, where r = rank ($\mathbf{X}$) and q = number of elements in $\mathbf{u}$. The linear dependencies relate to the coefficients of $\boldsymbol{\beta}$ but not of $\mathbf{u}$. Consequently $\hat{\mathbf{u}}$ is unique, but care needs to be exercised in solving for $\boldsymbol{\beta}^o$ and in writing $\mathbf{K}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}$, for $\mathbf{K}'\boldsymbol{\beta}$ must now be estimable under the augmented mixed model equations.

Chapter 15
Sampling from finite populations

C. R. Henderson

1984 - Guelph

1 Finite e

The populations from which samples have been drawn have been regarded as infinite in preceding chapters. Thus if a random sample of n is drawn from such a population with variance σ^2, the variance-covariance matrix of the sample vector is $\mathbf{I}_n\boldsymbol{\sigma}^2$. Suppose in contrast, the population has only t elements and a random sample of n is drawn. Then the variance-covariance matrix of the sample is

$$\begin{pmatrix} 1 & & -1/(t-1) \\ & \ddots & \\ -1/(t-1) & & 1 \end{pmatrix} \sigma^2. \quad (1)$$

If $t = n$, that is, the sample is the entire population, the variance-covariance matrix is singular. As an example, suppose that the population of observations on a fixed animal is a single observation on each day of the week. Then the model is

$$\mathbf{y}_i = \mu + \mathbf{e}_i. \quad (2)$$

$$Var(\mathbf{e}_i) = \begin{pmatrix} 1 & -1/6 & \cdots & -1/6 \\ -1/6 & 1 & \cdots & -1/6 \\ \vdots & \vdots & & \vdots \\ -1/6 & -1/6 & & 1 \end{pmatrix} \sigma^2. \quad (3)$$

Suppose we take n random observations. Then BLUE of μ is

$$\hat{\mu} = \bar{y},$$

and

$$Var(\hat{\mu}) = \frac{7-n}{6n}\,\sigma^2,$$

which equals 0 if $n = 7$. In general, with a population size, t, and a sample of n,

$$Var\ (\hat{\mu}) = \frac{t-n}{n(t-1)}\,\sigma^2,$$

which goes to σ^2/n when t goes to infinity, the latter being the usual result for a sample of n from an infinite population with $Var = \mathbf{I}\sigma^2$.

Suppose now that in this same problem we have a random sample of 3 unrelated animals with 2 observations on each and wish to estimate μ and to predict $\mathbf{a}$ when the model is

$$\begin{aligned} y_{ij} &= \mu + a_i + e_{ij}, \\ Var(\mathbf{a}) &= \mathbf{I}_3, \\ Var(\mathbf{e}) &= 6 \begin{pmatrix} 1 & -1/6 & 0 & 0 & 0 & 0 \\ -1/6 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1/6 & 0 & 0 \\ 0 & 0 & -1/6 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1/6 \\ 0 & 0 & 0 & 0 & -1/6 & 1 \end{pmatrix}. \end{aligned}$$

Then

$$\mathbf{R}^{-1} = \begin{pmatrix} 6 & 1 & 0 & 0 & 0 & 0 \\ 1 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 1 & 0 & 0 \\ 0 & 0 & 1 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 1 \\ 0 & 0 & 0 & 0 & 1 & 6 \end{pmatrix} /35.$$

The BLUP equations are

$$\begin{pmatrix} 1.2 & .4 & .4 & .4 \\ .4 & 1.4 & 0 & 0 \\ .4 & 0 & 1.4 & 0 \\ .4 & 0 & 0 & 1.4 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = .2 \begin{pmatrix} y.. \\ y_1. \\ y_2. \\ y_3. \end{pmatrix}.$$

2 Finite u

We could also have a finite number of breeding values from which a sample is drawn. If these are unrelated and are drawn at random from a population with t animals

$$Var(\mathbf{a}) = \begin{pmatrix} 1 & & -1/t \\ & \ddots & \\ -1/t & & 1 \end{pmatrix} \sigma^2_a. \tag{4}$$

If q are chosen <u>not</u> at random, we can either regard the resulting elements of $\mathbf{a}$ as fixed or we may choose to say we have a sample representing the entire population. Then

$$Var(\mathbf{a}) = \begin{pmatrix} 1 & & -1/q \\ & \ddots & \\ -1/q & & 1 \end{pmatrix} \sigma^2_{a*}, \tag{5}$$

where σ^2_{a*} probably is smaller than σ^2_a. Now $\mathbf{G}$ is singular, and we need to compute BLUP by the methods of Section 5.10. We would obtain exactly the same results if we assume $\mathbf{a}$ fixed but with levels that are unpatterned, and we then proceed to biased estimation as in Chapter 9, regarding the average values of squares and products of elements of $\mathbf{a}$ as

$$\mathbf{P} = \begin{pmatrix} 1 & & -1/q \\ & \ddots & \\ -1/q & & 1 \end{pmatrix} \sigma^2_{a*}. \tag{6}$$

3 Infinite By Finite Interactions

Much controversy has surrounded the problem of an appropriate model for the interactions in a 2 way mixed model. One commonly assumed model is that the interactions have $Var = \mathbf{I}\ \sigma^2_\gamma$. An alternative model is that the interactions in a row (rows being random and columns fixed) sum to zero. Then variance of interactions, ordered columns in rows, is

$$\begin{pmatrix} \mathbf{B} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{B} \end{pmatrix} \sigma^2_{\gamma*} \tag{7}$$

where B is $c \times c$ with 1's on the diagonal and $-1/(c-1)$ on all off- diagonals, where $c =$ number of columns. We will show in Chapter 17 how with appropriate adjustment of σ^2_r (= variance of rows) we can make them equivalent models. See Section 1.5 for definition of equivalence of models.

4 Finite By Finite Interactions

Suppose that we have a finite population of r rows and c columns. Then we might assume that the variance-covariance matrix of interactions is the following matrix multiplied by σ^2_γ.

All diagonals = 1.

Covariance between interactions in the same row $= -\sigma^2_\gamma/(c-1)$.

Covariance between interactions in the same column $= -\sigma^2_\gamma/(r-1)$.

Covariance between interations in neither the same row nor column $=$
$\sigma^2_\gamma/(r-1)(c-1)$.

If the sample involves r rows and c columns both regarded as fixed, and there is no assumed pattern of values of interactions, estimation biased by interactions can be accomplished by regarding these as pseudo-random variables and using the above "variances" for elements of $\mathbf{P}$, the average value of squares and products of interactions. This methodology was described in Chapter 9.

5 Finite, Factorial, Mixed Models

In previous chapters dealing with infinite populations from which $\mathbf{u}$ is drawn at random as well as infinite subpopulations from which subvectors $\mathbf{u}_i$ are drawn the assumption has been that the expectations of these vectors is null. In the case of a population with finite levels we shall assume that the sum of all elements of their population = 0. This results in a variance- covariance matrix with rank $\leq t-1$, where $t =$ the number of elements in the population. This is because every row (and column) of the variance-covariance matrix sums to 0. If the members of a finite population are mutually unrelated (for example, a set of unrelated sires), the variance-covariance matrix usually has d for diagonal elements and $-d/(t-1)$ for all off-diagonal elements. If the population refers to additive genetic values of a finite set of related animals, the variance-covariance matrix would be $\mathbf{A}\sigma_a^2$, but with every row (and column) of $\mathbf{A}$ summing to 0 and σ_a^2 having some value different from the infinite model value.

With respect to a factorial design with 2 factors with random and finite levels the following relationship exists. Let γ_{ij} represent the interaction variables. Then

$$\sum_{i=1}^{q_1} \gamma_{ij} = 0 \text{ for all } j = 1, \ldots, q_2,$$

and

$$\sum_{j=1}^{q_2} \gamma_{ij} = 0 \text{ for all } i = 1, \ldots, q_1, \quad (8)$$

where q_1 and q_2 are the numbers of levels of the first and second factors in the two populations.

Similarly for 3 factor interactions, γ_{ijk},

$$\begin{aligned}
\sum_{k=1}^{q_3} \gamma_{ijk} &= 0 \text{ for all } i = 1, \ldots, q_1,\ j = 1, \ldots, q_2, \\
\sum_{j=1}^{q_2} \gamma_{ijk} &= 0 \text{ for all } i = 1, \ldots, q_1,\ k = 1, \ldots, q_3, \quad \text{and} \\
\sum_{i=1}^{q_1} \gamma_{ijk} &= 0 \text{ for all } j = 1, \ldots, q_2,\ k = 1, \ldots, q_3. \quad (9)
\end{aligned}$$

This concept can be extended to any number of factors. The same principles regarding interactions can be applied to nesting factors if we visualize nesting as being a factorial design with planned disconnectedness. For example, let the first factor be sires and the second dams with 2 sires and 5 dams in the experiment. In terms of a factorial design the subclass numbers (numbers per litter, eg.) are

	Dams				
Sires	1	2	3	4	5
1	5	9	8	0	0
2	0	0	0	7	10

If this were a variance component estimation problem, we could estimate σ_s^2 and σ_e^2 but not σ_d^2 and σ_{sd}^2. We can estimate $\sigma_d^2 + \sigma_{sd}^2$ and this would usually be called $\sigma_{d/s}^2$.

6 Covariance Matrices

Consider the model

$$\mathbf{y} = \mathbf{X}\beta + \sum_i \mathbf{Z}_i \mathbf{u}_i + \text{possible interactions } + \ \mathbf{e}. \tag{10}$$

The $\mathbf{u}_i$ represent main effects. The i^{th} factor has t_i levels in the population. Under the traditional mixed model for variance components all $t_i \rightarrow$ infinity. In that case $Var(\mathbf{u}_i) = \mathbf{I}\sigma_i^2$ for all i, and all interactions have variance-covariance that are $\mathbf{I}$ times a scalar. Further, all subvectors of $\mathbf{u}_i$ and those subvectors for interactions are mutually uncorrelated.

Now with possible finite t_i

$$Var(\mathbf{u}_i) = \begin{pmatrix} 1 & & -1/(t_i - 1) \\ & \ddots & \\ -1/(t_i - 1) & & 1 \end{pmatrix} \sigma_i^2. \tag{11}$$

This notation denotes one's for diagonals and all off-diagonal elements $= -1/(t_i - 1)$. Now denote by γ_{gh} the interactions between levels of $\mathbf{u}_g$ and $\mathbf{u}_h$. Then there are $t_g t_h$ interactions in the population and the variance-covariance matrix has the following form, where i denotes the level of the g^{th} factor and j the level of the h^{th} factor. The diagonals are $Var(\gamma_{gh})$.

$$\begin{aligned} \text{All elements } ij \text{ with } ij' &= -Var(\gamma_{gh})/(t_h - 1). \\ \text{All elements } ij \text{ with } i'j &= -Var(\gamma_{gh})/(t_g - 1). \\ \text{All elements } ij \text{ with } i'j' &= -Var(\gamma_{gh})/(t_g - 1)(t_h - 1). \end{aligned} \tag{12}$$

i' denotes not equal to i, etc.

To illustrate suppose we have two levels of a first factor and 3 levels of a second. The variance-covariance matrix of

$$\begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \end{pmatrix} = \begin{pmatrix} 1 & -1/2 & -1/2 & -1 & 1/2 & 1/2 \\ & 1 & -1/2 & 1/2 & -1 & 1/2 \\ & & 1 & 1/2 & 1/2 & -1 \\ & & & 1 & -1/2 & -1/2 \\ & & & & 1 & -1/2 \\ & & & & & 1 \end{pmatrix} \sigma^2_{\gamma gh}$$

Suppose that $t_g \rightarrow$ infinity. Then the four types of elements of the variance-covariance matrix would be

$$[1, \ -1/(t_h - 1), \ 0, \ 0] \ \ Var(\gamma_{gh}).$$

This is a model sometimes used for interactions in the two way mixed model with levels of columns fixed.

Now consider 3 factor interactions, γ_{fgh}. Denote by i, j, k the levels of $\mathbf{u}_f$, $\mathbf{u}_g$, and $\mathbf{u}_h$, respectively. The elements of the variance-covariance matrix except for the scalar, $Var(\gamma_{fgh})$ are as follows.

$$\begin{aligned} \text{all diagonals} &= 1. \\ ijk \text{ with } ijk' &= -1/(t_h - 1). \\ ijk \text{ with } ij'k &= -1/(t_g - 1). \\ ijk \text{ with } i'jk &= -1/(t_f - 1). \\ ijk \text{ with } ij'k' &= 1/(t_g - 1)(t_h - 1) \\ ijk \text{ with } i'jk' &= 1/(t_f - 1)(t_h - 1) \\ ijk \text{ with } i'j'k &= 1/(t_f - 1)(t_g - 1) \\ ijk \text{ with } i'j'k' &= 1/(t_f - 1)(t_g - 1)(t_h - 1) \end{aligned} \qquad (13)$$

To illustrate, a mixed model with $\mathbf{u}_g$, $\mathbf{u}_h$ fixed and $t_f \rightarrow$ infinity, the above become 1, $-1/(t_h - 1)$, $-1/(t_g - 1)$, 0, k $1/(t_g - 1)(t_h - 1)$, 0, 0, 0. If levels of all factors $\rightarrow$ infinity, the variance-covariance matrix is $\mathbf{I}Var(\gamma_{fgh})$.

Finally let us look at 4 factor interactions γ_{efgh} with levels of $\mathbf{u}_e$, $\mathbf{u}_f$, $\mathbf{u}_g$, $\mathbf{u}_h$ denoted by i, j, k, m, respectively. Except for the scalar $Var(\gamma_{efgh})$ the variance-covariance matrix has elements like the following.

$$\begin{aligned} \text{all diagonals} &= 1. \\ ijkm \text{ with } ijkm' &= -1/(t_h - 1), \text{ and} \\ ijkm \text{ with } ijk'm &= -1/(t_g - 1), \text{ and} \\ etc. & \end{aligned}$$

$$
\begin{aligned}
ijkm \text{ with } ijk'm' &= 1/(t_g-1)(t_h-1), \text{ and} \\
ijkm \text{ with } ij'km' &= 1/(t_f-1)(t_h-1) \\
&\textit{etc.} \\
ijkm \text{ with } ij'k'm' &= 1/(t_f-1)(t_g-1)(t_h-1) \text{ and} \\
ijkm \text{ with } i'jk'm' &= 1/(t_e-1)(t_g-1)(t_h-1) \\
&\textit{etc.} \\
ijk \text{ with } i'j'k'm' &= 1/(t_e-1)(t_f-1)(t_g-1)(t_h-1).
\end{aligned} \tag{14}
$$

Note that for all interactions the numerator is 1, the denominator is the product of the $t-1$ for subscripts differing, and the sign is plus if the number of differing subscripts is even, and negative if the number of differing subscripts is odd. This set of rules applies to any interactions among any number of factors.

7 Estimability and Predictability

Previous chapters have emphasized the importance of consideration of estimability when $\mathbf{X}$ does not have full column rank, and this is usually the case in application. Now if we apply the same rules given in Chapter 2 for checking estimability and find that an element of $\boldsymbol{\beta}$, eg. μ, is estimable, the resulting estimate can be meaningless in sampling from finite populations. To illustrate suppose we have a model,

$$y_{ij} = \mu + s_i + e_{ij}.$$

Suppose that the s_i represent a random sample of 2 from a finite population of 5 correlated sires. Now $\mathbf{X}$ is a column vector of 1's and consequently μ is estimable by our usual rules. It seems obvious, however, that an estimate of μ has no meaning except as we define the population to which it refers. If we estimate μ by GLS does $\hat{\mu}$ refer to the mean averaged over the 2 sires in the sample or averaged over the 5 sires in the population? Looking at the problem in this manner suggests that we have a problem in prediction. Then the above question can be formulated as two alternatives, namely prediction of $\mu + \frac{1}{5}\sum_{t=1}^{5} s_i$ versus prediction of $\mu + \frac{1}{2}\sum_{i=1}^{2} s_i$, where the second alternative involves summing over the 2 sires in the sample. Of course we could, if we choose, predict $\mu + \mathbf{k}'\mathbf{s}$, where $\mathbf{k}$ is any vector with 5 elements and with $\mathbf{k}'\mathbf{l} = \mathbf{l}$. The variance of $\hat{\mu}$, the GLS estimator or the solution to μ in mixed model equations, is identical to the variance of error of prediction of $\mu + .2\sum_{i=1}^{5} s_i$ and not equal to the variance of error of prediction of $\mu + .5\sum_{i=1}^{2} s_i$. Let us illustrate with some data. Suppose there are 20, 5 observations on sires 1, 2 respectively. Suppose $\mathbf{R} = 50\,\mathbf{I}$ and

$$
\mathbf{G} = \begin{pmatrix}
4 & -1 & -1 & -1 & -1 \\
 & 4 & -1 & -1 & -1 \\
 & & 4 & -1 & -1 \\
 & & & 4 & -1 \\
 & & & & 4
\end{pmatrix}.
$$

Then the mixed model coefficient matrix (not including s_3,s_4,s_5) is

$$\frac{1}{30}\begin{pmatrix} 15 & 12 & 3 \\ & 20 & 2 \\ & & 11 \end{pmatrix}$$

with inverse

$$\frac{1}{9}\begin{pmatrix} 36 & -21 & -6 \\ & 26 & 1 \\ & & 26 \end{pmatrix}.$$

This gives the solution

$$\begin{pmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 6 & 3 \\ 2 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix}.$$

The variance of error of prediction of $\mu + .5(s_1 + s_2)$ is

$$(1\ .5\ .5)\ (\text{Inverse matrix})\ (1\ .5\ .5)' \ = \ 2.5.$$

This is not equal to 4, the variance of $\hat{\mu}$ from the upper left diagonal of the inverse.

Now let us set up equations for BLUP including all 5 sires. Since $\mathbf{G}$ is now singular we need to use one of the methods of Section 5.10. The non-symmetric set of equations is

$$\begin{pmatrix} .5 & .4 & .1 & 0 & 0 & 0 \\ 1.5 & 2.6 & -.1 & 0 & 0 & 0 \\ 0 & -.4 & 1.4 & 0 & 0 & 0 \\ -.5 & -.4 & -.1 & 1. & 0 & 0 \\ -.5 & -.4 & -.1 & 0 & 1. & 0 \\ -.5 & -.4 & -.1 & 0 & 0 & 1. \end{pmatrix}\begin{pmatrix} \hat{\mu} \\ \hat{s} \end{pmatrix} = \begin{pmatrix} .4 & .1 \\ 1.6 & -.1 \\ -.4 & .4 \\ -.4 & -.1 \\ -.4 & -.1 \\ -.4 & -.1 \end{pmatrix}\begin{pmatrix} \bar{y}_{1.} \\ \bar{y}_{2.} \end{pmatrix}.$$

Post-multiplying the inverse of this coefficient matrix by

$$\begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{G} \end{pmatrix}$$

we get as the prediction error variance matrix the following

$$9^{-1}\begin{pmatrix} 36 & -21 & -6 & 9 & 9 & 9 \\ & 26 & 1 & -9 & -9 & -9 \\ & & 26 & -9 & -9 & -9 \\ & & & 36 & -9 & -9 \\ & & & & 36 & -9 \\ & & & & & 36 \end{pmatrix}.$$

The upper 3×3 submatrix is the same as the inverse when only sires 1 and 2 are included. The solution is

$$\begin{pmatrix} \hat{\mu} \\ \hat{s} \end{pmatrix} = 9^{-1} \begin{pmatrix} 6 & 3 \\ 2 & -2 \\ -2 & 2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{y}_{1.} \\ \bar{y}_{2.} \end{pmatrix}.$$

$$\hat{s}_3, \ \hat{s}_4, \ \hat{s}_5 \ = \ 0$$

as would be expected because these sires are unrelated to the 2 with progeny relative to the population of 5 sires. The solution to $\hat{\mu}, \ \hat{s}_1, \ \hat{s}_2$ are the same as before. The prediction error variance of $\mu + .2 \ \sum \ s_i$ is

$$(1 \ .2 \ .2 \ .2 \ .2 \ .2) \text{ (Inverse matrix) } (1 \ .2 \ .2 \ .2 \ .2 \ .2)' = 4,$$

the value of the upper diagonal element of the inverse. By the same reasoning we find that $\hat{s}_j$ is BLUP of $s_j - .2 \sum_{i=1}^{5} s_i$ and not of $s_i - .5 \ (s_1 + s_2)$ for i=1,2. Using the former function with the inverse of the matrix of the second set of equations we obtain for s_1 the value, 2.889. This is also the value of the corresponding diagonal. In contrast the variance of the error of predition of $s_1 - .5 \ (s_1 + s_2)$ is 1.389. Thus $\hat{s}_j$ is the BLUP of $s_j \ - .2 \sum_{i=1}^{5} \ s_i$.

The following rules insure that one does not attempt to predict $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ that is not predictable.

1. $\mathbf{K}'\boldsymbol{\beta}$ must be estimable in a model in which $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$.

2. Pretend that there are no missing classes or subclasses involving all levels of $\mathbf{u}_i$ in the population.

3. Then if $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ is estimable in such a design with $\mathbf{u}$ regarded as fixed, $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ is predictable.

Use the rules of Chapter 2 in checking estimability.

For an example suppose we have sire $\times$ treatment design with 3 treatments and 2 sires regarded as a random sample from an infinite population of possibly related sires. Let the model be

$$\begin{aligned} y_{ijk} &= \mu + s_i + t_j + \gamma_{ij} + e_{ijk}. \\ & \quad \mu, t_j \text{ are fixed} \\ Var(\mathbf{s}) &= \mathbf{I}\sigma_s^2. \end{aligned}$$

Var ($\boldsymbol{\gamma}$) when $\boldsymbol{\gamma}$ are ordered treatments in sires is

$$\begin{pmatrix} \mathbf{B} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B} \end{pmatrix},$$

where

$$\mathbf{B} \ = \ \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix} \sigma^2_\gamma.$$

Suppose we have progeny on all 6 sire $\times$ treatment combinations except (2,3). This creates no problem in prediction due to rule 1 above. Now we can predict for example

$$t_1 \ + \ \sum_{i=1} c_i \ (s_i + \gamma_{i1}) \ - \ t_2 \ - \ \sum_{i=1} d_i \ (s_i + \gamma_{i2})$$

where

$$\sum_i c_i \ = \ \sum_i d_i \ = 1.$$

That is, we can predict the difference between treatments 1 and 2 averaged over <u>any</u> sires in the population, including some not in the sample of 2 sires if we choose to do so. In fact, as we shall see, BLUE of $(t_1 - t_2)$ is BLUP of treatment 1 averaged equally over all sires in the population minus treatment 2 averaged equally over all sires in the population.

Suppose we want to predict the merit of sire 1 versus sire 2. By the rules above, $(s_1 - s_2)$ is not predictable, but

$$s_1 \ + \ \sum_{j=1}^{3} c_j \ (t_j + \gamma_{ij}) \ - \ s_2 \ - \ \sum_{j=1}^{3} d_j \ (t_j + \gamma_{ij})$$

is predictable if $\sum_j \ c_j \ = \ \sum_j \ d_j \ = \ 1$. That is, we can predict sire differences only if we specify treatments, and obviously only treatments 1, 2, 3. We cannot predict unbiasedly, from the data, sire differences associated with some other treatment or treatments. But note that even though subclass (2,3) is missing we can still predict $s_1+t_3+\gamma_{13}-s_2-t_3-\gamma_{23}$. In contrast, if sires as well as treatments were fixed, this function could not be estimated unbiasedly.

8 BLUP When Some $\mathbf{u}_i$ Are Finite

Calculation of BLUE and BLUP when there are finite levels of random factors must take into account the fact that there may be singular $\mathbf{G}$. Consider the simple one way case with a population of 4 related sires. Suppose

$$\mathbf{A} = \begin{pmatrix} 1. & -.2 & -.3 & -.5 \\ & 1. & -.2 & -.6 \\ & & 1 & .5 \\ & & & 1.6 \end{pmatrix}.$$

Suppose we have progeny numbers on these sires that are 9, 5, 3, 0. Suppose the model is

$$\begin{aligned} y_{ijk} &= \mu + s_i + e_{ij}. \\ Var(\mathbf{s}) &= \mathbf{A}\sigma_s^2. \\ Var(\mathbf{e}) &= \mathbf{I}\sigma_e^2. \end{aligned}$$

Then if we wish to include all 4 sires in the mixed model equations we must resort to the methods of Sect. 5.10 since **G** is singular. One of those methods is to solve

$$\left(\begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{A}\sigma_s^2 \end{pmatrix} \begin{pmatrix} 17 & 9 & 5 & 3 & 0 \\ 9 & 9 & 0 & 0 & 0 \\ 5 & 0 & 5 & 0 & 0 \\ 3 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sigma_e^{-2} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \right)$$

$$\begin{pmatrix} \hat{\mu} \\ \hat{s}_1 \\ . \\ . \\ \hat{s}_4 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{A}\sigma_s^2 \end{pmatrix} \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \\ 0 \end{pmatrix} /\sigma_e^2. \qquad (15)$$

$$\hat{\mu} \text{ is BLUP of } \mu + \frac{1}{4} \sum_i s_i.$$

$$\hat{s}_j \text{ is BLUP of } s_j - \frac{1}{4} \sum_i s_i.$$

The inverse of the coefficient matrix post-multiplied by

$$\begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{A}\sigma_s^2 \end{pmatrix}$$

is the variance-covariance matrix of errors of predictions of these functions.

If we had chosen to include only the 3 sires with progeny, the mixed model equations would be

$$\begin{pmatrix} 17 & 9 & 5 & 3 \\ 9 & 9 & 0 & 0 \\ 5 & 0 & 5 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix} \sigma_e^{-2} + \begin{pmatrix} 0 & 0 \quad 0 \quad 0 \\ 0 & \\ 0 & \begin{pmatrix} 1 & -.2 & -.3. \\ -.2 & 1 & -.2 \\ -.3 & -.2 & 1 \end{pmatrix}^{-1} \sigma_s^{-2} \\ 0 & \end{pmatrix}$$

$$
\begin{pmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \end{pmatrix} = \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \end{pmatrix} \sigma_e^{-2}. \tag{16}
$$

This gives the same solution to $\hat{\mu},\ \hat{s}_1,\ \hat{s}_2,\ \hat{s}_3$ as the solution to (15), and the inverse of the coefficient matrix gives the same prediction variances. Even though s_4 is not included, $\hat{\mu}$ predicts $\mu + \frac{1}{4}\sum_{i=1}^4 s_i$, and $\hat{s}_j$ predicts $s_j - \frac{1}{4}\sum_{i=1}^4 s_i$. $\hat{s}_4$ can be computed by

$$
-[.5\ .6\ .5] \begin{pmatrix} 1 & -.2 & -.3 \\ -.2 & 1 & -.2 \\ -.3 & -.2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \end{pmatrix}.
$$

As another example suppose we have a sire by treatment model with an infinite population of sires. The n_{ij} are

	1	2
1	0	8
2	9	2
3	6	0

$$
\text{Var}\ (\mathbf{s}) = 2\mathbf{I},\ \text{Var}\ (\mathbf{e}) = 10\ \mathbf{I},
$$

Var (γ) including missing subclasses is

$$
\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & & 1 & -1 & 0 & 0 \\ & & & 1 & 0 & 0 \\ & & & & 1 & -1 \\ & & & & & 1 \end{pmatrix} /2.
$$

If we do not include γ_{11} and γ_{32} in the solution the only submatrix of $\mathbf{G}$ that is singular is the 2x2 block pertaining to $\gamma_{21},\ \gamma_{22}$. The GLS equations regarding $\mathbf{u}$ as fixed are

$$
\frac{1}{10} \begin{pmatrix} 8 & 0 & 0 & 0 & 8 & 8 & 0 & 0 & 0 \\ & 11 & 0 & 9 & 2 & 0 & 9 & 2 & 0 \\ & & 6 & 6 & 0 & 0 & 0 & 0 & 6 \\ & & & 15 & 0 & 0 & 9 & 0 & 6 \\ & & & & 10 & 8 & 0 & 2 & 0 \\ & & & & & 8 & 0 & 0 & 0 \\ & & & & & & 9 & 0 & 0 \\ & & & & & & & 2 & 0 \\ & & & & & & & & 6 \end{pmatrix} \begin{pmatrix} \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{t}_1 \\ \hat{t}_2 \\ \hat{\gamma}_{12} \\ \hat{\gamma}_{21} \\ \hat{\gamma}_{22} \\ \hat{\gamma}_{31} \end{pmatrix} = \begin{pmatrix} y_{1..} \\ y_{2..} \\ y_{3..} \\ y_{.1.} \\ y_{.2.} \\ y_{12.} \\ y_{21.} \\ y_{22.} \\ y_{31.} \end{pmatrix} \frac{1}{10} \tag{17}
$$

Then we premultiply the 7^{th} and 8^{th} equations of (17) by

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} /2$$

and add to the diagonal coefficients, (.5, .5, .5, 0, 0, 2, 1, 1, 2). The solution to the resulting equations is BLUP. If we had included γ_{12} and γ_{32}, we would premultiply the last 6 GLS equations (equations for γ) by Var (γ) and then add to the diagonals, (.5, .5, .5, 0, 0, 1, 1, 1, 1, 1, 1). When all elements of a population are included in a BLUP solution, an interesting property becomes apparent. The same summing to 0's occurs in the BLUP solution as is true in the corresponding elements of the finite populations described in Section 4.

9 An Easier Computational Method

Finite populations complicate computation of BLUE and BLUP because non-diagonal and singular **G** matrices exist. But if the model is that of Section 2, that is, finite populations of unrelated elements with common variance, computations can be carried out with diagonal submatrices for **G**. The resulting $\hat{\mathbf{u}}$ do not always predict the same functions predicted by using the actual **G** matrices, but appropriate linear functions of them do. We illustrate with a simple one way case.

$$\begin{aligned} y_{ij} &= \mu + a_i + e_{ij}. \quad i = 1, 2, 3. \\ Var(\mathbf{a}) &= \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}, \\ Var(\mathbf{e}) &= 10\mathbf{I}. \\ n_i &= (5, 3, 2), y_{i.} = (10, 8, 6). \end{aligned}$$

Using singular **G** the nonsymmetric mixed model equations are

$$\begin{pmatrix} 1. & .5 & .3 & .2 \\ .5 & 2. & -.3 & -.2 \\ -.1 & -.5 & 1.6 & -.2 \\ -.4 & -.5 & -.3 & 1.4 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = \begin{pmatrix} 2.4 \\ .6 \\ 0 \\ -.6 \end{pmatrix}. \tag{18}$$

The solution is [2.4768, -.2861, .0899, .1962]. Note that

$$\sum \hat{u}_i = 0.$$

We can obtain the same solution by pretending that $Var(\mathbf{a}) = 3\mathbf{I}$. Then the mixed model equations are

$$\begin{pmatrix} 1 & .5 & .3 & .2 \\ .5 & .5+3^{-1} & 0 & 0 \\ .3 & 0 & .3+3^{-1} & 0 \\ .2 & 0 & 0 & .2+3^{-1} \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = \begin{pmatrix} 2.4 \\ 1.0 \\ .8 \\ .6 \end{pmatrix} \tag{19}$$

The inverse of (15.19) is different from the inverse of (15.18) post-multiplied by

$$\begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{G} \end{pmatrix}.$$

The inverse of (19) does not yield prediction error variances. To obtain prediction error variances of $\mu + \bar{a}_.$ and of $a_i - \bar{a}_.$ pre-multiply it by

$$\frac{1}{3}\begin{pmatrix} 3 & 1 & 1 & 1 \\ 0 & 2 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}$$

and post-multiply that product by the transpose of this matrix. This is a consequence of the fact that the solution to (19) is BLUP of

$$\frac{1}{3}\begin{pmatrix} 3 & 1 & 1 & 1 \\ 0 & 2 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} \mu \\ s_1 \\ s_2 \\ s_3 \end{pmatrix}.$$

In most cases use of diagonal $\mathbf{G}$ does not result in the same solution as using the true $\mathbf{G}$, and the inverse never yields directly the prediction error variance-covariance matrix.

Rules for deriving diagonal submatrices of $\mathbf{G}$ to use in place of singular submatrices follow. For main effects say of $\mathbf{u}_i$ with t_i levels substitute for the $\mathbf{G}$ submatrix described in Section 6, $\mathbf{I}\sigma^2_{*i}$, where

$$\sigma^2_{*i} = \frac{t_i}{t_{i-1}}\sigma^2_i - \sum_j \frac{t_i}{(t_{i-1})(t_{j-1})}\sigma^2_{ij} + \sum_{j,k}\frac{t_i}{(t_{i-1})(t_{j-1})(t_{k-1})}\sigma^2_{ijk}$$
$$- \sum_{j,k,m}\frac{t_i}{(t_{i-1})(t_{j-1})(t_{k-1})(t_{m-1})}\sigma^2_{ijkm} \text{ etc.}$$
$$\text{for 5 factor, 6 factor interactions.} \tag{20}$$

σ^2_i refers to the scalar part of the variance of the i^{th} factor, σ^2_{ij} refers to 2 factor interactions involving $\mathbf{u}_i$, σ^2_{ijk} refers to 3 factor interactions involving $\mathbf{u}_i$, etc. Note that the signs alternate

$$\sigma^2_{*ij} = \frac{t_i\ t_j}{(t_{i-1})(t_{j-1})}\sigma^2_{ij} - \sum_k \frac{t_i\ t_j}{(t_{i-1})(t_{j-1})(t_{k-1})}\sigma^2_{ijk}$$

$$+ \sum_{k,m} \frac{t_i \ t_j}{(t_{i-1})(t_{j-1})(t_{k-1})(t_{m-1})} \ \sigma^2_{ijkm} \ \text{ etc.} \tag{21}$$

$$\sigma^2_{*ijk} \ = \frac{t_i \ t_j \ t_k}{(t_{i-1})(t_{j-1})(t_{k-1})} \ \sigma^2_{ijk} \ - \ \sum_m \frac{t_i \ t_j \ t_k}{(t_{i-1})(t_{j-1})(t_{k-1})(t_{m-1})}$$

$$\sigma^2_{ijk} \ + \ \text{etc.} \tag{22}$$

Higher order interactions for σ^2_* follow this same pattern with alternating signs. The sign is positive when the number of factors in the denominator minus the number in the numerator is even.

It appears superficially that one needs to estimate the different σ^2_i, σ^2_{ij}, σ^2_{ijk}, etc., and this is difficult because non-diagonal, singular submatrices of **G** are involved. But if one plans to use their diagonal representations, one might as well estimate the σ^2_* directly by any of the standard procedures for the conventional mixed model for variance components estimation. Then if for pedagogical or other reasons one wishes estimates of σ^2 rather than σ^2_*, one can use equations (20), (21), (22) that relate the two to affect the required linear transformation.

The solution using diagonal **G** should not be assumed to be the same as would have been obtained from use of the true **G** matrix. If we consider predictable functions as defined in Section 7 and take these same functions of the solution using diagonal **G** we do obtain BLUP. Similarly using these functions we can derive prediction error variances using a g-inverse of the coefficient matrix with diagonal **G**.

10 Biased Estimation

If we can legitimately assume that there is no expected pattern of values of the levels of a fixed factor and no expected pattern of values of interactions between levels of fixed factors, we can pretend that these fixed factors and interactions are populations with finite levels and proceed to compute biased estimators as though we are computing BLUP of random variables. Instead of prediction error variance as derived from the g-inverse of the coefficient matrix we obtain estimated mean squared errors.

Chapter 16
The One-Way Classification

C. R. Henderson

1984 - Guelph

This and subsequent chapters will illustrate principles of Chapter 1-15 as applied to specific designs and classification of data. This chapter is concerned with a model,

$$y_{ij} = \mu + a_i + e_{ij}. \tag{1}$$

Thus data can be classified with n_i observations on the i^{th} class and with the total of observations in that class $= y_{i.}$. Now (1) is not really a model until we specify what population or populations were sampled and what are the properties of these populations. One possibility is that in conceptual repeated sampling μ and a_i always have the same values, and the e_{ij} are random samples from an infinite population of uncorrelated variables with mean 0, and common variance, σ_e^2. That is, the variance of the population of $\mathbf{e}$ is $\mathbf{I}\sigma_e^2$, and the sample vector of n elements has expectation null and variance $= \mathbf{I}\sigma_e^2$. Note that $\text{Var}(e_{ij})$ is assumed equal to $\text{Var}(e_{i'j})$, $i \neq i'$.

1 Estimation and Tests For Fixed a

Estimation and tests of hypothesis are simple under this model. The mixed model equations are OLS equations since $\mathbf{Zu}$ does not exist and since $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$. They are

$$\frac{1}{\sigma_e^2}\begin{pmatrix} n_. & n_1 & n_2 & \ldots \\ n_1 & n_1 & 0 & \ldots \\ n_2 & 0 & n_2 & \ldots \\ \vdots & \vdots & \vdots & \end{pmatrix}\begin{pmatrix} \mu^o \\ a_1^o \\ a_2^o \\ \vdots \end{pmatrix} = \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ \vdots \end{pmatrix}\frac{1}{\sigma_e^2}. \tag{2}$$

The $\mathbf{X}$ matrix has $t+1$ columns, where $t =$ the number of levels of a, but the rank is t. None of the elements of the model is estimable. We can estimate

$$\mu + \sum_{i=1}^{t} k_i a_i,$$

where

$$\sum_i k_i = 1,$$

or

$$\sum_{i}^{t} k_i a_i,$$

if

$$\sum_{i} k_i = 0.$$

For example $\mu + a_i$ is estimable, $a_i - a_{i'}$ is estimable, and

$$a_1 - \sum_{i=2}^{t} k_i a_i,$$

with

$$\sum_{i=2}^{t} k_i = 1,$$

is estimable. The simplest solution to (2) is $\mu^o = 0$, $a_i^o = \overline{y}_{i.}$. This solution corresponds to the following g-inverse.

$$\begin{pmatrix} 0 & 0 & 0 & \ldots \\ 0 & n_1^{-1} & 0 & \ldots \\ 0 & 0 & n_2^{-1} & \ldots \\ \vdots & \vdots & \vdots & \end{pmatrix}.$$

Let us illustrate with the following example

$$\begin{aligned} (n_1,\ n_2,\ n_3) &= (8, 3, 4), \\ (y_{1.},\ y_{2.},\ y_{3.}) &= (49, 16, 13), \\ \mathbf{y}'\mathbf{y} &= 468. \end{aligned}$$

The OLS equations are

$$\frac{1}{\sigma_e^2} \begin{pmatrix} 15 & 8 & 3 & 4 \\ & 8 & 0 & 0 \\ & & 3 & 0 \\ & & & 4 \end{pmatrix} \begin{pmatrix} \mu^o \\ a_1^o \\ a_2^o \\ a_3^o \end{pmatrix} = \begin{pmatrix} 78 \\ 49 \\ 16 \\ 13 \end{pmatrix} \frac{1}{\sigma_e^2}. \qquad (3)$$

A solution is (0, 49/8, 16/3, 13/4). The corresponding g-inverse of the coefficient matrix is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ & 8^{-1} & 0 & 0 \\ & & 3^{-1} & 0 \\ & & & 4^{-1} \end{pmatrix} \sigma_e^2.$$

Suppose one wishes to estimate $a_1 - a_2$, $a_1 - a_3$, $a_2 - a_3$. Then from the above solution these would be $\frac{49}{8} - \frac{16}{3}$, $\frac{49}{8} - \frac{13}{4}$, $\frac{16}{3} - \frac{13}{4}$. The variance-covariance matrix of these estimators is

$$\begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 8^{-1} & 0 & 0 \\ 0 & 0 & 3^{-1} & 0 \\ 0 & 0 & 0 & 4^{-1} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \end{pmatrix} \sigma_e^2 . \qquad (4)$$

We do not know σ_e^2 but it can be estimated easily by

$$\begin{aligned} \hat{\sigma}_e^2 &= (\mathbf{y}'\mathbf{y} - \sum_i y_{i.}^2/n_i)/(15-3) \\ &= (468 \ - \ 427.708)/12 \\ &= 3.36. \end{aligned}$$

Then we can substitute this for σ_e^2 to obtain estimated sampling variances.

Suppose we want to test the hypothesis that the levels of a_i are equal. This can be expressed as a test that

$$\begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

$$Var(\mathbf{K}'\beta^o) = \mathbf{K}'(\text{g} - \text{inverse})\mathbf{K} = \begin{pmatrix} .45833 & .125 \\ .125 & .375 \end{pmatrix} \sigma_e^2$$

with

$$\text{inverse} = \begin{pmatrix} 2.4 & -.8 \\ -.8 & 2.9333 \end{pmatrix} \frac{1}{\sigma_e^2} .$$

$$\mathbf{K}'\boldsymbol{\beta}^o = (.79167 \ \ 2.875)'.$$

Then

$$\begin{aligned} \text{numerator SS} &= (.79167 \ \ 2.875) \begin{pmatrix} 2.4 & -.8 \\ & 2.9333 \end{pmatrix} \begin{pmatrix} .79167 \\ 2.875 \end{pmatrix} \\ &= 22.108. \end{aligned}$$

The same numerator can be computed from

$$\sum_i \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{n_.} = 427.708 \ - \ 405.6 = 22.108.$$

Then the test that a_i are equal is $\frac{22.108/2}{3.36}$ which is distributed as $F_{2,12}$ under the null hypothesis.

2 Levels of a Equally Spaced

In some experiments the levels of **a** (treatments) are chosen to be "equally spaced". For example, if treatments are percent protein in the diet, the levels chosen might be 10%, 12%, 14%, 16%, 18%. Suppose we have 5 such treatments with n_i = (5,2,1,3,8) and $y_{i.}$ = (10,7,3,8,33). Let the full model be

$$y_{ij} = \mu + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \beta_4 x_i^4 + e_{ij} \tag{5}$$

where x_i = (1,2,3,4,5). With Var(e) = $\mathbf{I}\sigma^2$ the OLS equations under the full model are

$$\begin{pmatrix} 19 & 64 & 270 & 1240 & 5886 \\ & 270 & 1240 & 5886 & 28,384 \\ & & 5886 & 28,384 & 138,150 \\ & & & 138,150 & 676,600 \\ & & & & 3,328,686 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{pmatrix} = \begin{pmatrix} 61 \\ 230 \\ 1018 \\ 4784 \\ 23,038 \end{pmatrix}. \tag{6}$$

The solution is [-4.20833, 9.60069, -3.95660, .58681, -.02257]. The reduction in SS is 210.958 which is exactly the same as $\sum_i y_{i.}^2/n_i$. A common set of tests is the following.

$\beta_1 = 0$ assuming β_2, β_3, β_4 non-existent.

$\beta_2 = 0$ assuming β_3, β_4 non-existent.

$\beta_3 = 0$ assuming β_4 non-existent.

$\beta_4 = 0$.

This can be done by computing the following reductions.

1. Red (full model).

2. Red (μ, β_1, β_2, β_3).

3. Red (μ, β_1, β_2).

4. Red (μ, β_1).

5. Red (μ).

Then the numerators for tests above are reductions 4-5, 3-4, 2-3, 1-2 respectively.

Red (2) is obtained by dropping the last equation of (6). This gives the solution (-3.2507, 7.7707, -2.8325, .3147) with reduction = 210.952. The other reductions by successive dropping of an equation are 207.011, 206.896, 195.842. This leads to mean squares each with 1 df.

Linear	11.054
Quadratic	.115
Cubic	3.941
Quartic	.006

The sum of these is equal to the reduction under the full model minus the reduction due to μ alone.

3 Biased Estimation of $\mu + a_i$

Now we consider biased estimation under the assumption that values of **a** are unpatterned. Using the same data as in the previous section we assume for purposes of illustration that $Var(\mathbf{e}) = \frac{5}{6}\mathbf{I}$, and that the average values of squares and products of the deviations of **a** from $\bar{\mathbf{a}}$ are

$$\frac{1}{8}\begin{pmatrix} 4 & -1 & -1 & -1 & -1 \\ & 4 & -1 & -1 & -1 \\ & & 4 & -1 & -1 \\ & & & 4 & -1 \\ & & & & 4 \end{pmatrix}. \tag{7}$$

Then the equations for minimum mean squared error estimation are

$$\begin{pmatrix} 22.8 & 6.0 & 2.4 & 1.2 & 3.6 & 9.6 \\ .9 & 4.0 & -.3 & -.15 & -.45 & -1.2 \\ -1.35 & -.75 & 2.2 & -.15 & -.45 & -1.2 \\ -2.1 & -.75 & -.3 & 1.6 & -.45 & -1.2 \\ -.6 & -.75 & -.3 & -.15 & 2.8 & -1.2 \\ 3.15 & -.75 & -.3 & -.15 & -.45 & 5.8 \end{pmatrix} (\boldsymbol{\beta}^o) = \begin{pmatrix} 73.2 \\ -1.65 \\ -3.9 \\ -6.9 \\ -3.15 \\ 15.6 \end{pmatrix}. \tag{8}$$

The solution is (3.072, -.847, .257, -.031, -.281, .902). Note that

$$\sum_i \hat{a}_i = 0.$$

The estimates of differences between a_i are

	2	3	4	5
1	-1.103	-.816	-.566	-1.749
2		.288	.538	-.645
3			.250	-.933
4				-1.183

Contrast these with the corresponding BLUE. These are

	2	3	4	5
1	-1.5	-1.0	-.667	-2.125
2		.5	.833	-.625
3			.333	-1.125
4				-1.458

Generally the absolute differences are larger for BLUE.

The mean squared error of these differences, assuming that σ_e^2 and products of deviations of **a** are correct, are obtained from a g-inverse post-multiplied by

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & -1 & -1 & -1 & -1 \\ 0 & -1 & 4 & -1 & -1 & -1 \\ 0 & -1 & -1 & 4 & -1 & -1 \\ 0 & -1 & -1 & -1 & 4 & -1 \\ 0 & -1 & -1 & -1 & -1 & 4 \end{pmatrix}.$$

These are

	2	3	4	5
1	.388	.513	.326	.222
2		.613	.444	.352
3			.562	.480
4				.287

The corresponding values for BLUE are

	2	3	4	5
1	.7	1.2	.533	.325
2		1.5	.833	.625
3			1.333	1.125
4				.458

If the priors used are really correct, the MSE for biased estimators of differences are considerably smaller than BLUE.

The same biased estimators can be obtained by use of a diagonal **P**, namely .625**I**, where

$$.625 = \frac{5}{5-1}\,(.5).$$

This gives the same solution vector, but the inverse elements are different. However, mean squared errors of estimable functions such as the $\hat{a}_i - \hat{a}_j$ and $\hat{\mu} + \overline{a}$ yield the same results when applied to the inverse.

4 Model with Linear Trend of Fixed Levels of a

Assume now the same data as section 2 and that the model is

$$y_{ij} = \mu + \beta x_i + a_i + e_{ij} \tag{9}$$

where x_i = (1,2,3,4,5). Suppose that the levels of **a** are assumed to have no pattern and we use a prior value on their squares and products =

$$\begin{pmatrix} .2 & & -.05 \\ & \ddots & \\ -.05 & & 2 \end{pmatrix}.$$

Assume as before Var($\mathbf{e}$) = $\frac{5}{6}\,\mathbf{I}$. Then the equations to solve are

$$\begin{pmatrix} 22.8 & 76.8 & 6. & 2.4 & 1.2 & 3.6 & 9.6 \\ 76.8 & 324 & 6 & 4.8 & 3.6 & 14.4 & 48 \\ .36 & -2.34 & 2.2 & -.12 & -.06 & -.18 & -.48 \\ -.54 & -2.64 & -.3 & 1.48 & -.06 & -.18 & -.48 \\ -.84 & -2.94 & -.3 & -.12 & 1.24 & -.18 & -.48 \\ -.24 & -.24 & -.3 & -.12 & -.06 & 1.72 & -.48 \\ 1.26 & 8.16 & -.3 & -.12 & -.06 & -.18 & 2.92 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{\beta} \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \end{pmatrix} = \begin{pmatrix} 73.2 \\ 276 \\ -.66 \\ -1.56 \\ -2.76 \\ -1.26 \\ 6.24 \end{pmatrix}. \tag{10}$$

The solution is [1.841, .400, -.145, .322, -.010, -.367, .200]. Note that

$$\sum_i \hat{a}_i = 0.$$

We need to observe precautions in interpreting the solution. β is not estimable and neither is $\mu + a_i$ nor $a_i - a_{i'}$.

We can only estimate treatment means associated with the particular level of x_i in the experiment. Thus we can estimate $\mu + a_i + x_i\beta$ where x_i = 1,2,3,4,5 for the 5 treatments respectively. The biased estimates of treatment means are

1. 1.841 + .400 - .145 = 2.096

2. 1.841 + .800 + .322 = 2.963

3. 1.841 + 1.200 - .010 = 3.031

4. 1.841 + 1.600 - .367 = 3.074

5. 1.841 + 2.000 + .200 = 4.041

The corresponding BLUE are the treatment means, (2.0, 3.5, 3.0, 2.667, 4.125).

If the true ratio of squares and products of a_i to σ_e^2 are as assumed above, the biased estimators have minimum mean squared error. Note that $E(\hat{\mu} + \hat{a}_i + x_i\hat{\beta})$ for the biased estimator is $\mu + x_i\beta +$ some function of **a** (not equal to a_i). The BLUE estimator has, of course, expectation, $\mu + x_i\beta + a_i$, that is, it is unbiased.

5 The Usual One Way Covariate Model

If, in contrast to x_i being constant for every observation on the i^{th} treatment as in Section 4, we have the more traditional covariate model,

$$y_{ij} = \mu + \beta x_{ij} + a_i + e_{ij}, \tag{11}$$

we can then estimate $\mu + a_i$ unbiasedly as well as $a_i - a_{i'}$. Again, however, if we think the a_i are unpatterned and we have some good prior value of their products, we can obtain smaller mean squared errors by using the biased method.

Now we need to consider the meaning of an estimator of $\mu + a_i$. This really is an estimator of treatment mean in hypothetical repeated sampling in which $\overline{x}_{i.} = 0$. What if the range of the x_{ij} is 5 to 21 in the sample? Can we infer from this that the the response to levels of x is that same linear function for a range of x_{ij} as low as 0? Strictly speaking we can draw inferences only for the values of x in the experiment. With this in mind we should really estimate $\mu + a_i + k\beta$, where k is some value in the range of x's in the experiment. With regard to treatment differences, $a_i - a_{i'}$, can be regarded as an estimate of $(\mu + a_i + k\beta) - (\mu + a_{i'} + k\beta)$, where k is in the range of the x's of the experiment.

6 Nonhomogenous Regressions

A still different covariate model is

$$y_{ij} = \mu + \beta_i x_{ij} + a_i + e_{ij}.$$

Note that in this model β is different from treatment to treatment. According to the rules for estimability $\mu + a_i$, $a_i - a_{i'}$, and β_i are all estimable. However, it is now obvious that $a_i - a_{i'}$ has no practical meaning as an estimate of treatment difference. We must specify what levels of x we assume to be present for each treatment. In terms of a treatment mean these are

$$\mu + a_i + k_i \beta_i$$

and

$$\mu + a_j + k_j \beta_j$$

and the difference is

$$a_i + k_i \beta_i - a_j - k_j \beta_j.$$

Suppose $k_i = k_j = k$. Then the treatment difference is

$$a_i - a_j + k(\beta_i - \beta_j),$$

and this is not invariant to the choice of k when $\beta_i \neq \beta_j$. In contrast when all $\beta_i = \beta$, the treatment difference is invariant to the choice of k.

Let us illustrate with two treatments.

Treatment	n_i	$y_{i.}$	$x_{i.}$	$\sum_j x_{ij}^2$	$\sum_j x_{ij}y_{ij}$
1	8	38	36	220	219
2	5	43	25	135	208

This gives least squares equations

$$\begin{pmatrix} 8 & 0 & 36 & 0 \\ & 5 & 0 & 25 \\ & & 220 & 0 \\ & & & 135 \end{pmatrix} \begin{pmatrix} \hat{\mu} + \hat{t}_1 \\ \hat{\mu} + \hat{t}_2 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} 38 \\ 43 \\ 219 \\ 208 \end{pmatrix}.$$

The solution is (1.0259, 12.1, .8276, -.7). Then the estimated difference, treatment 1 minus treatment 2 for various values for x, the same for each treatment, are as follows

x	Estimated Difference
0	-11.07
2	-8.02
4	-4.96
6	-1.91
8	1.15
10	4.20
12	7.26

It is obvious from this example that treatment differences are very sensitive to the average value of x.

7 The Usual One Way Random Model

Next we consider a model

$$\begin{aligned} y &= \mu + a_i + e_{ij}. \\ Var(\mathbf{a}) &= \mathbf{I}\sigma_a^2, \\ Var(\mathbf{e}) &= \mathbf{I}\sigma_e^2, \\ Cov(\mathbf{a}, \mathbf{e}') &= \mathbf{0}. \end{aligned}$$

In this case it is assumed that the levels of **a** in the sample are a random sample from an infinite population with var $\mathbf{I}\sigma_a^2$, and similarly for the sample of **e**. The experiment may have been conducted to do one of several things, estimate μ, predict **a**, or to estimate σ_a^2 and σ_e^2. We illustrate these with the following data.

Levels of a	n_i	$y_{i.}$
1	5	10
2	2	7
3	1	3
4	3	8
5	8	33

Let us estimate μ and predict $\mathbf{a}$ under the assumption that $\sigma_e^2/\sigma_a^2 = 10$. Then we need to solve these equations.

$$\begin{pmatrix} 19 & 5 & 2 & 1 & 3 & 8 \\ & 15 & 0 & 0 & 0 & 0 \\ & & 12 & 0 & 0 & 0 \\ & & & 11 & 0 & 0 \\ & & & & 13 & 0 \\ & & & & & 18 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \end{pmatrix} = \begin{pmatrix} 61 \\ 10 \\ 7 \\ 3 \\ 8 \\ 33 \end{pmatrix}. \quad (12)$$

The solution is [3.137, -.379, .061, -.012, -.108, .439]. Note that $\sum \hat{a}_i = 0$. This could have been anticipated by noting that the sum of the last 4 equations minus the first equation gives

$$10 \sum \hat{a}_i = 0.$$

The inverse of the coefficient matrix is

$$\begin{pmatrix} .0790 & -.0263 & -.0132 & -.0072 & -.0182 & -.0351 \\ & .0754 & .0044 & .0024 & .0061 & .0117 \\ & & .0855 & .0012 & .0030 & .0059 \\ & & & .0916 & .0017 & .0032 \\ & & & & .0811 & .0081 \\ & & & & & .0712 \end{pmatrix}. \quad (13)$$

This matrix premultiplied by (0 1 1 1 1 1) equals (-1 1 1 1 1 1)(σ_a^2/σ_e^2). This is always a check on the inverse of the coefficient matrix in a model of this kind. From the inverse

$$\begin{aligned} Var(\hat{\mu}) &= .0790\, \sigma_e^2, \\ Var(\hat{a}_1 - a_1) &= .0754\, \sigma_e^2. \end{aligned}$$

$\hat{\mu}$ is BLUP of μ + the mean of all a in the infinite population. Similarly $\hat{a}_i$ is BLUP of a_i minus the mean of all a_i in the infinite population.

Let us estimate σ_a^2 by Method 1. For this we need $\sum_i y_{i.}^2/n_i$ and $y_{..}^2/n_.$ and their expectations. These are 210.9583 and 195.8421 with expectations, $19\sigma_a^2 + 5\sigma_e^2$ and $5.4211\sigma_a^2 + \sigma_e^2$ respectively ignoring 19 μ^2 in both.

$$\hat{\sigma}_e^2 = (\mathbf{y}'\mathbf{y} - 210.9583)/(19 - 5).$$

Suppose this is 2.8. Then $\hat{\sigma}_a^2 = .288$.

Let us next compute an approximate MIVQUE estimate using the prior $\sigma_e^2/\sigma_a^2 = 10$, the ratio used in the BLUP solution. We shall use $\hat{\sigma}_e^2 = 2.8$ from the least squares residual rather than a MIVQUE estimate. Then we need to compute $\hat{\mathbf{a}}'\hat{\mathbf{a}} = .35209$ and its expectation. The expectation is $trVar(\hat{\mathbf{a}})$. But $Var(\hat{\mathbf{a}}) = \mathbf{C}_a Var(\mathbf{r})\mathbf{C}_a'$, where $\mathbf{C}_a$ is the last 5 rows of the inverse of the mixed model equations (12), and $\mathbf{r}$ is the vector of right hand sides.

$$
\begin{aligned}
Var(\mathbf{r}) \;=\; & \begin{pmatrix} 5 & 2 & 1 & 3 & 8 \\ 5 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 & 3 & 8 \\ 5 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix}' \sigma_a^2 + \\
& \begin{pmatrix} 19 & 5 & 2 & 1 & 3 & 8 \\ & 5 & 0 & 0 & 0 & 0 \\ & & 2 & 0 & 0 & 0 \\ & & & 1 & 0 & 0 \\ & & & & 3 & 0 \\ & & & & & 8 \end{pmatrix} \sigma_e^2 \, .
\end{aligned}
$$

This gives

$$E(\hat{\mathbf{a}}'\hat{\mathbf{a}}) = .27163\,\sigma_a^2 + .06802\,\sigma_e^2,$$

and using $\hat{\sigma}_e^2 = 2.8$, we obtain $\hat{\sigma}_a^2 = .595$.

8 Finite Levels of a

Suppose now that the five a_i in the sample of our example of Section 7 comprise all of the elements of the population and that they are unrelated. Then

$$Var(\mathbf{a}) = \begin{pmatrix} 1 & & -.25 \\ & \ddots & \\ -.25 & & 1 \end{pmatrix} \sigma_a^2 \, .$$

Let us assume that $\sigma_e^2/\sigma_a^2 = 12.5$. Then the mixed model equations are the OLS equations premultiplied by

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ & .08 & -.02 & -.02 & -.02 & -.02 \\ & & .08 & -.02 & -.02 & -.02 \\ & & & .08 & -.02 & -.02 \\ & & & & .08 & -.02 \\ & & & & & .08 \end{pmatrix} . \tag{14}$$

This gives the same solution as that to (11). This is because σ_a^2 of the infinite model is $\frac{5}{4}$ times σ_a^2 of the finite model. See Section 15.9. Now $\hat{\mu}$ is a predictor of

$$\mu + \frac{1}{5}\sum_i a_i$$

and $\hat{a}_j$ is a predictor of

$$a_j - \frac{1}{5}\sum_i a_i.$$

Let us find the Method 1 estimate of σ_a^2 in the finite model. Again we compute $\sum_i y_{i.}^2/n_i$ and $y_{..}^2/n_{..}$. Then the coefficient of σ_e^2 in each of these is the same as in the infinite model, that is 5 and 1 respectively. For the coefficients of σ_a^2 we need the contribution of σ_a^2 to Var(rhs). This is

$$\begin{pmatrix} 5 & 2 & 1 & 3 & 8 \\ 5 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} 1 & & -\frac{1}{4} \\ & \ddots & \\ -\frac{1}{4} & & 1 \end{pmatrix} \text{(left matrix)}'$$

$$= \begin{pmatrix} 38.5 & 7.5 & -4.5 & -3.5 & -3.0 & 42. \\ & 25.0 & -2.5 & -1.25 & -3.75 & -10. \\ & & 4.0 & -.5 & -1.5 & -4. \\ & & & 1. & -.75 & -2. \\ & & & & 9. & -6. \\ & & & & & 64. \end{pmatrix}. \quad (15)$$

Then the coefficient of σ_a^2 in $\sum_i y_{i.}^2/n_i$ is $tr[dg(5,2,1,3,8)]^{-1}$ times the lower 5×5 submatrix of (15) = 19.0. The coefficient of σ_a^2 in $y_{..}^2/n_{..} = 38.5/19 = 2.0263$. Thus we need only the diagonals of (15). Assuming again that $\sigma_e^2 = 2.8$, we find $\hat{\sigma}_a^2 = .231$. Note that in the infinite model $\hat{\sigma}_a^2 = .288$ and that $\frac{5}{4}(.231) = .288$ except for rounding error. This demonstrates that we could estimate σ_a^2 as though we had an infinite model and estimate μ and predict **a** using $\hat{\sigma}_a^2/\hat{\sigma}_e^2$ in mixed model equations for the infinite model. Remember that the resulting inverse does not yield directly $Var(\hat{\mu})$ and $Var(\hat{\mathbf{a}} - \mathbf{a})$. For this pre- and post-multiply the inverse by

$$\frac{1}{5}\begin{pmatrix} 5 & 1 & 1 & 1 & 1 & 1 \\ 0 & 4 & -1 & -1 & -1 & -1 \\ 0 & -1 & 4 & -1 & -1 & -1 \\ 0 & -1 & -1 & 4 & -1 & -1 \\ 0 & -1 & -1 & -1 & 4 & -1 \\ 0 & -1 & -1 & -1 & -1 & 4 \end{pmatrix}.$$

This is in accord with the idea that in the finite model $\hat{\mu}$ is BLUP of $\mu + \overline{a}.$ and $\hat{a}_i$ is BLUP of $a_i - \overline{a}.$.

9 One Way Random and Related Sires

We illustrate the use of the numerator relationship matrix in evaluating sires in a simple one way model,

$$\begin{aligned} y_{ij} &= \mu + s_i + e_{ij}. \\ Var(\mathbf{s}) &= \mathbf{A}\sigma_s^2, \\ Var(\mathbf{e}) &= \mathbf{I}\sigma_e^2, \\ Cov(\mathbf{s}, \mathbf{e}') &= \mathbf{0}, \\ \sigma_e^2/\sigma_s^2 &= 10. \end{aligned}$$

Then mixed model equations for estimation of μ and prediction of $\mathbf{s}$ are

$$\left(\begin{pmatrix} n_{.} & n_{1.} & n_{2.} & \ldots \\ n_{1.} & n_{1.} & 0 & \ldots \\ n_{2.} & 0 & & \\ \vdots & \vdots & & \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & \ldots \\ 0 & & \mathbf{A}^{-1} & \sigma_e^2/\sigma_s^2 \\ 0 & & & \\ \vdots & & & \end{pmatrix} \right)$$

$$\begin{pmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \vdots \end{pmatrix} = (y_{..} \ \ y_{1.} \ \ y_{2.} \ \ \ldots)' \ . \tag{16}$$

We illustrate with the numerical example of section 7 but now with

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & .5 & .5 & 0 \\ & 1. & 0 & 0 & .5 \\ & & 1. & .25 & 0 \\ & & & 1 & 0 \\ & & & & 1 \end{pmatrix} .$$

The resulting mixed model equations are

$$\begin{pmatrix} 19 & 5 & 2 & 1 & 3 & 8 \\ & 65/3 & 0 & -20/3 & -20/3 & 0 \\ & & 46/3 & 0 & 0 & -20/3 \\ & & & 43/3 & 0 & 0 \\ & & & & 49/3 & 0 \\ & & & & & 64/3 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix} = \begin{pmatrix} 61 \\ 10 \\ 7 \\ 3 \\ 8 \\ 33 \end{pmatrix} . \tag{17}$$

The solution is (3.163, -.410, .232, -.202, -.259, .433). Note that $\sum_i \hat{s}_i \neq 0$ in contrast to the case in which $\mathbf{A} = \mathbf{I}$. Unbiased estimators of σ_e^2 and σ_s^2 can be obtained by computing Method 1 type quadratics, that is

$$\mathbf{y}'\mathbf{y} - \sum_i y_{i.}^2/n_i$$

and

$$\sum_i y_{i.}^2/n_i - \text{C.F.}$$

However, the expectations must take into account the fact that $Var(\mathbf{s}) \neq \mathbf{I}\sigma_s^2$, but rather $\mathbf{A}\sigma_s^2$. In a non-inbred population

$$E(\mathbf{y}'\mathbf{y}) = n_.(\sigma_s^2 + \sigma_e^2).$$

For an inbred population the expectation is

$$\sum_i n_i a_{ii} \sigma_a^2 + n_. \sigma_e^2,$$

where a_{ii} is the ith diagonal element of $\mathbf{A}$. The coefficients of σ_e^2 in $\sum y_{i.}^2/n_i$ and $y_{..}^2/n_.$ are the same as in an unrelated sample of sires. The coefficients of σ_s^2 require the diagonals of Var(rhs). For our example, these coefficients are

$$\begin{pmatrix} 5 & 2 & 1 & 3 & 8 \\ 5 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix} \mathbf{A} \text{ (left matrix)}'$$

$$= \begin{pmatrix} 140.5 & 35. & 12. & 4.25 & 17.25 & 72. \\ & 25. & 0 & 2.5 & 7.5 & 0 \\ & & 4. & 0 & 0 & 8. \\ & & & 1. & .75 & 0 \\ & & & & 9. & 0 \\ & & & & & 64. \end{pmatrix}. \quad (18)$$

Then the coefficient of σ_a^2 in $\sum_i y_{i.}^2/n_i$ is $tr(dg(0,\ 5^{-1},\ 2^{-1},\ 1,\ 3^{-1},\ 8^{-1}))$ times the matrix in (18) = 19. The coefficient of σ_a^2 in $y_{..}^2/n_. = 140.5/19 = 7.395$.

If we wanted an approximate MIVQUE we could compute rather than

$$\sum_i \frac{y_{1.}^2}{n_i} - \frac{y_{..}^2}{n_.}$$

of Method 1, the quadratic,

$$\hat{\mathbf{u}}'\mathbf{A}^{-1}\hat{\mathbf{u}} = .3602.$$

The expectation of this is

$$tr(\mathbf{A}^{-1}\ Var(\hat{\mathbf{s}})).$$
$$Var(\hat{\mathbf{s}}) = \mathbf{C}_s\ Var(\text{rhs})\ \mathbf{C}_s'.$$

$\mathbf{C}_s$ is the last 5 rows of the inverse of the mixed model coefficient matrix.

$$Var(\text{rhs}) = \text{Matrix (18) } \sigma_s^2 + (\text{OLS coefficient matrix}) \ \sigma_e^2.$$

Then

$$Var(\mathbf{s}) = \begin{pmatrix} .0788 & -.0527 & .0443 & .0526 & -.0836 \\ & .0425 & -.0303 & -.0420 & .0561 \\ & & .0285 & .0283 & -.0487 \\ & & & .0544 & -.0671 \\ & & & & .1014 \end{pmatrix} \sigma_s^2 + \begin{pmatrix} .01284 & -.00774 & .00603 & .00535 & -.01006 \\ & .00982 & -.00516 & -.00677 & .00599 \\ & & .00731 & .00159 & -.00675 \\ & & & .01133 & -.00883 \\ & & & & .01462 \end{pmatrix} \sigma_e^2.$$

$\hat{\mathbf{s}}'\mathbf{A}^{-1}\hat{\mathbf{s}} = .36018$, with expectation $.05568\,\sigma_e^2 + .22977\,\sigma_s^2$. $\hat{\sigma}_e^2$ for approximate MIVQUE can be computed from

$$\mathbf{y}'\mathbf{y} - \sum_i y_{i.}^2/n_{i.}.$$

Chapter 17
The Two Way Classification

C. R. Henderson

1984 - Guelph

This chapter is concerned with a linear model in which

$$y_{ijk} = \mu + a_i + b_j + \gamma_{ij} + e_{ijk}. \tag{1}$$

For this to be a model we need to specify whether **a** is fixed or random, **b** is fixed or random, and accordingly whether $\boldsymbol{\gamma}$ is fixed or random. In the case of random subvectors we need to specify the variance-covariance matrix, and that is determined in part by whether the vector sampled is finite or infinite.

1 The Two Way Fixed Model

We shall be concerned first with a model in which **a** and **b** are both fixed, and as a consequence so is $\boldsymbol{\gamma}$. For convenience let

$$\mu_{ij} = \mu + a_i + b_j + \gamma_{ij}. \tag{2}$$

Then it is easy to prove that the only estimable linear functions are linear functions of μ_{ij} that are associated with filled subclasses ($n_{ij} > 0$). Further notations and definitions are:

$$\text{Row mean} = \bar{\mu}_{i.}. \tag{3}$$

Its estimate is sometimes called a least squares mean, but I agree with Searle *et al*. (1980) that this is not a desirable name.

$$\begin{aligned}
\text{Column mean} &= \bar{\mu}_{.j}. & (4)\\
\text{Row effect} &= \bar{\mu}_{i.} - \bar{\mu}_{..}. & (5)\\
\text{Column effect} &= \bar{\mu}_{.j} - \bar{\mu}_{..}. & (6)\\
\text{General mean} &= \bar{\mu}_{..}. & (7)\\
\text{Interaction effect} &= \mu_{ij} - \bar{\mu}_{i.}. - \bar{\mu}_{.j} + \bar{\mu}_{..}. & (8)
\end{aligned}$$

From the fact that only μ_{ij} for filled subclasses are estimable, missing subclasses result in the parameters of (17.3) ... (17.8) being non-estimable.

$\bar{\mu}_{i'.}$ is not estimable if any $n_{i'j} = 0$.
$\bar{\mu}_{.j'}$ is not estimable if any $n_{ij'} = 0$.
$\bar{\mu}_{..}$ is not estimable if one or more $n_{ij} = 0$.

All row effects, columns effects, and interaction effects are non-estimable if one or more $n_{ij} = 0$. Due to these non-estimability considerations, mimicking of either the balanced or the filled subclass estimation and tests of hypotheses wanted by many experimenters present obvious difficulties. We shall present biased methods that are frequently used and a newer method with smaller mean squared error of estimation given certain assumptions.

2 BLUE For The Filled Subclass Case

Assuming that $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$, it is easy to prove that $\hat{\mu}_{ij} = \bar{y}_{ij.}$. Then it follows that BLUE of the i^{th} row mean in the filled subclass case is

$$\frac{1}{c} \sum_{j=1}^{c} \bar{y}_{ij.}. \tag{9}$$

BLUE of j^{th} column mean is

$$\frac{1}{r} \sum_{i=1}^{r} \bar{y}_{ij.}. \tag{10}$$

$$\begin{aligned} r &= \text{number of rows, and} \\ c &= \text{number of columns.} \end{aligned}$$

BLUE of i^{th} row effect is

$$\frac{1}{c} \sum_j \bar{y}_{ij.} - \frac{1}{rc} \sum_i \sum_j \bar{y}_{ij.}. \tag{11}$$

Thus BLUE of any of (17.3), ... , (17.8) is that same function of $\hat{\mu}_{ij}$, where $\hat{\mu}_{ij} = \bar{y}_{ij.}$.

The variances of any of these functions are simple to compute. Any of them can be expressed as $\sum_i \sum_j k_{ij} \; \mu_{ij}$ with BLUE =

$$\sum_i \sum_j k_{ij}\bar{y}_{ij.}. \tag{12}$$

The variance of this is

$$\sigma_e^2 \sum_i \sum_j k_{ij}^2/n_{ij}. \tag{13}$$

The covariance between BLUE's of linear functions,

$$\sum_i \sum_j k_{ij}\bar{y}_{ij.} \text{ and } \sum_i \sum_j t_{ij}\bar{y}_{ij.'}$$

is

$$\sigma_e^2 \sum_i \sum_j k_{ij}t_{ij}/n_{ij}. \tag{14}$$

The numbers required for tests of hypotheses are (17.13) and (17.14) and the associated BLUE's. Consider a standard ANOVA, that is, mean squares for rows, columns, $R \times C$. The $R \times C$ sum of squares with $(r-1)(c-1)$ d.f. can be computed by

$$\sum_i \sum_j \frac{y_{ij.}^2}{n_{ij}} - \text{ Reduction under model with no interaction.} \tag{15}$$

The last term of (17.15) can be obtained by a solution to

$$\begin{pmatrix} \mathbf{D}_i & \mathbf{N}_{ij} \\ \mathbf{N}_{ij} & \mathbf{D}_j \end{pmatrix} \begin{pmatrix} \mathbf{a}^o \\ \mathbf{b}^o \end{pmatrix} = \begin{pmatrix} \mathbf{y}_i \\ \mathbf{y}_j \end{pmatrix}. \tag{16}$$

$$\begin{aligned} \mathbf{D}_i &= \text{diag } (n_{1.},\ n_{2.}, \ldots, n_{r.}). \\ \mathbf{D}_j &= \text{diag } (n_{.1},\ n_{.2}, \ldots, n_{.c}). \\ \mathbf{N}_{ij} &= \text{matrix of all } n_{ij}. \\ \mathbf{y}_i' &= (y_{1..}, \ldots, y_{r..}). \\ \mathbf{y}_j' &= (y_{.1.}, \ldots, y_{.c.}). \end{aligned}$$

Then the reduction is

$$(\mathbf{a}^o)'\ \mathbf{y}_i + (\mathbf{b}^o)'\mathbf{y}_j. \tag{17}$$

Sums of squares for rows and columns can be computed conveniently by the method of weighted squares of means, due to Yates (1934). For rows compute

$$\alpha_i = \frac{1}{c} \sum_j \bar{y}_{ij.}\ (i = 1, \ldots, r),\ \text{and} \tag{18}$$

$$k_i^{-1} = \frac{1}{c^2} \sum_j \frac{1}{n_{ij}}.$$

Then the row S.S. with $r - 1$ d.f. is

$$\sum_i k_i\ \alpha_i^2\ -\ (\textstyle\sum_i k_i \alpha_i)^2 / \sum_i k_i. \tag{19}$$

The column S.S. with $c - 1$ d.f. is computed in a similar manner. The "error" mean square for tests of these mean squares is

$$(\mathbf{y}'\mathbf{y} - \textstyle\sum_i \sum_j y_{ij.}^2 / n_{ij}) / (n_{..} - rc). \tag{20}$$

An obvious limitation of the weighted squares of means for testing rows is that the test refers to equal weighting of subclasses across columns. This may not be what is desired by the experimenter.

An illustration of a filled subclass 2 way fixed model is a breed by treatment design with the following n_{ij} and $y_{ij.}$.

	Treatments					
	n_{ij}			$y_{ij.}$		
Breeds	1	2	3	1	2	3
1	5	2	1	68	29	19
2	4	2	2	55	30	36
3	5	1	4	61	13	61
4	4	5	4	47	65	75

$$\sum_i \sum_j y_{ij.}^2/n_{ij} = 8207.5.$$

Let us test the hypothesis that interaction is negligible. The reduction under a model with no interaction can be obtained from a solution to equation (17.21).

$$\begin{pmatrix} 8 & 0 & 0 & 0 & 5 & 2 \\ & 8 & 0 & 0 & 4 & 2 \\ & & 10 & 0 & 5 & 1 \\ & & & 13 & 4 & 5 \\ & & & & 18 & 0 \\ & & & & & 10 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} 116 \\ 121 \\ 135 \\ 187 \\ 231 \\ 137 \end{pmatrix}. \qquad (21)$$

The solution is (18.5742, 18.5893, 16.3495, 17.4624, -4.8792, -4.0988)$'$. The reduction is 8187.933. Then $R \times C$ S.S. = 8207.5 - 8187.923 = 19.567. S. S. for rows can be formulated as a test of the hypothesis

$$\mathbf{K}'\boldsymbol{\mu} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \mu_{11} \\ \vdots \\ \mu_{43} \end{pmatrix} = \mathbf{0}.$$

The $\hat{\mu}_{ij}$ are (13.6, 14.5, 19.0, 13.75, 15.0, 18.0, 12.2, 13.0, 15.25, 11.75, 13.0, 18.75).

$$\begin{aligned} \mathbf{K}'\hat{\boldsymbol{\mu}} &= (3.6 \ 3.25 \ -3.05)'. \\ Var(\mathbf{K}'\hat{\boldsymbol{\mu}}) &= \mathbf{K}' \ [\text{diag } (5 \ 2 \ \ldots 4)]^{-1} \ \mathbf{K}\sigma_e^2 \\ &= \begin{pmatrix} 2.4 & .7 & .7 \\ & 1.95 & .7 \\ & & 2.15 \end{pmatrix} \sigma_e^2. \end{aligned}$$

$$\sigma_e^{-2} \ (\mathbf{K}'\hat{\boldsymbol{\mu}})' \ [Var(\mathbf{K}'\hat{\boldsymbol{\mu}})]^{-1} \ \mathbf{K}'\hat{\boldsymbol{\mu}} = 20.54 \ = \text{SS for rows.}$$

SS for cols. is a test of

$$\mathbf{K}'\boldsymbol{\mu} = \begin{pmatrix} 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\mu} = \mathbf{0}.$$

$$\mathbf{K}'\hat{\boldsymbol{\mu}} = \begin{pmatrix} -19.7 \\ -15.5 \end{pmatrix}.$$

$$Var(\mathbf{K}'\hat{\boldsymbol{\mu}}) = \begin{pmatrix} 2.9 & 2.0 \\ & 4.2 \end{pmatrix} \sigma_e^2.$$

$$\sigma_e^{-2} \ (\mathbf{K}'\hat{\boldsymbol{\mu}})' \ [Var(\mathbf{K}'\hat{\boldsymbol{\mu}})]^{-1} \ \mathbf{K}'\hat{\boldsymbol{\mu}} = 135.12 = \text{SS for Cols.}$$

Next we illustrate weighted squares of means to obtain these same results. Sums of squares for rows uses the values below

	α_i	k_i
1	15.7	5.29412
2	15.5833	7.2
3	13.4833	6.20690
4	14.5	12.85714

$$\begin{aligned}\textstyle\sum_i k_i\ \alpha_i^2 &= 6885.014.\\ (\textstyle\sum_i k_i\ \alpha_i)^2/\textstyle\sum_i k_i &= 6864.478.\\ \text{Diff.} &= 20.54 \text{ as before.}\end{aligned}$$

Sums of squares for columns uses the values below

	b_j	k_j
1	12.825	17.7778
2	13.875	7.2727
3	17.75	8.

$$\begin{aligned}\textstyle\sum k_j\ b_j^2 &= 6844.712.\\ (\textstyle\sum k_j)^2/\textstyle\sum k_j &= 6709.590.\\ \text{Diff.} &= 135.12 \text{ as before.}\end{aligned}$$

Another interesting method for obtaining estimates and tests involves setting up least squares equations using Lagrange multipliers to impose the following restrictions

$$\begin{aligned}\textstyle\sum_i \gamma_{ij} &= 0 \text{ for } i = 1, \ldots, r.\\ \textstyle\sum_i \gamma_{ij} &= 0 \text{ for } j = 1, \ldots, c.\\ \mu^o &= 0\end{aligned}$$

A solution is

$$\begin{aligned}\mathbf{b}' &= (0,\ -14,\ -266,\ -144)/120.\\ \mathbf{t}' &= (1645,\ 1771,\ 2236)/120.\\ \boldsymbol{\gamma}' &= (-13,\ -31,\ 44,\ 19,\ 43,\ -62,\ 85,\ 55,\ -140,\ -91,\ -67,\ 158)/120.\end{aligned}$$

Using these values, $\hat{\mu}_{ij}$ are the $\bar{y}_{ij.}$, and the reduction in SS is

$$\textstyle\sum_i \sum_j y_{ij.}^2/n_{ij.} = 8207.5.$$

Next the SS for rows is this reduction minus the reduction when $\mathbf{b}^o$ is dropped from the equations restricted as before. A solution in that case is

$$
\begin{aligned}
\mathbf{t}' &= (12.8133,\ 14.1223,\ 17.3099). \\
\boldsymbol{\gamma}' &= (.4509,\ -.4619,\ .0110,\ .4358,\ -.1241,\ -.3117, \\
&\quad -.0897,\ 1.4953,\ -1.4055,\ -.7970, .9093,\ 1.7063),
\end{aligned}
$$

and the reduction is 8186.960. The row sums of squares is

$$8207.5 - 8186.960 = 20.54 \text{ as before.}$$

Now drop $\mathbf{t}^o$ from the equations. A solution is

$$
\begin{aligned}
\hat{\mathbf{b}}' &= (13.9002,\ 15.0562,\ 13.5475,\ 14.4887). \\
\hat{\boldsymbol{\gamma}}' &= (.9648,\ .9390,\ -1.9039,\ .2751,\ .2830,\ -.5581, \\
&\quad -.0825, .1309,\ -.0485,\ -1.1574,\ -1.3530,\ 2.5104),
\end{aligned}
$$

and the reduction is 8072.377, giving the column sums of squares as

$$8207.5 - 8072.377 = 135.12 \text{ as before.}$$

An interesting way to obtain estimates under the sum to 0 restrictions in $\boldsymbol{\gamma}$ is to solve

$$\bar{\mathbf{X}}_0'\ \bar{\mathbf{X}}_0 \begin{pmatrix} \mathbf{b}^o \\ \mathbf{t}^o \end{pmatrix} = \bar{\mathbf{X}}_0 \bar{\mathbf{y}},$$

where $\bar{\mathbf{X}}_0$ is the submatrix of $\bar{\mathbf{X}}$ referring to $\mathbf{b}$, $\mathbf{t}$ only, and $\bar{\mathbf{y}}$ is a vector of subclass means. These equations are

$$
\begin{pmatrix}
3 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & 3 & 0 & 0 & 1 & 1 & 1 \\
 & & 3 & 0 & 1 & 1 & 1 \\
 & & & 3 & 1 & 1 & 1 \\
 & & & & 4 & 0 & 0 \\
 & & & & & 4 & 0 \\
 & & & & & & 4
\end{pmatrix}
\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ t_1 \\ t_2 \\ t_3 \end{pmatrix}
=
\begin{pmatrix} 47.1 \\ 46.75 \\ 40.45 \\ 43.5 \\ 51.3 \\ 55.5 \\ 71.0 \end{pmatrix}. \tag{22}
$$

A solution is

$$
\begin{aligned}
\mathbf{b}' &= (0,\ -14,\ -266,\ -144)/120, \\
\mathbf{t}' &= (1645,\ 1771,\ 2236)/120.
\end{aligned}
$$

This is the same as in the restricted least squares solution. Then

$$\hat{\gamma}_{ij} = \bar{y}_{ij.} - b_i^o - t_j^o,$$

which gives the same result as before. More will be said about these alternative methods in the missing subclass case.

3 The Fixed, Missing Subclass Case

When one or more subclasses is missing, the estimates and tests described in Section 2 cannot be effected. What should be done in this case? There appears to be no agreement among statisticians. It is of course true that any linear functions of μ_{ij} in which $n_{ij} > 0$ can be estimated by BLUE and can be tested, but these may not be of any particular interest to the researcher. One method sometimes used, and this is the basis of a SAS Type 4 analysis, is to select a subset of subclasses, all filled, and then to do a weighted squares of means analysis on this subset. For example, suppose that in a 3×4 design, subclass (1,2) is missing. Then one could discard all data from the second column, leaving a 3×3 design with filled subclasses. This would mean that rows are compared by averaging over columns 1,3,4 and only columns 1,3,4 are compared, these averaged over the 3 rows. One could also discard the first row leaving a 2×4 design. The columns are compared by averaging over only rows 2 and 3, and only rows 2 and 3 are compared, averaging over all 4 columns. Consequently this method is not unique because usually more than one filled subset can be chosen. Further, most experimenters are not happy with the notion of discarding data that may have been costly to obtain.

Another possibility is to estimate μ_{ij} for missing subclasses by some biased procedure. For example, one can estimate μ_{ij} such that $E(\hat{\mu}_{ij}) = \mu + a_i + b_j +$ some function of the γ_{ij} associated with filled subclasses. One way of doing this is to set up least squares equations with the following restrictions.

$$\begin{aligned} \textstyle\sum_j \gamma_{ij} &= 0 \text{ for } i = 1, \ldots, r. \\ \textstyle\sum_i \gamma_{ij} &= 0 \text{ for } j = 1, \ldots, c. \\ \gamma_{ij} &= 0 \text{ if } n_{ij} = 0. \end{aligned}$$

This is the method used in Harvey's computer package. When equations with these restrictions are solved,

$$\hat{\mu}_{ij} = \mu + a_i^o + b_j^o + \gamma_{ij}^o = \bar{y}_{ij.},$$

when $n_{ij} > 0$ and thus is unbiased. A biased estimator for a missing subclass is $\mu^o + a_i^o + b_j^o$, and this has expectation $\mu + a_i + b_j + \sum_i \sum_j k_{ij}\gamma_{ij}$, where summation in the last term is over filled subclass and $\sum_i \sum_j k_{ij} = 1$. Harvey's package does not compute this but does produce "least squares means" for main effects and some of these are biased.

Thus $\hat{\mu}_{ij}$ is BLUE for filled subclasses and is biased for empty subclasses. In the class of estimators of μ_{ij} with expectation $\mu + a_i + b_j +$ some linear function of μ_{ij} associated with filled subclasses, this method minimizes the contribution of quadratics in $\boldsymbol{\gamma}$ to mean squared error when the squares and products of the elements of $\boldsymbol{\gamma}$ are in accord with no particular pattern of values. This minimization might appear to be a desirable property, but unfortunately the method does not control contributions of σ_e^2 to MSE. If one wishes to minimize the contribution of σ_e^2, but not to control on quadratics in $\boldsymbol{\gamma}$, while still having $E(\hat{\mu}_{ij})$ contain $\mu + a_i + b_j$, the way to accomplish this is to solve least squares equations with $\boldsymbol{\gamma}$ dropped. Then the biased estimators in this case for filled as well as empty subclasses, are

$$\hat{\mu}_{ij} = \mu^o + a_i^o + b_j^o. \tag{23}$$

A third possibility is to assume some prior values of σ_e^2 and squares and products of γ_{ij} and compute as in Section 9.1. Then all $\hat{\mu}_{ij}$ are biased by γ_{ij} but have in their expectations $\mu + a_i + b_j$. Finally one could relax the requirement of $\mu + a_i + b_j$ in the expectation of $\hat{\mu}_{ij}$. In that case one would assume average values of squares and products of the a_i and b_j as well as for the γ_j and use the method described in Section 9.1.

Of these biased methods, I would usually prefer the one in which priors on the $\boldsymbol{\gamma}$, but not on **a** and **b** are used. In most fixed, 2 way models the number of levels of **a** and **b** are too small to obtain a good estimate of the pseudo-variances of **a** and **b**.

We illustrate these methods with a 4×3 design with 2 missing subclasses as follows.

	n_{ij}				$y_{ij.}$			
	1	2	3	4	1	2	3	4
1	5	2	3	2	30	11	13	7
2	4	2	0	5	21	6	–	9
3	3	0	1	4	12	–	3	15

4 A Method Based On Assumption $\gamma_{ij} = 0$ If $n_{ij} = 0$

First we illustrate estimation under sum to 0 model for $\boldsymbol{\gamma}$ and in addition the assumption that $\gamma_{23} = \gamma_{32} = 0$. The simplest procedure for this set of restrictions is to solve for $\mathbf{a}^o$, $\mathbf{b}^o$ in equations (17.24).

$$\begin{pmatrix} 4 & 0 & 0 & 1 & 1 & 1 & 1 \\ & 3 & 0 & 1 & 1 & 0 & 1 \\ & & 3 & 1 & 0 & 1 & 1 \\ & & & 3 & 0 & 0 & 0 \\ & & & & 2 & 0 & 0 \\ & & & & & 2 & 0 \\ & & & & & & 3 \end{pmatrix} \begin{pmatrix} \mathbf{a}^o \\ \mathbf{b}^o \end{pmatrix} = \begin{pmatrix} 19.333 \\ 10.05 \\ 10.75 \\ 15.25 \\ 8.5 \\ 7.333 \\ 9.05 \end{pmatrix}. \tag{24}$$

The first right hand side is $\frac{30}{5} + \frac{11}{2} + \frac{13}{3} + \frac{7}{2} = 19.333$, etc. for others. A solution is (3.964, 2.286, 2.800, 2.067, 1.125, .285, 0). The estimates of μ_{ij} are $\bar{y}_{ij.}$ for filled subclasses and 2.286 + .285 for $\hat{\mu}_{23}$ and 2.800 + 1.125 for $\hat{\mu}_{32}$. If $\hat{\gamma}_{ij}$ are wanted they are

$$\hat{\gamma}_{11} = \bar{y}_{11} - 3.964 - 1.125$$

etc, for filled subclasses, and 0 for $\hat{\gamma}_{23}$ and $\hat{\gamma}_{32}$.

The same results can be obtained, but with much heavier computing by solving least squares equations with restrictions on γ that are $\sum_j \gamma_{ij} = 0$ for all i, $\sum_i \gamma_{ij} = 0$ for all j, and $\gamma_{ij} = 0$ for subclasses with $n_{ij} = 0$. From these equations one can obtain sums of squares that mimic weighted squares of means. A solution to the restricted equations is

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{a}^o &= (3.964,\ 2.286,\ 2.800,\ 2.067)'. \\
\mathbf{b}^o &= (1.125,\ .285,\ 0)'. \\
\boldsymbol{\gamma}^o &= (-.031,\ .411,\ .084,\ -.464,\ .897,\ -.411, \\
& \quad\ 0,\ -.486,\ -.866,\ \ 0,\ -.084,\ .950)'.
\end{aligned}$$

Note that the solution to $\boldsymbol{\gamma}$ conforms to the restrictions imposed. Also note that this solution is the same as the one previously obtained. Further, $\hat{\mu}_{ij} = \mu^o + a_i^o + b_j^o + \gamma_{ij}^o = \bar{y}_{ij.}$ for filled subclasses.

A test of hypothesis that the main effects are equal, that is $\bar{\mu}_{i.} = \bar{\mu}_{i'}$ for all pairs of i, i$'$, can be effected by taking a new solution to the restricted equations with $\mathbf{a}^o$ dropped. Then the SS for rows is

$$(\boldsymbol{\beta}^o)' \text{ RHS} - (\boldsymbol{\beta}_*^o)' \text{ RHS}_*, \tag{25}$$

where $\boldsymbol{\beta}^o$ is a solution to the full set of equations, and this reduction is simply $\sum_i \sum_j y_{ij}^2/n_{ij.}$, $\boldsymbol{\beta}_*^o$ is a solution with $\mathbf{a}$ deleted from the set of equations, and RHS_* is the right hand side. This tests a nontestable hypothesis inasmuch as the main effects are not estimable when subclasses are missing. The test is valid only if γ_{ij} are truly 0 for all missing subclasses, and this is not a testable assumption, Henderson and McAllister (1978). If one is to use a test based on non-estimable functions, as is done in this case, there should be some attempt to evaluate the numerator with respect to quadratics in fixed effects other than those being tested and use this in the denominator. That is, a minimum requirement could seem to be a test of this sort.

$$\begin{aligned}
E(\text{numerator}) &= Q_t(\mathbf{a}) + Q \text{ (fixed effects causing bias in the estimator)} \\
& \quad + \text{ linear functions of random variables.}
\end{aligned}$$

Then the denominator should have the same expectation except that $Q_t(\mathbf{a})$, the quadratic in fixed effects being tested, would not be present. In our example the reduction under the full model with restrictions on $\boldsymbol{\gamma}$ is 579.03, and this is the same as the uncorrected subclass sum of squares. A solution with $\boldsymbol{\gamma}$ restricted as before and with $\mathbf{a}$ dropped is

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{b}^o &= (5.123,\ 4.250,\ 4.059,\ 2.790)', \\
\boldsymbol{\gamma}^o &= (.420,\ .129,\ -.119,\ -.430,\ .678,\ -.129, \\
& \quad\ 0,\ -.549,\ -1.098,\ \ 0,\ .119,\ .979)'.
\end{aligned}$$

This gives a reduction of 566.32. Then the sum of squares with 2 df for the numerator is 579.03-566.32, but $\hat{\sigma}_e^2$ is not an appropriate denominator MS, when $\hat{\sigma}_e^2$ is the within subclass mean square, unless γ_{23} and γ_{32} are truly equal to zero, and we cannot test this assumption.

Similarly a solution when $\mathbf{b}$ is dropped is

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{a}^o &= (5.089, 3.297, 3.741)', \\
\boldsymbol{\gamma}^o &= (.098,\ .254,\ -.355,\ .003,\ 1.035,\ -.254, \\
&\quad 0,\ -.781,\ -1.133,\ \ 0,\ .355,\ .778)'.
\end{aligned}$$

The reduction is 554.81. Then if γ_{23} and $\gamma_{32} = 0$, the numerator sum of squares with 3 df is 579.03-554.81. The sum of squares for interaction with (3-1)(4-1)-2 = 4 df. is 579.03 - reduction with $\boldsymbol{\gamma}$ and the Lagrange multiplier deleted. This latter reduction is 567.81 coming from a solution

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{a}^o &= (3.930,\ 2.296,\ 2.915)',\ \text{and} \\
\boldsymbol{\beta}^o &= (2.118, 1.137, .323, 0)'.
\end{aligned}$$

5 Biased Estimation By Ignoring γ

Another biased estimation method sometimes suggested is to ignore $\boldsymbol{\gamma}$. That is, least squares equations with only μ^o, $\mathbf{a}^o$, $\mathbf{b}^o$ are solved. This is sometimes called the method of fitting constants, Yates (1934). This method has quite different properties than the method of Section 17.4. Both obtain estimators of μ_{ij} with expectations $\mu + a_i + b_j$ + linear functions of γ_{ij}. The method of section 17.4 minimizes the contribution of quadratics in $\boldsymbol{\gamma}$ to MSE, but does a poor job of controlling on the contribution of σ_e^2. In contrast, the method of fitting constants minimizes the contribution of σ_e^2 but does not control quadratics in $\boldsymbol{\gamma}$. The method of the next section is a compromise between these two extremes.

A solution for our example for the method of this section is

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{a}^o &= (3.930,\ 2.296,\ 2.915)', \\
\mathbf{b}^o &= (2.118,\ 1.137,\ .323,\ \ 0)'.
\end{aligned}$$

Then if we wish $\hat{\mu}_{ij}$ these are $\mu^o + a_i^o + b_j^o$.

A test of row effects often suggested is to compute the reduction in SS under the model with $\boldsymbol{\gamma}$ dropped minus the reduction when $\mathbf{a}$ and $\boldsymbol{\gamma}$ are dropped, the latter being simply $\sum_j y_{.j.}^2 / n_{.j.}$. Then this is tested against some denominator. If $\hat{\sigma}_e^2$ is used, the denominator is too small unless $\boldsymbol{\gamma}$ is $\mathbf{0}$. If $R \times C$ for MS is used, the denominator is probably too large. Further, the numerator is not a test of rows averaged in some logical way across columns, but rather each row is averaged differently depending upon the pattern of subclass numbers. That is, $\mathbf{K}'\boldsymbol{\beta}$ is dependent upon the incidence matrix, an obviously undesirable property.

6 Priors On Squares And Products Of γ

The methods of the two preceding sections control in the one case on $\boldsymbol{\gamma}$ and the other on σ_e^2 as contributors to MSE. The method of this section is an attempt to control on both. The logic of the method depends upon the assumption that there is no pattern of values of $\boldsymbol{\gamma}$, such, for example as linear by columns or linear by rows. Then consider the matrix of squares and products of elements of γ_{ij} for all possible permutations of rows and columns. The average values are found to be

$$\begin{aligned} \gamma_{ij}^2 &= \alpha. \\ \gamma_{ij}\gamma_{ij'} &= -\alpha/(c-1). \\ \gamma_{ij}\gamma_{i'j} &= -\alpha/(r-1). \\ \gamma_{ij}\gamma_{i'j'} &= \alpha/(r-1)(c-1). \end{aligned} \tag{26}$$

Note that if we substitute σ_γ^2 for α, this is the same matrix as that for $Var(\boldsymbol{\gamma})$ in the finite random rows and finite random columns model. Then if we have estimates of σ_e^2 and α or an estimate of the relative magnitudes of these parameters, we can proceed to estimate with $\mathbf{a}$ and $\mathbf{b}$ regarded as fixed and $\boldsymbol{\gamma}$ regarded as a pseudo random variable.

We illustrate with our same numerical example. Assume that $\sigma_e^2 = 20$ and $\alpha = 6$. Write the least squares equations that include γ_{23} and γ_{32}, the missing subclasses. Premultiply the last 12 equations by

$$\begin{pmatrix} 6 & -2 & -2 & -2 & -3 & 1 & 1 & 1 & -3 & 1 & 1 & 1 \\ & 6 & -2 & -2 & 1 & -3 & 1 & 1 & 1 & -3 & 1 & 1 \\ & & 6 & -2 & 1 & 1 & -3 & 1 & 1 & 1 & -3 & 1 \\ & & & 6 & 1 & 1 & 1 & -3 & 1 & 1 & 1 & -3 \\ & & & & 6 & -2 & -2 & -2 & -3 & 1 & 1 & 1 \\ & & & & & 6 & -2 & -2 & 1 & -3 & 1 & 1 \\ & & & & & & 6 & -2 & 1 & 1 & -3 & 1 \\ & & & & & & & 6 & 1 & 1 & 1 & -3 \\ & & & & & & & & 6 & -2 & -2 & -2 \\ & & & & & & & & & 6 & -2 & -2 \\ & & & & & & & & & & 6 & -2 \\ & & & & & & & & & & & 6 \end{pmatrix}. \tag{27}$$

Then add 1 to each of the last 12 diagonals. The resulting coefficient matrix is (17.28) ... (17.31). The right hand side vector is (3.05, 1.8, 1.5, 3.15, .85, .8, 2.6, .4, 1.8, -4.8, .95, 1.15, -2.25, .15, -3.55, -1.55, .45, 4.65)′ $\boldsymbol{\beta}' = (a_1 a_2 a_3 b_1 b_2 b_3 \boldsymbol{\gamma}')$. Thus μ and b_4 are deleted, which is equivalent to obtaining a solution with $\mu^o = 0$, $b_4^o = 0$.

Upper left 9 × 9

$$\begin{pmatrix} .6 & 0 & 0 & .25 & .1 & .15 & .25 & .1 & .15 \\ 0 & .55 & 0 & .2 & .1 & 0 & 0 & 0 & 0 \\ 0 & 0 & .4 & .15 & 0 & .05 & 0 & 0 & 0 \\ .25 & .2 & .15 & .6 & 0 & 0 & .25 & 0 & 0 \\ .1 & .1 & 0 & 0 & .2 & 0 & 0 & .1 & 0 \\ .15 & 0 & .05 & 0 & 0 & .2 & 0 & 0 & .15 \\ .8 & -.25 & -.2 & .45 & -.1 & -.25 & 2.5 & -.2 & -.3 \\ -.4 & .15 & .4 & -.15 & .3 & -.25 & -.5 & 1.6 & -.3 \\ 0 & .55 & .2 & -.15 & -.1 & .75 & -.5 & -.2 & 1.9 \end{pmatrix} \tag{28}$$

Upper right 9 × 9

$$\begin{pmatrix} .1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .2 & .1 & 0 & .25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .15 & 0 & .05 & .2 \\ 0 & .2 & 0 & 0 & 0 & .15 & 0 & 0 & 0 \\ 0 & 0 & .1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .05 & 0 \\ -.2 & -.6 & .1 & 0 & .25 & -.45 & 0 & .05 & .2 \\ -.2 & .2 & -.3 & 0 & .25 & .15 & 0 & .05 & .2 \\ -.2 & .2 & .1 & 0 & .25 & .15 & 0 & -.15 & .2 \end{pmatrix} \tag{29}$$

Lower left 9 × 9

$$\begin{pmatrix} -.4 & -.45 & -.4 & -.15 & -.1 & -.25 & -.5 & -.2 & -.3 \\ -.4 & .5 & -.2 & 0 & -.1 & .2 & -.75 & .1 & .15 \\ .2 & -.3 & .4 & 0 & .3 & .2 & .25 & -.3 & .15 \\ 0 & -1.1 & .2 & 0 & -.1 & -.6 & .25 & .1 & -.45 \\ .2 & .9 & -.4 & 0 & -.1 & .2 & .25 & .1 & .15 \\ -.4 & -.25 & .4 & -.45 & .2 & .05 & -.75 & .1 & .15 \\ .2 & .15 & -.8 & .15 & -.6 & .05 & .25 & -.3 & .15 \\ 0 & .55 & -.4 & .15 & .2 & -.15 & .25 & .1 & -.45 \\ .2 & -.45 & .8 & .15 & .2 & .05 & .25 & .1 & .15 \end{pmatrix} \tag{30}$$

Lower right 9×9

$$\begin{pmatrix}
1.6 & .2 & .1 & 0 & -.75 & .15 & 0 & .05 & -.6 \\
.1 & 2.2 & -.2 & 0 & -.5 & -.45 & 0 & .05 & .2 \\
.1 & -.4 & 1.6 & 0 & -.5 & .15 & 0 & .05 & .2 \\
.1 & -.4 & -.2 & 1.0 & -.5 & .15 & 0 & -.15 & .2 \\
-.3 & -.4 & -.2 & 0 & 2.5 & .15 & 0 & .05 & -.6 \\
.1 & -.6 & .1 & 0 & .25 & 1.9 & 0 & -.1 & -.4 \\
.1 & .2 & -.3 & 0 & .25 & -.3 & 1.0 & -.1 & -.4 \\
.1 & .2 & .1 & 0 & .25 & -.3 & 0 & 1.3 & -.4 \\
-.3 & .2 & .1 & 0 & -.75 & -.3 & 0 & -.1 & 2.2
\end{pmatrix} \tag{31}$$

The solution is

$$\begin{aligned}
\mathbf{a}^o &= (3.967,\ 2.312,\ 2.846)'. \\
\mathbf{b}^o &= (2.068,\ 1.111,\ .288,\ \ 0)'.
\end{aligned}$$

$\boldsymbol{\gamma}^o$ displayed as a table is

	1	2	3	4
1	-.026	.230	.050	-.255
2	.614	-.230	0	-.384
3	-.588	0	-.050	.638

Note that the γ_{ij}^o sum to 0 by rows and columns. Now the $\hat{\mu}_{ij} = a_i^o + b_j^o + \gamma_{ij}^o$. The same solution can be obtained more easily by treating $\boldsymbol{\gamma}$ as a random variable with $Var = 12\mathbf{I}$. The value 12 comes from $\frac{rc}{(r-1)(c-1)}\ 6 = \frac{(3)4}{(2)3}\ (6) = 12$. The resulting coefficient matrix (times 60) is in (17.32). The right hand side vector is (3.05, 1.8, 1.5, 3.15, .85, .8, 1.5, .55, .65, .35, 1.05, .3, 0, .45, .6, 0, .15, .75)$'$. μ and b_4 are dropped as before.

$$\begin{pmatrix}
36 & 0 & 0 & 15 & 6 & 9 & 15 & 6 & 9 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & 33 & 0 & 12 & 6 & 0 & 0 & 0 & 0 & 0 & 12 & 6 & 0 & 15 & 0 & 0 & 0 & 0 \\
 & & 24 & 9 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 3 & 12 \\
 & & & 36 & 0 & 0 & 15 & 0 & 0 & 0 & 12 & 0 & 0 & 0 & 9 & 0 & 0 & 0 \\
 & & & & 12 & 0 & 0 & 6 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 12 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0
\end{pmatrix} \tag{32}$$

$$\text{diag}\ (20,11,14,11,17,11,5,20,14,5,8,17)$$

The solution is the same as before. This is clearly an easier procedure than using the equations of (17.28). The inverse of the matrix of (17.28) post-multiplied by

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix},$$

where $\mathbf{P}$ = the matrix of (17.27), is not the same as the inverse of the matrix of (17.32) with diagonal $\mathbf{G}$, but if we pre-multiply each of them by $\mathbf{K}'$ and then post-multiply by $\mathbf{K}$, where $\mathbf{K}'$ is the representation of μ_{ij} in terms of $\mathbf{a}$, $\mathbf{b}$, $\boldsymbol{\gamma}$, we obtain the same matrix, which is the mean squared error for the $\hat{\mu}_{ij}$ under the priors used, $\sigma_e^2 = 20$ and $\alpha = 6$. Biased estimates of $\hat{\mu}_{ij}$ are in both methods

$$\begin{pmatrix} 6.009 & 5.308 & 4.305 & 3.712 \\ 4.994 & 3.192 & 2.600 & 1.928 \\ 4.327 & 3.957 & 3.084 & 3.484 \end{pmatrix}.$$

The estimated MSE matrix of this vector is
Upper left 8 × 8

$$\begin{pmatrix} 3.58 & .32 & .18 & .46 & .28 & -.32 & -.87 & -.09 \\ & 8.34 & .15 & .64 & -.43 & 1.66 & -2.29 & -.32 \\ & & 6.01 & .38 & .02 & -.15 & 4.16 & .04 \\ & & & 7.64 & -.29 & -.64 & -1.77 & .49 \\ & & & & 4.40 & .43 & 1.93 & .31 \\ & & & & & 8.34 & 2.29 & .32 \\ & & & & & & 33.72 & 1.54 \\ & & & & & & & 3.63 \end{pmatrix}$$

Upper right 8 × 4

$$\begin{pmatrix} .33 & -.70 & -.55 & -.11 \\ .04 & 5.21 & -.45 & .08 \\ -.33 & -1.33 & 1.97 & -.24 \\ -.37 & -1.47 & -1.14 & .57 \\ .34 & -1.09 & -.06 & -.24 \\ -.04 & 4.79 & .45 & -.08 \\ -1.12 & -4.67 & 7.51 & -1.04 \\ -.25 & -1.04 & -.13 & .22 \end{pmatrix}$$

Lower right 4 × 4

$$\begin{pmatrix} 5.66 & 2.62 & 1.00 & .50 \\ & 33.60 & 4.00 & 2.03 \\ & & 14.08 & .73 \\ & & & 4.44 \end{pmatrix}$$

Suppose we wish an approximate test of the hypothesis that $\bar{\mu}_{i.}$ are equal. In this case we could write $\mathbf{K}'$ as

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{-1}{4} & \frac{-1}{4} & \frac{-1}{4} & \frac{-1}{4} \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{-1}{4} & \frac{-1}{4} & \frac{-1}{4} \end{pmatrix}.$$

Then compute $\mathbf{K}'\mathbf{C}\mathbf{K}$, where $\mathbf{C}$ is either the g-inverse of (17.28) post-multiplied by

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix},$$

or the g-inverse of the matrix using diagonal $\mathbf{G}$. This 2×2 matrix gives the MSE for $\sigma_e^2 = 20$, $\alpha = 6$. Finally premultiply the inverse of this matrix by $(\mathbf{K}'\boldsymbol{\beta}^o)'$ and post-multiply by $\mathbf{K}'\boldsymbol{\beta}^o$. This quantity is distributed approximately as χ_2^2 under the null hypothesis.

7 Priors On Squares And Products Of a, b, And γ

Another possibility for biased estimation is to require only that

$$E(\hat{\mu}_{ij}) = \mu + \text{linear functions of } \mathbf{a}, \mathbf{b},\ \boldsymbol{\gamma}.$$

We do this by assuming prior values of squares and products of $\mathbf{a}$ and of $\mathbf{b}$ as

$$\begin{pmatrix} 1 & & \frac{-1}{r-1} \\ & \ddots & \\ \frac{-1}{r-1} & & 1 \end{pmatrix} \sigma_a^2 \text{ and } \begin{pmatrix} 1 & & \frac{-1}{c-1} \\ & \ddots & \\ \frac{-1}{c-1} & & 1 \end{pmatrix} \sigma_b^2,$$

respectively, where σ_a^2 and σ_b^2 are pseudo-variances. The prior on $\boldsymbol{\gamma}$ is the same as in Section 17.6. Then we apply the method for singular G.

To illustrate in our example, let the priors be $\alpha_e^2 = 20$, $\alpha_a^2 = 4$, $\alpha_b^2 = 9$, $\alpha_\gamma^2 = 6$. Then we multiply all equations except the first pertaining to μ by

$$\begin{pmatrix} \mathbf{P}_a\sigma_a^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_b\sigma_b^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_\gamma\sigma_\gamma^2 \end{pmatrix},$$

and add 1's to all diagonals except the first. This yields the equations with coefficient matrix in (17.33) . . . (17.36) and right hand vector = (6.35, 5.60, -1.90, -3.70, 18.75, -8.85, -9.45, -.45, 2.60, .40, 1.80, -4.80, .95, 1.15, -2.25, .15, -3.55, -1.55, .45, 4.65)′.

Upper left 10 × 10

$$\begin{pmatrix}
1.55 & .6 & .55 & .4 & .6 & .2 & .2 & .55 & .25 & .1 \\
.5 & 3.4 & -1.1 & -.8 & .3 & .2 & .5 & -.5 & 1.0 & .4 \\
.2 & -1.2 & 3.2 & -.8 & 0 & .2 & -.4 & .4 & -.5 & -.2 \\
-.7 & -1.2 & -1.1 & 2.6 & -.3 & -.4 & -.1 & .1 & -.5 & -.2 \\
2.55 & 1.2 & .75 & .6 & 6.4 & -.6 & -.6 & -1.65 & 2.25 & -.3 \\
-2.25 & -.6 & -.45 & -1.2 & -1.8 & 2.8 & -.6 & -1.65 & -.75 & .9 \\
-2.25 & 0 & -1.65 & -.6 & -1.8 & -.6 & 2.8 & -1.65 & -.75 & -.3 \\
1.95 & -.6 & 1.35 & 1.2 & -1.8 & -.6 & -.6 & 5.95 & -.75 & -.3 \\
.35 & .8 & -.25 & -.2 & .45 & -.1 & -.25 & .25 & 2.5 & -.2 \\
.15 & -.4 & .15 & .4 & -.15 & .3 & -.25 & .25 & -.5 & 1.6
\end{pmatrix} \quad (33)$$

Upper right 10 × 10

$$\begin{pmatrix} .15 & .1 & .2 & .1 & 0 & .25 & .15 & 0 & .05 & .2 \\ .6 & .4 & -.4 & -.2 & 0 & -.5 & -.3 & 0 & -.1 & -.4 \\ -.3 & -.2 & .8 & .4 & 0 & 1.0 & -.3 & 0 & -.1 & -.4 \\ -.3 & -.2 & -.4 & -.2 & 0 & -.5 & .6 & 0 & .2 & .8 \\ -.45 & -.3 & 1.8 & -.3 & 0 & -.75 & 1.35 & 0 & -.15 & -.6 \\ -.45 & -.3 & -.6 & .9 & 0 & -.75 & -.45 & 0 & -.15 & -.6 \\ 1.35 & -.3 & -.6 & -.3 & 0 & -.75 & -.45 & 0 & .45 & -.6 \\ -.45 & .9 & -.6 & -.3 & 0 & 2.25 & -.45 & 0 & -.15 & 1.8 \\ -.3 & -.2 & -.6 & .1 & 0 & .25 & -.45 & 0 & .05 & .2 \\ -.3 & -.2 & .2 & -.3 & 0 & .25 & .15 & 0 & .05 & .2 \end{pmatrix} \tag{34}$$

Lower left 10 × 10

$$\begin{pmatrix} .75 & 0 & .55 & .2 & -.15 & -.1 & .75 & .25 & -.5 & -.2 \\ -1.25 & -.4 & -.45 & -.4 & -.15 & -.1 & -.25 & -.75 & -.5 & -.2 \\ -.1 & -.4 & .5 & -.2 & 0 & -.1 & .2 & -.2 & -.75 & .1 \\ .3 & .2 & -.3 & .4 & 0 & .3 & .2 & -.2 & .25 & -.3 \\ -.9 & 0 & -1.1 & .2 & 0 & -.1 & -.6 & -.2 & .25 & .1 \\ .7 & .2 & .9 & -.4 & 0 & -.1 & .2 & .6 & .25 & .1 \\ -.25 & -.4 & -.25 & .4 & -.45 & .2 & .05 & -.05 & -.75 & .1 \\ -.45 & .2 & .15 & -.8 & .15 & -.6 & .05 & -.05 & .25 & -.3 \\ .15 & 0 & .55 & -.4 & .15 & .2 & -.15 & -.05 & .25 & .1 \\ .55 & .2 & -.45 & .8 & .15 & .2 & .05 & .15 & .25 & .1 \end{pmatrix} \tag{35}$$

Lower right 10 × 10

$$\begin{pmatrix} 1.9 & -.2 & .2 & .1 & 0 & .25 & .15 & 0 & -.15 & .2 \\ -.3 & 1.6 & .2 & .1 & 0 & -.75 & .15 & 0 & .05 & -.6 \\ .15 & .1 & 2.2 & -.2 & 0 & -.5 & -.45 & 0 & .05 & .2 \\ .15 & .1 & -.4 & 1.6 & 0 & -.5 & .15 & 0 & .05 & .2 \\ -.45 & .1 & -.4 & -.2 & 1.0 & -.5 & .15 & 0 & -.15 & .2 \\ .15 & -.3 & -.4 & -.2 & 0 & 2.5 & .15 & 0 & .05 & -.6 \\ .15 & .1 & -.6 & .1 & 0 & .25 & 1.9 & 0 & -.1 & -.4 \\ .15 & .1 & .2 & -.3 & 0 & .25 & -.3 & 1.0 & -.1 & -.4 \\ -.45 & .1 & .2 & .1 & 0 & .25 & -.3 & 0 & 1.3 & -.4 \\ .15 & -.3 & .2 & .1 & 0 & -.75 & -.3 & 0 & -.1 & 2.2 \end{pmatrix} \tag{36}$$

The solution is

$$\begin{aligned} \hat{\mu} &= 4.014, \\ \hat{\mathbf{a}} &= (.650,\ -.467,\ -.183), \\ \hat{\mathbf{b}} &= (.972,\ .120,\ -.208,\ -.885), \\ \hat{\boldsymbol{\gamma}} &= \begin{pmatrix} .111 & .225 & -.170 & -.166 \\ .489 & -.276 & .303 & -.515 \\ -.599 & .051 & -.133 & .681 \end{pmatrix}. \end{aligned}$$

Note that $\sum_i \hat{a}_i = \sum_j \hat{b}_j = 0$ and the $\hat{\gamma}_{ij}$ sum to 0 by rows and columns. A solution can be obtained by pretending that $\mathbf{a}$, $\mathbf{b}$, $\boldsymbol{\gamma}$ are random variables with $Var(\mathbf{a}) = 3\mathbf{I}$, $Var(\mathbf{b}) = 8\mathbf{I}$, $Var(\boldsymbol{\gamma}) = 12\mathbf{I}$. The coefficient matrix of these is in (17.37) ... (17.39) and the right hand side is (6.35, 3.05, 1.8, 1.5, 3.15, .85, .8, 1.55, 1.5, .55, .65, .35, 1.05, .3, 0, .45, .6, 0, .15, .75)'. The solution is

$$\begin{aligned}
\hat{\mu} &= 4.014, \\
\hat{\mathbf{a}} &= (.325,\ -.233,\ -.092)', \\
\hat{\mathbf{b}} &= (.648,\ .080,\ -.139,\ -.590)', \\
\hat{\boldsymbol{\gamma}} &= \begin{pmatrix} .760 & .590 & .086 & -.136 \\ .580 & -.470 & 0 & -1.043 \\ -.367 & 0 & -.294 & .295 \end{pmatrix}.
\end{aligned}$$

This is a different solution from the one above, but the $\hat{\mu}_{ij}$ are identical for the two. These are as follows, in table form,

$$\begin{pmatrix} 5.747 & 5.009 & 4.286 & 3.613 \\ 5.009 & 3.391 & 3.642 & 2.148 \\ 4.204 & 4.003 & 3.490 & 3.627 \end{pmatrix}.$$

Note that $\sum_i \hat{a}_i = \sum_j \hat{b}_j = 0$, but the $\hat{\gamma}_{ij}$ do not sum to 0 by rows and columns.

$$\begin{aligned}
\sum_j \hat{\gamma}_{ij} &= \sigma^2_\gamma\ \hat{a}_i / \sigma^2_a = 4\ \hat{a}_i. \\
\sum_i \hat{\gamma}_{ij} &= \sigma^2_\gamma\ \hat{b}_j / \sigma^2_b = 1.5\ \hat{b}_j.
\end{aligned}$$

The pseudo variances come from (15.18) and (15.19).

$$\begin{aligned}
\sigma^2_{*a} &= \frac{3}{2}\ (4) - \frac{3}{(2)3}\ (6) = 3. \\
\sigma^2_{*b} &= \frac{4}{3}\ (9) - \frac{4}{(2)3}\ (6) = 8. \\
\sigma^2_{*\gamma} &= \frac{(3)4}{(2)3}\ (6) = 12.
\end{aligned}$$

Upper left 8 × 8

$$120^{-1} \begin{pmatrix}
186 & 72 & 66 & 48 & 72 & 24 & 24 & 66 \\
 & 112 & 0 & 0 & 30 & 12 & 18 & 12 \\
 & & 106 & 0 & 24 & 12 & 0 & 30 \\
 & & & 88 & 18 & 0 & 6 & 24 \\
 & & & & 87 & 0 & 0 & 0 \\
 & & & & & 39 & 0 & 0 \\
 & & & & & & 39 & 0 \\
 & & & & & & & 81
\end{pmatrix} \tag{37}$$

Lower left 12 × 8 and (upper right 12 × 8)'

$$120^{-1} \begin{pmatrix} 30 & 30 & 0 & 0 & 30 & 0 & 0 & 0 \\ 12 & 12 & 0 & 0 & 0 & 12 & 0 & 0 \\ 18 & 18 & 0 & 0 & 0 & 0 & 18 & 0 \\ 12 & 12 & 0 & 0 & 0 & 0 & 0 & 12 \\ 24 & 0 & 24 & 0 & 24 & 0 & 0 & 0 \\ 12 & 0 & 12 & 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 30 & 0 & 30 & 0 & 0 & 0 & 0 & 30 \\ 18 & 0 & 0 & 18 & 18 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 6 & 0 & 0 & 6 & 0 \\ 24 & 0 & 0 & 24 & 0 & 0 & 0 & 24 \end{pmatrix} \tag{38}$$

Lower 12 × 12

$$= 120^{-1} \text{ diag } (40,\ 22,\ 28,\ 22,\ 34,\ 22,\ 10,\ 40,\ 28,\ 10,\ 16,\ 34) \tag{39}$$

Approximate tests of hypotheses can be effected as described in the previous section. $\mathbf{K}'$ for SS Rows is (times .25)

$$\begin{pmatrix} 0 & 4 & 0 & -4 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 \\ 0 & 0 & 4 & -4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{pmatrix}.$$

$\mathbf{K}'$ for SS columns is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 3 & 0 & 0 & 3 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 3 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 3 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & -1 \end{pmatrix} /3.$$

8 The Two Way Mixed Model

The two way mixed model is one in which the elements of the rows (or columns) are a random sample from some population of rows (or columns), and the levels of columns (or rows) are fixed. We shall deal with random rows and fixed columns. There is really more than one type of mixed model, as we shall see, depending upon the variance-covariance matrices, $Var(\mathbf{a})$ and $Var(\boldsymbol{\gamma})$, and consequently $Var(\boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ = vector of elements, $\mu + a_i + b_j + \gamma_{ij}$. The most commonly used model is

$$Var(\boldsymbol{\alpha}) = \begin{pmatrix} \mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{C} \end{pmatrix}, \tag{40}$$

where $\mathbf{C}$ is $q \times q$, q being the number of columns. There are p such blocks down the diagonal, where p is the number of rows. $\mathbf{C}$ is a matrix with every diagonal = v and every off-diagonal = c. If the rows were sires and the columns were traits and if $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$, this would imply that the heritability is the same for every trait, 4 $v/(4v + \sigma_e^2)$, and the genetic correlation between any pair of traits is the same, c/v. This set of assumptions should be questioned in most mixed models. Is it logical to assume that $Var(\alpha_{ij}) = Var(\alpha_{ij'})$ and that $Cov(\alpha_{ij},\ \alpha_{ik}) = Cov(\alpha_{ij},\ \alpha_{im})$? Also is it logical to assume that $Var(e_{ijk}) = Var(e_{ij'k})$? Further we cannot necessarily assume that α_{ij} is uncorrelated with $\alpha_{i'j}$. This would not be true if the i^{th} sire is related to the i' sire. We shall deal more specifically with these problems in the context of multiple trait evaluation.

Now let us consider what assumptions regarding

$$Var\begin{pmatrix} \mathbf{a} \\ \boldsymbol{\gamma} \end{pmatrix}$$

will lead to $Var(\boldsymbol{\alpha})$ like (17.40). Two models commonly used in statistics accomplish this. The first is based on the model for unrelated interactions and main effects formulated in Section 15.4.

$$Var(\mathbf{a}) = \mathbf{I}\sigma_a^2,$$

since the number of levels of $\mathbf{a}$ in the population $\rightarrow \infty$, and

$$\begin{aligned}
Var(\gamma_{ij}) &= \sigma_\gamma^2. \\
Cov(\gamma_{ij},\ \gamma_{ij'}) &= -\sigma_\gamma^2/(q-1). \\
Cov(\gamma_{ij},\ \gamma_{i'j}) &= -\sigma_\gamma^2/(\text{one less than population levels of } a) = 0. \\
Cov(\gamma_{ij},\ \gamma_{i'j}) &= -\sigma_\gamma^2/(q-1)\ (\text{one less than population levels of } a) = 0.
\end{aligned}$$

This leads to

$$Var(\boldsymbol{\gamma}) = \begin{pmatrix} \mathbf{P} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{P} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \end{pmatrix} \sigma_y^2, \tag{41}$$

where $\mathbf{P}$ is a matrix with $1'$s in diagonals and $-1/(q-1)$ in all off-diagonals. Under this model

$$\begin{aligned}
Var(\alpha_{ij}) &= \sigma_a^2 + \sigma_\gamma^2. \\
Cov(\alpha_{ij},\ \alpha_{ij'}) &= \sigma_a^2 - \sigma_\gamma^2/(q-1).
\end{aligned} \tag{42}$$

An equivalent model often used that is easier from a computational standpoint, but less logical is

$$\begin{aligned}
Var(\mathbf{a}_*) &= \mathbf{I}\sigma_{*a}^2,\ \text{where } \sigma_{*a}^2 = \sigma_a^2 - \sigma_\gamma^2/(q-1). \\
Var(\gamma_*) &= \mathbf{I}\sigma_{*\gamma}^2,\ \text{where } \sigma_{*\gamma}^2 = q\sigma_\gamma^2/(q-1).
\end{aligned} \tag{43}$$

Note that we have re-labelled the row and interaction effects because these are not the same variables as $\mathbf{a}$ and $\boldsymbol{\gamma}$.

The results of (17.43) come from principles described in Section 15.9. We illustrate these two models (and estimation and prediction methods) with our same two way example. Let

$$\begin{aligned} Var(\mathbf{e}) &= 20\mathbf{I}, Var(\mathbf{a}) = 4\mathbf{I}, \text{ and} \\ Var(\boldsymbol{\gamma}) &= 6\begin{pmatrix} \mathbf{P} & 0 & 0 \\ 0 & \mathbf{P} & 0 \\ 0 & 0 & \mathbf{P} \end{pmatrix}, \end{aligned}$$

where $\mathbf{P}$ is a 4×4 matrix with 1's for diagonals and $-1/3$ for all off-diagonals. We set up the least squares equations with μ deleted, multiply the first 3 equations by 4 $\mathbf{I}_3$ and the last 12 equations by $Var(\boldsymbol{\gamma})$ described above. Then add 1 to the first 4 and the last 12 diagonal coefficients. This yields equations with coefficient matrix in (17.44) ... (17.47). The right hand side is (12.2, 7.2, 6.0, 3.15, .85, .8, 1.55, 5.9, -1.7, -.9, -3.3, 4.8, -1.2, -3.6, 0, 1.8, -3.0, -1.8, 3.0)′.

Upper left 10 × 10

$$\begin{pmatrix} 3.4 & 0 & 0 & 1.0 & .4 & .6 & .4 & 1.0 & .4 & .6 \\ 0 & 3.2 & 0 & .8 & .4 & 0 & 1.0 & 0 & 0 & 0 \\ 0 & 0 & 2.6 & .6 & 0 & .2 & .8 & 0 & 0 & 0 \\ .25 & .2 & .15 & .6 & 0 & 0 & 0 & .25 & 0 & 0 \\ .1 & .1 & 0 & 0 & .2 & 0 & 0 & 0 & .1 & 0 \\ .15 & 0 & .05 & 0 & 0 & .2 & 0 & 0 & 0 & .15 \\ .1 & .25 & .2 & 0 & 0 & 0 & .55 & 0 & 0 & 0 \\ .8 & 0 & 0 & 1.5 & -.2 & -.3 & -.2 & 2.5 & -.2 & -.3 \\ -.4 & 0 & 0 & -.5 & .6 & -.3 & -.2 & -.5 & 1.6 & -.3 \\ 0 & 0 & 0 & -.5 & -.2 & .9 & -.2 & -.5 & -.2 & 1.9 \end{pmatrix} \tag{44}$$

Upper right 10 × 9

$$\begin{pmatrix} .4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .8 & .4 & 0 & 1.0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .6 & 0 & .2 & .8 \\ 0 & .2 & 0 & 0 & 0 & .15 & 0 & 0 & 0 \\ 0 & 0 & .1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .05 & 0 \\ .1 & 0 & 0 & 0 & .25 & 0 & 0 & 0 & .2 \\ -.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{45}$$

Lower left 9 × 10

$$\begin{pmatrix} -.4 & 0 & 0 & -.5 & -.2 & -.3 & .6 & -.5 & -.2 & -.3 \\ 0 & .5 & 0 & 1.2 & -.2 & 0 & -.5 & 0 & 0 & 0 \\ 0 & -.3 & 0 & -.4 & .6 & 0 & -.5 & 0 & 0 & 0 \\ 0 & -1.1 & 0 & -.4 & -.2 & 0 & -.5 & 0 & 0 & 0 \\ 0 & .9 & 0 & -.4 & -.2 & 0 & 1.5 & 0 & 0 & 0 \\ 0 & 0 & .4 & .9 & 0 & -.1 & -.4 & 0 & 0 & 0 \\ 0 & 0 & -.8 & -.3 & 0 & -.1 & -.4 & 0 & 0 & 0 \\ 0 & 0 & -.4 & -.3 & 0 & .3 & -.4 & 0 & 0 & 0 \\ 0 & 0 & .8 & -.3 & 0 & -.1 & 1.2 & 0 & 0 & 0 \end{pmatrix} \quad (46)$$

Lower right 9 × 9

$$\begin{pmatrix} 1.6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2.2 & -.2 & 0 & -.5 & 0 & 0 & 0 & 0 \\ 0 & -.4 & 1.6 & 0 & -.5 & 0 & 0 & 0 & 0 \\ 0 & -.4 & -.2 & 1.0 & -.5 & 0 & 0 & 0 & 0 \\ 0 & -.4 & -.2 & 0 & 2.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.9 & 0 & -.1 & -.4 \\ 0 & 0 & 0 & 0 & 0 & -.3 & 1.0 & -.1 & -.4 \\ 0 & 0 & 0 & 0 & 0 & -.3 & 0 & 1.3 & -.4 \\ 0 & 0 & 0 & 0 & 0 & -.3 & 0 & -.1 & 2.2 \end{pmatrix} \quad (47)$$

The solution is

$$\begin{aligned} \hat{\mathbf{a}} &= (.563,\ -.437,\ -.126)'. \\ \hat{\mathbf{b}} &= (5.140,\ 4.218,\ 3.712,\ 2.967)'. \\ \hat{\boldsymbol{\gamma}} &= \begin{pmatrix} .104 & .163 & -.096 & -.170 \\ .421 & -.226 & .219 & -.414 \\ -.524 & .063 & -.122 & .584 \end{pmatrix}. \end{aligned}$$

The $\hat{\gamma}_{ij}$ sum to 0 by rows and columns.

When we employ the model with $Var(\mathbf{a}_*) = 2\mathbf{I}$ and $Var(\boldsymbol{\gamma}_*) = 8\mathbf{I}$, the coefficient matrix is in (17.48) ... (17.50) and the right hand side is (3.05, 1.8, 1.5, 3.15, .85, .8, 1.55, 1.5, .55, .65, .35, 1.05, .3, 0, .45, .6, 0, .15, .75)′.

Upper left 7 × 7

$$\begin{pmatrix} 1.1 & 0 & 0 & .25 & .1 & .15 & .1 \\ & 1.05 & 0 & .2 & .1 & 0 & .25 \\ & & .9 & .15 & 0 & .05 & .2 \\ & & & .6 & 0 & 0 & 0 \\ & & & & .2 & 0 & 0 \\ & & & & & .2 & 0 \\ & & & & & & .55 \end{pmatrix} \quad (48)$$

Lower left 12 $\times$ 7 and (upper right 7 $\times$ 12)$'$

$$\begin{pmatrix} .25 & 0 & 0 & .25 & 0 & 0 & 0 \\ .1 & 0 & 0 & 0 & .1 & 0 & 0 \\ .15 & 0 & 0 & 0 & 0 & .15 & 0 \\ .1 & 0 & 0 & 0 & 0 & 0 & .1 \\ 0 & .2 & 0 & .2 & 0 & 0 & 0 \\ 0 & .1 & 0 & 0 & .1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .25 & 0 & 0 & 0 & 0 & .25 \\ 0 & 0 & .15 & .15 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .05 & 0 & 0 & .05 & 0 \\ 0 & 0 & .2 & 0 & 0 & 0 & .2 \end{pmatrix} \tag{49}$$

Lower right 12 $\times$ 12

$$= \text{diag}\ (.375,\ .225,\ .275,\ .225,\ .325,\ .225,\ .125,\ .375,\ .275,\ .125,\ .175,\ .325). \tag{50}$$

The solution is

$$\begin{aligned} \hat{\mathbf{a}} &= (.282,\ -.219,\ -.063)',\ \text{different from above.} \\ \hat{\mathbf{b}} &= (5.140,\ 4.218,\ 3.712,\ 2.967)',\ \ \text{the same as before.} \\ \hat{\boldsymbol{\gamma}} &= \begin{pmatrix} .385 & .444 & .185 & .112 \\ .202 & -.444 & 0 & -.632 \\ -.588 & 0 & -.185 & .521 \end{pmatrix}, \end{aligned}$$

different from above. Now the $\hat{\gamma}$ sum to 0 by columns, but not by rows. This sum is

$$\sigma^2_{*\gamma}\hat{a}_i/\sigma^2_{*a} = 4\hat{a}_i.$$

As we should expect, the predictions of subclass means are identical in the two solutions. These are

$$\begin{pmatrix} 5.807 & 4.945 & 4.179 & 3.360 \\ 5.124 & 3.555 & 3.493 & 2.116 \\ 4.490 & 4.155 & 3.463 & 3.425 \end{pmatrix}.$$

These are all unbiased, including missing subclasses. This is in contrast to the situation in which both rows and columns are fixed. Note, however, that we should not predict μ_{ij} except for $j = 1, 2, 3, 4$. We could predict μ_{ij} (j=1,2,3,4) for $i > 3$, that is for rows not in the sample. BLUP would be $\hat{b}_j$. Remember, that $\hat{b}_j$ is BLUP of b_j + the mean of all $\mathbf{a}_i$ in the infinite population, and $\mathbf{a}_i$ is BLUP of $\mathbf{a}_i$ minus the mean of all $\mathbf{a}_i$ in the population.

We could, if we choose, obtain biased estimators and predictors by using some prior on the squares and products of $\mathbf{b}$, say

$$\begin{pmatrix} 1 & & \frac{-1}{3} \\ & \ddots & \\ \frac{-1}{3} & & 1 \end{pmatrix} \sigma^2_b,$$

where σ_b^2 is a pseudo-variance.

Suppose we want to estimate the variances. In that case the model with

$$Var(\mathbf{a}_*) = \mathbf{I}\sigma_{*a}^2 \text{ and } Var(\boldsymbol{\gamma}_*) = \mathbf{I}\sigma_{*\gamma}^2$$

is obviously easier to deal with than the pedagogically more logical model with $Var(\boldsymbol{\gamma})$ not a diagonal matrix. If we want to use that model, we can estimate σ_{*a}^2 and $\sigma_{*\gamma}^2$ and then by simple algebra convert those to estimates of σ_a^2 and σ_γ^2 .

Chapter 18
The Three Way Classification

C. R. Henderson

1984 - Guelph

This chapter deals with a 3 way classification model,

$$y_{ijkm} = \mu + a_i + b_j + c_k + ab_{ij} + ac_{ik} + bc_{jk} + abc_{ijk} + e_{ijkm}. \tag{1}$$

We need to specify the distributional properties of the elements of this model.

1 The Three Way Fixed Model

We first illustrate a fixed model with $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$. A simple way to approach this model is to write it as

$$y_{ijkm} = \mu_{ijk} + e_{ijkm}. \tag{2}$$

Then BLUE of μ_{ijk} is $\overline{y}_{ijk.}$ provided $n_{ijk} > 0$. Also BLUE of

$$\sum_i \sum_j \sum_k p_{ijk}\, \mu_{ijk} = \sum_i \sum_j \sum_k p_{ijk}\, \overline{y}_{ijk.},$$

where summation is over subclasses that are filled. But if subclasses are missing, there may not be linear functions of interest to the experimenter. Analogous to the two-way fixed model we have these definitions.

$$\begin{aligned}
a \text{ effects} &= \overline{\mu}_{i..} - \overline{\mu}_{...}, \\
b \text{ effects} &= \overline{\mu}_{.j.} - \overline{\mu}_{...}, \\
c \text{ effects} &= \overline{\mu}_{..k} - \overline{\mu}_{...}, \\
ab \text{ interactions} &= \overline{\mu}_{ij.} - \overline{\mu}_{i..} - \overline{\mu}_{.j.} + \overline{\mu}_{...}, \\
abc \text{ interactions} &= \mu_{ijk} - \overline{\mu}_{ij.} - \overline{\mu}_{i.k} - \overline{\mu}_{.jk} \\
&\quad + \overline{\mu}_{i..} + \overline{\mu}_{.j.} + \overline{\mu}_{..k} - \overline{\mu}_{...}\,.
\end{aligned} \tag{3}$$

None of these is estimable if a single subclass is missing. Consequently, the usual tests of hypotheses cannot be effected exactly.

2 The Filled Subclass Case

Suppose we wish to test the hypotheses that a effects, b effects, c effects, ab interactions, ac interactions, bc interactions, and abc interactions are all zero where these are defined as in (18.3). Three different methods will be described. The first two involve setting up least squares equations reparameterized by

$$\begin{aligned}\sum_i a_i &= \Sigma\, b_j = \Sigma\, c_k = 0 \\ \sum_j ab_{ij} &= 0 \text{ for all } i, \text{ etc.} \\ \sum_{jk} abc_{ijk} &= 0 \text{ for all } i, \text{ etc.}\end{aligned} \tag{4}$$

We illustrate this with a $2 \times 3 \times 4$ design with subclass numbers and totals as follows

	n_{ijk}											
	b_1				b_2				b_3			
a	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4
1	3	5	2	6	5	2	1	4	5	2	1	1
2	7	2	5	1	6	2	4	3	3	4	6	1

	$y_{ijk.}$											
	b_1				b_2				b_3			
a	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4
1	53	110	41	118	91	31	9	55	96	31	8	12
2	111	43	89	9	95	26	61	35	52	55	97	10

The first 7 columns of $\overline{\mathbf{X}}$ are

$$\begin{pmatrix}
1 & 1 & 1 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & -1 & -1 & -1 \\
1 & 1 & 0 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 0 & 1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & 1 & 0 & 0 \\
1 & 1 & -1 & -1 & 0 & 1 & 0 \\
1 & 1 & -1 & -1 & 0 & 0 & 1 \\
1 & 1 & -1 & -1 & -1 & -1 & -1 \\
1 & -1 & 1 & 0 & 1 & 0 & 0 \\
1 & -1 & 1 & 0 & 0 & 1 & 0 \\
1 & -1 & 1 & 0 & 0 & 0 & 1 \\
1 & -1 & 1 & 0 & -1 & -1 & -1 \\
1 & -1 & 0 & 1 & 1 & 0 & 0 \\
1 & -1 & 0 & 1 & 0 & 1 & 0 \\
1 & -1 & 0 & 1 & 0 & 0 & 1 \\
1 & -1 & 0 & 1 & -1 & -1 & -1 \\
1 & -1 & -1 & -1 & 1 & 0 & 0 \\
1 & -1 & -1 & -1 & 0 & 1 & 0 \\
1 & -1 & -1 & -1 & 0 & 0 & 1 \\
1 & -1 & -1 & -1 & -1 & -1 & -1
\end{pmatrix}.$$

The first column pertains to μ, the second to $\mathbf{a}$, the next two to $\mathbf{b}$, and the last 3 to $\mathbf{c}$. The remaining 17 columns are formed by operations on columns 2-7. Column 8 is formed by taking the products of corresponding elements of columns 2 and 3. Thus these are 1(1), 1(1), 1(1), 1(1), 1(0), ..., -1(-1). The other columns are as follows: $9 = 2 \times 4$, $10 = 2 \times 5$, $11 = 2 \times 6$, $12 = 2 \times 7$, $13 = 3 \times 5$, $14 = 3 \times 6$, $15 = 3 \times 7$, $16 = 4 \times 5$, $17 = 4 \times 6$, $18 = 4 \times 7$, $19 = 2 \times 13$, $20 = 2 \times 14$, $21 = 2 \times 15$, $22 = 2 \times 16$, $23 = 2 \times 17$, $24 = 2 \times 18$.

This gives the following for columns 8-16 of $\overline{\mathbf{X}}$

$$\begin{pmatrix}
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & -1 \\
-1 & -1 & 1 & 0 & 0 & -1 & 0 & 0 & -1 \\
-1 & -1 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\
-1 & -1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\
-1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\
-1 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\
-1 & 0 & 1 & 1 & 1 & -1 & -1 & -1 & 0 \\
0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & -1 \\
1 & 1 & -1 & 0 & 0 & -1 & 0 & 0 & -1 \\
1 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\
1 & 1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}$$

and for columns 17-24 of $\overline{\mathbf{X}}$,

$$\begin{pmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
-1 & -1 & 0 & 0 & 0 & -1 & -1 & -1 \\
0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\
-1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\
0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\
-1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & -1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & -1 & -1 & -1 & -1 & -1 & -1
\end{pmatrix}.$$

Then the least squares coefficient matrix is $\overline{\mathbf{X}}'\mathbf{N}\overline{\mathbf{X}}$, where $\mathbf{N}$ is a diagonal matrix of n_{ijk}. The right hand sides are $\overline{\mathbf{X}}'\mathbf{y}.$, where $\mathbf{y}.$ is the vector of subclass totals. The coefficient matrix of the equations is in (18.5) … (18.7). The right hand side is (1338, -28, 213, 42, 259, 57, 66, 137, 36, -149, -83, -320, -89, -38, -80, -30, -97, -103, -209, -16, -66, -66, 11, 19)'.

Upper left 12 × 12

$$
\begin{pmatrix}
81 & -7 & 8 & 4 & 13 & 1 & 3 & 6 & 2 & -9 & -5 & -17 \\
 & 81 & 6 & 2 & -9 & -5 & -17 & 8 & 4 & 13 & 1 & 3 \\
 & & 54 & 23 & -3 & -4 & -5 & -4 & -5 & -11 & 0 & -3 \\
 & & & 50 & -2 & -7 & -7 & -5 & -8 & -4 & 1 & 1 \\
 & & & & 45 & 16 & 16 & -11 & -4 & 3 & 6 & 6 \\
 & & & & & 33 & 16 & 0 & 1 & 6 & 7 & 6 \\
 & & & & & & 35 & -3 & 1 & 6 & 6 & -5 \\
 & & & & & & & 54 & 23 & -3 & -4 & -5 \\
 & & & & & & & & 50 & -2 & -7 & -7 \\
 & & & & & & & & & 45 & 16 & 16 \\
 & & & & & & & & & & 33 & 16 \\
 & & & & & & & & & & & 35
\end{pmatrix}. \tag{5}
$$

Upper right 12 × 12 and (lower left 12 × 12)'

$$
\begin{pmatrix}
-3 & -4 & -5 & -2 & -7 & -7 & -11 & 0 & -3 & -4 & 1 & 1 \\
-11 & 0 & -3 & -4 & 1 & 1 & -3 & -4 & -5 & -2 & -7 & -7 \\
9 & 4 & 5 & 6 & 4 & 5 & -7 & -4 & -13 & 2 & -2 & -5 \\
6 & 4 & 5 & 10 & 1 & 3 & 2 & -2 & -5 & 0 & -3 & -9 \\
7 & 5 & 5 & 8 & 5 & 5 & -1 & 5 & 5 & -2 & 1 & 1 \\
5 & 6 & 5 & 5 & 3 & 5 & 5 & 10 & 5 & 1 & 3 & 1 \\
5 & 5 & 5 & 5 & 5 & 3 & 5 & 5 & 7 & 1 & 1 & 3 \\
-7 & -4 & -13 & 2 & -2 & -5 & 9 & 4 & 5 & 6 & 4 & 5 \\
2 & -2 & -5 & 0 & -3 & -9 & 6 & 4 & 5 & 10 & 1 & 3 \\
-1 & 5 & 5 & -2 & 1 & 1 & 7 & 5 & 5 & 8 & 5 & 5 \\
5 & 10 & 5 & 1 & 3 & 1 & 5 & 6 & 5 & 5 & 3 & 5 \\
5 & 5 & 7 & 1 & 1 & 3 & 5 & 5 & 5 & 5 & 5 & 3
\end{pmatrix}. \tag{6}
$$

Lower right 12 × 12

$$
\begin{pmatrix}
27 & 9 & 9 & 10 & 2 & 2 & 3 & 5 & 5 & 2 & 0 & 0 \\
 & 22 & 9 & 2 & 8 & 2 & 5 & 6 & 5 & 0 & -2 & 0 \\
 & & 23 & 2 & 2 & 9 & 5 & 5 & -3 & 0 & 0 & -5 \\
 & & & 28 & 9 & 9 & 2 & 0 & 0 & 2 & 1 & 1 \\
 & & & & 19 & 9 & 0 & -2 & 0 & 1 & -1 & 1 \\
 & & & & & 21 & 0 & 0 & -5 & 1 & 1 & -7 \\
 & & & & & & 27 & 9 & 9 & 10 & 2 & 2 \\
 & & & & & & & 22 & 9 & 2 & 8 & 2 \\
 & & & & & & & & 23 & 2 & 2 & 9 \\
 & & & & & & & & & 28 & 9 & 9 \\
 & & & & & & & & & & 19 & 9 \\
 & & & & & & & & & & & 21
\end{pmatrix}. \tag{7}
$$

The resulting solution is (15.3392, .5761, 2.6596, -1.3142, 2.0092, 1.5358, -.8864, 1.3834, -.4886, .4311, .2156, -2.5289, -3.2461, 2.2154, 2.0376, .9824, -1.3108, -1.0136, -1.4858, -1.9251, 1.9193, .6648, .9469, -6836).

One method for finding the numerator sums of squares is to compare reductions, that is, subtracting the reduction when each factor and interaction is deleted from the reduction under the full model. For A, equation and unknown 2 is deleted, for B equations 3 and 4 are deleted, ..., for ABC equations 19-24 are deleted. The reduction under the full model is 22879.49 which is also simply

$$\sum_i \sum_j \sum_k y_{ijk.}^2 / n_{ijk}.$$

The sums of squares with their d.f. are as follows.

	d.f.	SS
A	1	17.88
B	2	207.44
C	3	192.20
AB	2	55.79
AC	3	113.25
BC	6	210.45
ABC	6	92.73

The denominator MS to use is $\hat{\sigma}_e^2 = (\mathbf{y}'\mathbf{y}$ - reduction in full model)/(81-24), where 81 is n, and 24 is the rank of the full model coefficient matrix.

A second method, usually easier, is to compute for the numerator

$$\text{SS} = (\boldsymbol{\beta}_i^o)'(Var(\boldsymbol{\beta}_i^o))^{-1}\boldsymbol{\beta}_i^o\sigma_e^2. \tag{8}$$

$\boldsymbol{\beta}_i^o$ is a subvector of the solution, β_2^o for A; β_3^o, β_4^o for B, ..., β_{17}^o, ..., β_{24}^o for ABC. $Var(\boldsymbol{\beta}_i^o)$ is the corresponding diagonal block of the inverse of the 24 $\times$ 24 coefficient matrix, not shown, multiplied by σ_e^2 . Thus

$$\text{SS for } A = .5761\ (.0186)^{-1}\ .5761,$$

$$\text{SS for } B = (2.6596 \quad -1.3142) \begin{pmatrix} .0344 & -.0140 \\ -.0140 & .0352 \end{pmatrix}^{-1} \begin{pmatrix} 2.6596 \\ -1.3142 \end{pmatrix},$$

etc. The terms inverted are diagonal blocks of the inverse of the coefficient matrix. These give the same results as by the first method.

The third method is to compute

$$(\mathbf{K}_i'\hat{\boldsymbol{\mu}})'\,(Var(\mathbf{K}_i'\hat{\boldsymbol{\mu}}))^{-1}\,\mathbf{K}'\hat{\boldsymbol{\mu}}\sigma_e^2. \tag{9}$$

$\mathbf{K}_i'\boldsymbol{\mu} = 0$ is the hypothesis tested for the i^{th} SS. $\hat{\boldsymbol{\mu}}$ is BLUE of $\boldsymbol{\mu}$, the vector of μ_{ijk}, and this is the vector of $\overline{y}_{ijk}$.

$\mathbf{K}_A$ is the 2nd column of $\overline{\mathbf{X}}$.

$\mathbf{K}_B$ is columns 3 and 4 of $\overline{\mathbf{X}}$.

$\vdots$

$\mathbf{K}_{ABC}$ is the last 6 columns of $\overline{\mathbf{X}}$.

For example, $\mathbf{K}_B'$ for SSB is

$$\begin{pmatrix} \mathbf{1} & \mathbf{0} & -\mathbf{1} & \mathbf{1} & \mathbf{0} & -\mathbf{1} \\ \mathbf{0} & \mathbf{1} & -\mathbf{1} & -\mathbf{1} & \mathbf{1} & -\mathbf{1} \end{pmatrix},$$

where $\mathbf{1} = (1\ 1\ 1\ 1)$ and $\mathbf{0} = (0\ 0\ 0\ 0)$.

$$Var(\hat{\boldsymbol{\mu}})/\sigma_e^2 = \mathbf{N}^{-1},$$

where $\mathbf{N}$ is the diagonal matrix of n_{ijk}. Then

$$Var(\mathbf{K}'\hat{\boldsymbol{\mu}})^{-1}\sigma_e^2 = (\mathbf{K}'\mathbf{N}^{-1}\mathbf{K})^{-1}.$$

This method leads to the same sums of squares as the other 2 methods.

3 Missing Subclasses In The Fixed Model

When one or more subclasses is missing, the usual estimates and tests of main effects and interactions cannot be made. If one is satisfied with estimating and testing functions like $\mathbf{K}'\boldsymbol{\mu}$, where $\boldsymbol{\mu}$ is the vector of μ_{ijk} corresponding to filled subclasses, BLUE and exact tests are straightforward. BLUE of

$$\mathbf{K}'\boldsymbol{\mu} = \mathbf{K}'\overline{\mathbf{y}}, \tag{10}$$

where $\overline{\mathbf{y}}.$ is the vector of means of filled 3 way subclasses. The numerator SS for testing the hypothesis that $\mathbf{K}'\boldsymbol{\mu} = \mathbf{c}$ is

$$(\mathbf{K}'\overline{\mathbf{y}}. - \mathbf{c})'\ \ Var(\mathbf{K}'\overline{\mathbf{y}}.)^{-1}(\mathbf{K}'\overline{\mathbf{y}}. - \mathbf{c})\sigma_e^2. \tag{11}$$

$$Var(\mathbf{K}'\overline{\mathbf{y}}.)/\sigma_e^2 = \mathbf{K}'\mathbf{N}^{-1}\mathbf{K}, \tag{12}$$

where $\mathbf{N}$ is a diagonal matrix of subclass numbers. The statistic of (18.11) is distributed as central $\chi^2\sigma_e^2$ with d.f. equal to the number of linearly independent rows of $\mathbf{K}'$. Then the corresponding MS divided by $\hat{\sigma}_e^2$ is distributed as F under the null hypothesis.

Unfortunately, if many subclasses are missing, the experimenter may have difficulty in finding functions of interest to estimate and test. Most of them wish correctly or otherwise to find estimates and tests that mimic the filled subclass case. Clearly this is possible only if one is prepared to use biased estimators and approximate tests of the functions whose estimators are biased.

We illustrate some biased methods with the following $2 \times 3 \times 4$ example.

	n_{ijk}											
	b_1				b_2				b_3			
a	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4
1	3	5	2	6	5	2	0	4	5	2	0	0
2	7	2	5	0	6	2	4	3	3	4	6	0

	$y_{ijk.}$											
	b_1				b_2				b_3			
a	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4
1	53	110	41	118	91	31	–	55	96	31	–	–
2	111	43	89	–	95	26	61	35	52	55	97	–

Note that 5 of the potential 24 *abc* subclasses are empty and one of the potential 12 *bc* subclasses is empty. All *ab* and *ac* subclasses are filled. Some common procedures are

1. Estimate and test main effects pretending that no interactions exist.

2. Estimate and test main effects, *ac* interactions, and *bc* interactions pretending that *bc* and *abc* interactions do not exist.

3. Estimate and test under a model in which interactions sum to 0 and in which each of the 5 missing *abc* and the one missing *bc* interactions are assumed = 0.

All of these clearly are biased methods, and their "goodness" depends upon the closeness of the assumptions to the truth. If one is prepared to use biased estimators, it seems more logical to me to attempt to minimize mean squared errors by using prior values for average sums of squares and products of interactions. Some possibilities for our example are:

1. Priors on **abc** and **bc**, the interactions associated with missing subclasses.

2. Priors on all interactions.

3. Priors on all interactions and on all main effects.

Obviously there are many other possibilities, e.g. priors on **c** and all interactions.

The first method above might have the greatest appeal since it results in biases due only to **bc** and **abc** interactions. No method for estimating main effects exists that does not contain biases due to these. But the first method does avoid biases due to main effects, **ab**, and **ac** interactions. This method will be illustrated. Let μ, **a**, **b**, **c**, **ab**, **ac** be treated as fixed. Consequently we have much confounding among them. The rank of the submatrix of $\mathbf{X}'\mathbf{X}$ pertaining to them is 1 + (2-1) + (3-1) + (4-1) + (2-1)(3-1) + (2-1)(4-1) = 12. We set up least squares equations with **ab**, **ac**, **bc**, and **abc** including missing subclasses for **bc** and **abc**. The submatrix for **ab** and **ac** has order, 14 and rank, 12. Treating **bc** and **abc** as random results in a mixed model coefficient matrix with order 50, and rank 48. The OLS coefficient matrix is in (18.13) to (18.18). The upper 26 $\times$ 26 block is in (18.13) to (18.15), the upper right 26 $\times$ 24 block is in (18.16) to (18.17), and the lower 24 $\times$ 24 block is in (18.18).

$$\begin{pmatrix}
16 & 0 & 0 & 0 & 0 & 0 & 3 & 5 & 2 & 6 & 0 & 0 & 0 \\
 & 11 & 0 & 0 & 0 & 0 & 5 & 2 & 0 & 4 & 0 & 0 & 0 \\
 & & 7 & 0 & 0 & 0 & 5 & 2 & 0 & 0 & 0 & 0 & 0 \\
 & & & 14 & 0 & 0 & 0 & 0 & 0 & 0 & 7 & 2 & 5 \\
 & & & & 15 & 0 & 0 & 0 & 0 & 0 & 6 & 2 & 4 \\
 & & & & & 13 & 0 & 0 & 0 & 0 & 3 & 4 & 6 \\
 & & & & & & 13 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 9 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & 2 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 10 & 0 & 0 & 0 \\
 & & & & & & & & & & 16 & 0 & 0 \\
 & & & & & & & & & & & 8 & 0 \\
 & & & & & & & & & & & & 15
\end{pmatrix} \tag{13}$$

$$\begin{pmatrix}
0 & 3 & 5 & 2 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 5 & 2 & 0 & 4 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 2 & 0 & 0 \\
0 & 7 & 2 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 6 & 2 & 4 & 3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 4 & 6 & 0 \\
0 & 3 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 5 & 0 & 0 & 0 \\
0 & 0 & 5 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\
0 & 7 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 4 & 0 & 0 \\
0 & 0 & 0 & 5 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 6 & 0
\end{pmatrix} \tag{14}$$

$$
\begin{pmatrix}
3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\
 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 11 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 4 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & 7 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 8 & 0 & 0 & 0 \\
 & & & & & & & & & & 6 & 0 & 0 \\
 & & & & & & & & & & & 6 & 0 \\
 & & & & & & & & & & & & 0
\end{pmatrix} \tag{15}
$$

$$
\begin{pmatrix}
3 & 5 & 2 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 5 & 2 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 \\
0 & 5 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 6 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 \\
0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix} \tag{16}
$$

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 6 & 2 & 4 & 3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 4 & 6 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 6 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 4 & 0 & 0 \\
0 & 5 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 6 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}, \tag{17}
$$

$$
\text{diag}(3, 5, 2, 6, 5, 2, 0, 4, 5, 2, 0, 0, 7, 2, 5, 0, 6, 2, 4, 3, 3, 4, 6, 0). \tag{18}
$$

The right hand side is

$$
\begin{array}{rrrrrrrrrrr}
[322, & 177, & 127, & 243, & 217, & 204, & 240, & 172, & 41, & 173, & 258, \\
124, & 247, & 35, & 164, & 153, & 130, & 118, & 186, & 57, & 61, & 90, \\
148, & 86, & 97, & 0, & 53, & 110, & 41, & 118, & 91, & 31, & 0, \\
55, & 96, & 31, & 0, & 0, & 111, & 43, & 89, & 0, & 95, & 26, \\
61, & 35, & 52, & 55, & 97, & 0] & & & & &
\end{array}
$$

We use the diagonalization method and assume that the pseudo-variances are $\sigma^2_{bc} = .3\ \sigma^2_e$, $\sigma^2_{abc} = .6\ \sigma^2_e$. Accordingly we add $.3^{-1}$ to the 15-26 diagonals and $.6^{-1}$ to the 27-50 diagonals of the OLS equations. This gives the following solution

$$
\begin{aligned}
\mathbf{ab} &= (20.664,\ 16.724,\ 17.812,\ 0,\ -3.507,\ -2.487)' \\
\mathbf{ac} &= (.047,\ -.618,\ 0,\ -1.949,\ 18.268,\ 17.976,\ 18.401,\ 15.441)' \\
\mathbf{bc} &= (-1.132,\ 1.028,\ -.164,\ .268,\ .541,\ -.366,\ .093,\ -.268, \\
&\quad\ .591,\ -.662,\ .071,\ 0)'
\end{aligned}
$$

$$\mathbf{abc} = (-1.229,\ .694,\ 0,\ .535,\ .666,\ -.131,\ 0,\ -.535,\ .563,$$
$$-.563,\ 0,\ 0,\ -1.034,\ 1.362,\ -.328,\ 0,\ .416,\ -.601,$$
$$.186,\ 0,\ .618,\ -.760,\ .142,\ 0)'.$$

The biased estimator of μ_{ijk} is $ab^o_{ij} + ac^o_{ik} + bc^o_{jk} + abc^o_{ijk}$. These are in tabular form ordered c in b in a by rows.

$$\mathbf{K} = 8^{-1} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ -1 & -1 \\ -1 & -1 \\ -1 & -1 \\ -1 & -1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ -1 & -1 \\ -1 & -1 \\ -1 & -1 \\ -1 & -1 \end{pmatrix}, \quad \hat{\boldsymbol{\mu}} = \begin{pmatrix} 18.35 \\ 21.77 \\ 20.50 \\ 19.52 \\ 17.98 \\ 15.61 \\ 16.82 \\ 13.97 \\ 19.01 \\ 15.97 \\ 17.88 \\ 15.86 \\ 16.10 \\ 20.37 \\ 17.91 \\ 15.71 \\ 15.72 \\ 13.50 \\ 15.17 \\ 11.67 \\ 16.99 \\ 14.07 \\ 16.13 \\ 12.95 \end{pmatrix}$$

The variance-covariance matrix of these $\hat{\mu}_{ijk}$ is $\overline{\mathbf{X}}\mathbf{C}\overline{\mathbf{X}}'\sigma^2_e$, where $\overline{\mathbf{X}}$ is the 24 x 50 incidence matrix for $\overline{y}_{ijk}$, and $\mathbf{C}$ is a g-inverse of the mixed model coefficient matrix. Approximate tests of hypotheses of $\mathbf{K}'\boldsymbol{\mu} = \mathbf{c}$ can be effected by computing

$$(\mathbf{K}'\hat{\boldsymbol{\mu}} - \mathbf{c})'[\mathbf{K}'\overline{\mathbf{X}}\mathbf{C}\overline{\mathbf{X}}'\mathbf{K}]^{-1}(\mathbf{K}'\hat{\boldsymbol{\mu}} - \mathbf{c})/(\text{rank }(\mathbf{K}'\overline{\mathbf{X}})\hat{\sigma}^2_e).$$

Under the null hypothesis this is distributed approximately as F.

To illustrate suppose we wish to test that all $\overline{\mu}_{.j.}$ are equal. $\mathbf{K}'$ and $\hat{\boldsymbol{\mu}}$ for this test are shown above and $\mathbf{c} = \mathbf{0}$. $\mathbf{K}'\hat{\boldsymbol{\mu}} = (2.66966 \quad -1.05379)'$. The pseudo-variances, σ^2_{bc} and σ^2_{abc}, could be estimated quite easily by Method 3. One could estimate σ^2_e by $\mathbf{y}'\mathbf{y}$ - reduction under full model, and this is simply

$$\mathbf{y}'\mathbf{y} - \sum_i \sum_j \sum_k y^2_{ijk.}/n_{ijk}.$$

Then we divide by n - the number of filled subclasses. Three reductions are needed to estimate σ^2_{bc} and σ^2_{abc}. The easiest ones are probably

Red (full model) described above.

Red (ab,ac,bc).

Red (ab,ac).

Partition the OLS coefficient matrix as

$$(\mathbf{C}_1 \quad \mathbf{C}_2 \quad \mathbf{C}_3).$$

$\mathbf{C}_1$ represents the first 14 cols., $\mathbf{C}_2$ the next 12, and $\mathbf{C}_3$ the last 24. Then compute $\mathbf{C}_2\mathbf{C}_2'$ and $\mathbf{C}_3\mathbf{C}_3'$. Let $\mathbf{Q}_2$ be the g-inverse of the matrix for Red (ab,ac,bc), which is the LS coefficient matrix with rows (and cols.) 27-50 set to 0. $\mathbf{Q}_3$ is the g-inverse for Red (ab,ac), which is the LS coefficient matrix with rows (and cols.) 15-50 set to 0. Then

$$\begin{aligned} E[\text{Red (full)}] &= 19\sigma^2_e + n(\sigma^2_{bc} + \sigma^2_{abc}) + t, \\ E[\text{Red }(ab, ac, bc)] &= 17\sigma^2_e + n\sigma^2_{bc} + tr\mathbf{Q}_2\mathbf{C}_3\mathbf{C}_3'\sigma^2_{abc} + t, \\ E[\text{Red }(ab, ac) &= 12\sigma^2_e + tr\mathbf{Q}_3\mathbf{C}_2\mathbf{C}_2'\sigma^2_{bc} + tr\mathbf{Q}_3\mathbf{C}_3\mathbf{C}_3'\sigma^2_{abc} + t. \end{aligned}$$

t is a quadratic in the fixed effects. The coefficient of σ^2_e is in each case the rank of the coefficient matrix used in the reduction.

4 The Three Way Mixed Model

Mixed models could be of two general types, namely one factor fixed and two random such as **a** fixed and **b** and **c** random, or with two factors fixed and one factor random, e.g. **a** and **b** fixed with **c** random. In either of these we would need to consider whether the populations are finite or infinite and whether the elements are related in any way. With **a** and **b** fixed and **c** random we would have fixed **ab** interaction and random **ac**, **bc**, **abc** interactions. With **a** fixed and **b** and **c** random all interactions would be random.

We also need to be careful about what we can estimate and predict. With **a** fixed and **b** and **c** random we can predict elements of **ab**, **ac**, and **abc** only for the levels of **a** in the experiment. With **a** and **b** fixed we can predict elements of **ac**, **bc**, **abc** only for the levels of both **a** and **b** in the experiment. For infinite populations of **b** and **c** in the first case and **c** in the second we can predict for levels of **b** and **c** (or **c**) outside the experiment. BLUP of them is 0. Thus in the case with **c** random, **a** and **b** fixed, BLUP of the 1,2,20 subclass when the number of levels of **c** in the experiment <20, is

$$\mu^o + a^o_1 + b^o_2 + ab^o_{12}.$$

In contrast, if the number of levels of **c** in the experiment >19, BLUP is

$$\mu^o + a_1^o + b_2^o + \hat{c}_{20} + \hat{ac}_{1,20} + \hat{bc}_{2,20} + ab_{12}^o + \hat{abc}_{12,20}.$$

In the case with $\mathbf{a}, \mathbf{b}$ fixed and $\mathbf{c}$ random, we might choose to place a prior on **ab**, especially if **ab** subclasses are missing in the data. The easiest way to do this would be to treat **ab** as a pseudo random variable with variance $= \mathbf{I}\sigma^2_{ab}$, which could be estimated. We could also use priors on **a** and **b** if we choose, and then the mixed model equations would mimic the 3 way random model.

Chapter 19
Nested Classifications

C. R. Henderson

1984 - Guelph

The nested classification can be described as cross-classification with disconnectedness. For example, we could have a cross-classified design with the main factors being sires and dams. Often the design is such that a set of dams is mated to sire 1 a second set to sire 2, etc. Then σ^2_d and σ^2_{ds}, dams assumed random, cannot be estimated separately, and the sum of these is defined as $\sigma^2_{d/s}$. As is the case with cross-classified data, estimability and methods of analysis depend upon what factors are fixed versus random. We assume that the only possibilities are random within random, random within fixed, and fixed within fixed. Fixed within random is regarded as impossible from a sampling viewpoint.

1 Two Way Fixed Within Fixed

A linear model for fixed effects nested within fixed effects is

$$y_{ijk} \; = \; t_i + a_{ij} + e_{ijk}$$

with t_i and a_{ij} fixed. The j subscript has no meaning except in association with some i subscript. None of the t_i is estimable nor are differences among the t_i. So far as the a_{ij} are concerned

$$\textstyle\sum_j \alpha_j a_{ij} \text{ for } \sum_j \alpha_j \; = \; 0 \text{ can be estimated.}$$

Thus we can estimate $2a_{i1} - a_{i2} - a_{i3}$. In contrast it is not possible to estimate differences between a_{ij} and a_{gh} $(i \neq g)$ or between a_{ij} and a_{gh} $(i \neq g, j \neq h)$. Obviously main effects can be defined only as some averaging over the nested factors. Thus we could define the mean of the i^{th} main factor as $\alpha_i = t_i + \sum_j k_j a_{ij}$ where $\sum_j k_j = 1$. Then the i^{th} main effect would be defined as α_i -$\tilde{\alpha}$. Tests of hypotheses of estimable linear functions can be effected in the usual way, that is, by utilizing the variance-covariance matrix of the estimable functions.

Let us illustrate with the following simple example

t	a	n_{ij}	y_{ij}	$\bar{y}_{ij.}$
1	1	4	20	5
	2	5	15	3
2	3	1	8	8
	4	10	70	7
	5	2	12	6
3	6	5	45	9
	7	2	16	8

Assume that $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$.

Main effects	$\hat{\alpha}_i$	$Var(\hat{\alpha}_i)$
1	4.	$\sigma_e^2\ (4^{-1} + 5^{-1})/4 = .1125\ \sigma_e^2$
2	7.	$\sigma_e^2\ (1 + 10^{-1} + 2^{-1})/9 = .177\ \sigma_e^2$
3	8.5	$\sigma_e^2\ (5^{-1} + 2^{-1})/4 = .175\ \sigma_e^2$

Test

$$\begin{aligned}
\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \boldsymbol{\alpha} &= \mathbf{0} \\
\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \hat{\boldsymbol{\alpha}} &= \begin{pmatrix} -3.5 \\ -.5 \end{pmatrix} \\
\sigma_e^{-2} \text{ Var } (\hat{\boldsymbol{\alpha}}) &= \text{dg } (.1125, .177 \ldots, .175) \\
Var(\mathbf{K}'\hat{\boldsymbol{\alpha}})\sigma_e^{-2} &= \begin{pmatrix} .2875 & .175 \\ .175 & .35278 \end{pmatrix}
\end{aligned}$$

with inverse

$$\begin{pmatrix} 4.98283 & -2.47180 \\ -2.47180 & 4.06081 \end{pmatrix}.$$

Then the numerator MS is

$$(-3.5 \quad -.5) \begin{pmatrix} 4.98283 & -2.47180 \\ -2.47180 & 4.06081 \end{pmatrix} \begin{pmatrix} -3.5 \\ -.5 \end{pmatrix} /2 = 26.70$$

Estimate σ_e^2 as $\hat{\sigma}_e^2$ = within subclass mean square. Then the test is numerator MS/$\hat{\sigma}_e^2$ with 2,26 d.f. A possible test of differences among a_{ij} could be

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \mathbf{a},$$

the estimate of which is (2 2 1 1)' with

$$Var = \begin{pmatrix} .45 & 0 & 0 & 0 \\ 0 & 1.5 & .5 & 0 \\ 0 & .5 & .6 & 0 \\ 0 & 0 & 0 & .7 \end{pmatrix},$$

the inverse of which is

$$\begin{pmatrix} 2.22222 & 0 & 0 & 0 \\ 0 & .92308 & -.76923 & 0 \\ 0 & -.76923 & 2.30769 & 0 \\ 0 & 0 & 0 & 1.42857 \end{pmatrix}.$$

This gives the numerator MS = 13.24.

The "usual" ANOVA described in many text books is as follows.

$$\begin{aligned} \text{S.S. for T} &= \textstyle\sum_i y_{i..}^2/n_{i.} \;-\; y_{...}^2/n_{..}. \\ \text{S.S. for A} &= \textstyle\sum_i\sum_j y_{ij.}^2/n_{ij} \;-\; \sum_i y_{i..}^2/n_{i.}. \end{aligned}$$

In our example,

$$\begin{aligned} \text{MST} &= (1290.759 - 1192.966)/2 = 48.897. \\ \text{MSA} &= (1304 - 1290.759)/1 = 13.24. \end{aligned}$$

Note that the latter is the same as in the previous method. They do in fact test the same hypothesis. But MST is different from the result above which tests treatments averaged equally over the **a** nested within it. The second method tests differences among **t** weighted over **a** according to the number of observations. Thus the weights for t_1 are (4,5)/9.

To illustrate this test,

$$\mathbf{K}' = \begin{pmatrix} .444444 & .555555 & 0 & 0 & 0 & -.71429 & -.28572 \\ 0 & 0 & .07692 & .76923 & .15385 & -.71429 & -.28572 \end{pmatrix}$$

$$Var(\mathbf{K}'\bar{y}_{ij})/\sigma_e^2 = \begin{pmatrix} .253968 & .142857 \\ .142857 & .219780 \end{pmatrix}$$

$$\text{with inverse } = \begin{pmatrix} 6.20690 & -4.03448 \\ -4.03448 & 7.17242 \end{pmatrix}$$

Then the MS is

$$(-4.82540 \;\; -1.79122) \begin{pmatrix} 6.20690 & -4.03448 \\ -4.03448 & 7.17242 \end{pmatrix} \begin{pmatrix} -4.82540 \\ -1.79122 \end{pmatrix} /2 \;=\; 48.897$$

as in the regular ANOVA. Thus ANOVA weights according to the n_{ij}. This does not appear to be a particularly interesting test.

2 Two Way Random Within Fixed

There are two different sampling schemes that can be envisioned in the random nested within fixed model. In one case, the random elements associated with every fixed factor are assumed to be a sample from the same population. A different situation is one in which the elements within each fixed factor are assumed to be from separate populations. The first type could involve treatments as the fixed factors and then a random sample of sires is drawn from a common population to assign to a particular treatment. In contrast, if the main factors are breeds, then the sires sampled would be from separate populations, namely the particular breeds. In the first design we can estimate the difference among treatments, each averaged over the same population of sires. In the second case we would compare breeds defined as the average of all sires in each of the respective breeds.

2.1 Sires within treatments

We illustrate this design with a simple example

	n_{ij} Treatments			$y_{ij.}$ Treatments		
Sires	1	2	3	1	2	3
1	5	0	0	7	-	-
2	2	0	0	6	-	-
3	0	3	0	-	7	-
4	0	8	0	-	9	-
5	0	0	5	-	-	8

Let us treat this first as a multiple trait problem with $Var(\mathbf{e}) = 40\mathbf{I}$,

$$Var\begin{pmatrix} s_{i1} \\ s_{i2} \\ s_{i3} \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 5 \end{pmatrix},$$

where s_{ij} refers to the value of the i^{th} sire with respect to the j^{th} treatment. Assume that the sires are unrelated. The inverse is

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 5 \end{pmatrix}^{-1} = \begin{pmatrix} .5 & -.25 & 0 \\ -.25 & .4375 & -.125 \\ 0 & -.125 & .25 \end{pmatrix}.$$

Then the mixed model equations are (19.1).

$$
80^{-1}\begin{pmatrix}
50 & -20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0 & 0 \\
-20 & 35 & -10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -10 & 20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 44 & -20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\
0 & 0 & 0 & -20 & 35 & -10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -10 & 20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 40 & -20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -20 & 41 & -10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -10 & 20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 40 & -20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -20 & 51 & -10 & 0 & 0 & 0 & 0 & 16 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -10 & 20 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 40 & -20 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -20 & 35 & -10 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -10 & 30 & 0 & 0 & 10 \\
10 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 14 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 16 & 0 & 0 & 0 & 0 & 0 & 22 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0 & 0 & 10
\end{pmatrix}
$$

$$
\begin{gathered}
(s_{11},\ s_{12},\ s_{13},\ s_{21},\ s_{22},\ s_{23},\ s_{31},\ s_{32},\ s_{33}, \\
s_{41},\ s_{42},\ s_{43}, s_{51},\ s_{52},\ s_{53}, t_1,\ t_2,\ t_3)' \\
= [.175,\ 0,\ 0, .15,\ 0,\ 0,\ 0,\ .175, 0, 0,\ .225, \\
0,\ 0,\ 0,\ .2,\ .325,\ .4,\ .2]'.
\end{gathered} \tag{1}
$$

The solution is

$$(-.1412,\ -.0941,\ -.0471,\ .1412,\ .0941,\ .0471,\ .0918,\ .1835,$$
$$.0918,\ -.0918,\ -.1835, -.0918,\ 0,\ 0,\ 0,\ 1.9176,\ 1.5380,\ 1.600)'. \tag{2}$$

Now if we treat this as a nested model, $\mathbf{G}$ = diag (3,3,4,4,5). Then the mixed model equations are in (19.3).

$$120^{-1}\begin{pmatrix} 55 & 0 & 0 & 0 & 0 & 15 & 0 & 0 \\ & 46 & 0 & 0 & 0 & 6 & 0 & 0 \\ & & 39 & 0 & 0 & 0 & 9 & 0 \\ & & & 54 & 0 & 0 & 24 & 0 \\ & & & & 39 & 0 & 0 & 15 \\ & & & & & 21 & 0 & 0 \\ & & & & & & 33 & 0 \\ & & & & & & & 15 \end{pmatrix}\begin{pmatrix} \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{s}_4 \\ \hat{s}_5 \\ \hat{t}_1 \\ \hat{t}_2 \\ \hat{t}_3 \end{pmatrix} = 120^{-1}\begin{pmatrix} 21 \\ 18 \\ 21 \\ 27 \\ 24 \\ 39 \\ 48 \\ 24 \end{pmatrix} \tag{3}$$

The solution is

$$(-.1412,\ .1412,\ .1835,\ -.1835,\ 0,\ 1.9176,\ 1.5380,\ 1.6000)'. \tag{4}$$

Note that $\hat{s}_{11} = \hat{s}_1$, $\hat{s}_{21} = \hat{s}_2$, $\hat{s}_{32} = \hat{s}_3$, $\hat{s}_{42} = \hat{s}_4$, $\hat{s}_{53} = \hat{s}_5$ from the solution in (19.2) and (19.4). Also note that t_j are equal in the two solutions. The second method is certainly easier than the first but it does not predict values of sires for treatments in which they had no progeny.

2.2 Sires within breeds

Now we assume that we have a population of sires unique to each breed. Then the first model of Section 19.2.1 would be useless. The second method illustrated would be appropriate if sires were unrelated and σ_s^2 = 3,4,5 for the 3 breeds. If σ_s^2 were the same for all breeds $\mathbf{G} = \mathbf{I}_5\sigma_s^2$.

3 Random Within Random

Let us illustrate this model by dams within sires. Suppose the model is

$$\begin{aligned} y_{ijk} &= \mu + s_i + d_{ij} + e_{ijk}. \\ Var\begin{pmatrix} \mathbf{s} \\ \mathbf{d} \\ \mathbf{e} \end{pmatrix} &= \begin{pmatrix} \mathbf{I}\sigma_s^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_d^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2 \end{pmatrix}. \end{aligned}$$

Let us use the data of Section 19.2.1 but now let **t** refer to sires and **s** to dams. Suppose $\sigma_e^2/\sigma_s^2 = 12$, $\sigma_e^2/\sigma_d^2 = 10$. Then the mixed model equations are in (19.5).

$$\begin{pmatrix} 23 & 7 & 11 & 5 & 5 & 2 & 3 & 8 & 5 \\ & 19 & 0 & 0 & 5 & 2 & 0 & 0 & 0 \\ & & 23 & 0 & 0 & 0 & 3 & 8 & 0 \\ & & & 17 & 0 & 0 & 0 & 0 & 5 \\ & & & & 15 & 0 & 0 & 0 & 0 \\ & & & & & 12 & 0 & 0 & 0 \\ & & & & & & 13 & 0 & 0 \\ & & & & & & & 18 & 0 \\ & & & & & & & & 15 \end{pmatrix} \begin{pmatrix} \hat{\mu}_1 \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{d}_1 \\ \hat{d}_2 \\ \hat{d}_3 \\ \hat{d}_4 \\ \hat{d}_5 \end{pmatrix} = \begin{pmatrix} 37 \\ 13 \\ 16 \\ 8 \\ 7 \\ 6 \\ 7 \\ 9 \\ 8 \end{pmatrix} \tag{5}$$

The solution is (l.6869, .0725, -.0536, -.0l89, -.ll98, .2068, .l6l6, -.2259, -.0227)$'$. Note that $\sum \hat{s}_i = 0$ and that 10 (Sum of $\hat{d}$ within i^{th} sire)/12 $= \hat{s}_i$.

Chapter 20
Analysis of Regression Models

C. R. Henderson

1984 - Guelph

A regression model is one in which $\mathbf{Zu}$ does not exist, the first column of $\mathbf{X}$ is a vector of 1's, and all other elements of $\mathbf{X}$ are general (not 0's and 1's) as in the classification model. The elements of $\mathbf{X}$ other than the first column are commonly called covariates or independent variables. The latter is not a desirable description since they are not variables but rather are constants. In hypothetical repeated sampling the value of $\mathbf{X}$ remains constant. In contrast $\mathbf{e}$ is a sample from a multivariate population with mean $= \mathbf{0}$ and variance $= \mathbf{R}$, often $\mathbf{I}\sigma_e^2$. Accordingly $\mathbf{e}$ varies from one hypothetical sample to the next. It is usually assumed that the columns of $\mathbf{X}$ are linearly independent, that is, $\mathbf{X}$ has full column rank. This should not be taken for granted in all situations, for it could happen that linear dependencies exist. A more common problem is that near but not complete dependencies exist. In that case, $(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-1}$ can be quite inaccurate, and the variance of some or all of the elements of $\hat{\boldsymbol{\beta}}$ can be extremely large. Methods for dealing with this problem are discussed in Section 20.2.

1 Simple Regression Model

The most simple regression model is

$$y_i = \mu + w_i\gamma + e_i,$$

where

$$\mathbf{X} = \begin{pmatrix} 1 & w_1 \\ 1 & w_2 \\ \vdots & \vdots \\ 1 & w_n \end{pmatrix}.$$

The most simple form of $Var(\mathbf{e}) = \mathbf{R}$ is $\mathbf{I}\sigma_e^2$. Then the BLUE equations are

$$\begin{pmatrix} n & w. \\ w. & \sum w_i^2 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} y. \\ \sum w_i y_i \end{pmatrix}. \quad (1)$$

To illustrate suppose n=5,

$$\mathbf{w}' = (6, 5, 3, 4, 2), \;\; \mathbf{y}' = (8, 6, 5, 6, 5).$$

The BLUE equations are

$$\frac{1}{\sigma_e^2}\begin{pmatrix} 5 & 20 \\ 20 & 90 \end{pmatrix}\begin{pmatrix} \hat{\mu} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} 30 \\ 127 \end{pmatrix} /\sigma_e^2.$$

The inverse of the coefficient matrix is

$$\begin{pmatrix} 1.8 & -.4 \\ -.4 & .1 \end{pmatrix} \sigma_e^2.$$

The solution is (3.2, .7).

$$Var(\hat{\mu}) = 1.8\ \sigma_e^2,\ Var(\hat{\gamma}) = .1\ \sigma_e^2, Cov(\hat{\mu},\hat{\gamma}) = -.4\ \sigma_e^2.$$

Some text books describe the model above as

$$y_i = \alpha + (w - \bar{w}_.)\gamma + e_i.$$

The BLUE equations in this case are

$$\begin{pmatrix} n & 0 \\ 0 & \sum(w_i - \bar{w}_.)^2 \end{pmatrix}\begin{pmatrix} \hat{\alpha} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} y_. \\ \sum(w_i - w_.)\bar{y}_i \end{pmatrix}. \qquad (2)$$

This gives the same solution to $\hat{\gamma}$ as (19.1) but $\hat{\mu} \neq \hat{\alpha}$ except when $\bar{w}_. = 0$. The equations of (20.2) in our example are

$$\frac{1}{\sigma_e^2}\begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix}\begin{pmatrix} \hat{\mu} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} 30 \\ 7 \end{pmatrix} /\sigma_e^2.$$

$$\hat{\alpha} = 6,\ \hat{\gamma} = .7.$$

$$Var(\hat{\alpha}) = .2\sigma_e^2,\ Var(\hat{\gamma}) = .1\sigma_e^2,\ Cov(\hat{\alpha},\hat{\gamma}) = 0.$$

It is easy to verify that $\hat{\mu} = \hat{\alpha} - \bar{w}_.\hat{\gamma}_.$. These two alternative models meet the requirements of linear equivalence, Section 1.5.

BLUP of a future y say y_0 with $w_i = w_0$ is

$$\hat{\mu} + w_0\hat{\gamma} + \hat{e}_0 \text{ or } \hat{\alpha} + (w_0 - \bar{w})\hat{\gamma} + \hat{e}_0,$$

where $\hat{e}_0$ is BLUP of $e_0 = 0$, with prediction error variance, σ_e^2. If $w_0 = 3$, y_0 would be 5.3 in our example. This result assumes that future μ or α) have the same value as in the population from which the original sample was taken. The prediction error variance is

$$(1\ \ 3)\begin{pmatrix} 1.8 & -.4 \\ -.4 & .1 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} \sigma_e^2 + \sigma_e^2 = 1.3\ \sigma_e^2.$$

Also using the second model it is

$$(1\ \ -1)\begin{pmatrix} .2 & 0 \\ 0 & .1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \sigma_e^2 + \sigma_e^2 = 1.3\ \sigma_e^2$$

as in the equivalent model.

2 Multiple Regression Model

In the multiple regression model the first column of $\mathbf{X}$ is a vector of 1's, and there are 2 or more additional columns of covariates. For example, the second column could represent age in days and the third column could represent initial weight, while $\mathbf{y}$ represents final weight. Note that in this model the regression on age is asserted to be the same for every initial weight. Is this a reasonable assumption? Probably it is not. A possible modification of the model to account for effect of initial weight upon the regression of final weight on age and for effect of age upon the regression of final weight on initial weight is

$$\begin{aligned} y_i &= \mu + \gamma_1 w_1 + \gamma_2 w_2 + \gamma_3 w_3 + e_i, \\ \text{where} \quad w_3 &= w_1 w_2. \end{aligned}$$

This model implies that the regression coefficient for $\mathbf{y}$ on w_1 is a simple linear function of w_2 and the regression coefficient for $\mathbf{y}$ on w_2 is a simple linear function of w_1. A model like this sometimes gives trouble because of the relationship between columns 2 and 3 with column 4 of $\mathbf{X}$. We illustrate with

$$\mathbf{X} = \begin{pmatrix} 1 & 6 & 8 & 48 \\ 1 & 5 & 9 & 45 \\ 1 & 5 & 8 & 40 \\ 1 & 6 & 7 & 42 \\ 1 & 7 & 9 & 63 \end{pmatrix}.$$

The elements of column 4 are the products of the corresponding elements of columns 2 and 3. The coefficient matrix is

$$\begin{pmatrix} 5 & 29 & 41 & 238 \\ & 171 & 238 & 1406 \\ & & 339 & 1970 \\ & & & 11662 \end{pmatrix}. \tag{3}$$

The inverse of this is

$$\begin{pmatrix} 4780.27 & -801.54 & -548.45 & 91.73 \\ & 135.09 & 91.91 & -15.45 \\ & & 63.10 & -10.55 \\ & & & 1.773 \end{pmatrix}. \tag{4}$$

Suppose that we wish to predict y for $w_1 = \bar{w}_{1.} = 5.8$, $\bar{w}_2 = 8.2$, $w_3 = 47.56 = (5.8)(8.2)$. The variance of the error of prediction is

$$(1 \;\; 5.8 \;\; 8.2 \;\; 47.56)(\text{matrix } 20.4) \begin{pmatrix} 1 \\ 5.8 \\ 8.2 \\ 47.56 \end{pmatrix} \sigma_e^2 \; + \; \sigma_e^2 \; = \; 1.203 \; \sigma_e^2$$

Suppose we predict y for $w_1 = 3$, $w_2 = 5$, $w_3 = 15$. Then the variance of the error of prediction is 215.77 σ_e^2, a substantial increase. The variance of the prediction error is extremely vulnerable to departures of w_i from $\bar{w}_i$.

Suppose we had not included w_3 in the model. Then the inverse of the coefficient matrix is

$$\begin{pmatrix} 33.974 & -1.872 & -2.795 \\ & .359 & -.026 \\ & & .359 \end{pmatrix}.$$

The variances of the errors of prediction of the two predictors above would then be 1.20 and 7.23, the second of which is much smaller than when w_3 is included. But if $w_3 \neq 0$, the predictor is biased when w_3 is not included.

Let us look at the solution when w_3 is included and $\mathbf{y}' = (6, 4, 8, 7, 5)$. The solution is

$$(157.82, \quad -23.64, \quad -17.36, \quad 2.68).$$

This is a strange solution that is the consequence of the large elements in $(\mathbf{X}'\mathbf{X})^{-1}$. A better solution might result if a prior is placed on w_3. When the prior is 1, we add 1 to the lower diagonal element of the coefficient matrix. The resulting solution is

$$(69.10, \quad -8.69, \quad -7.16, \quad .967).$$

This type of solution is similar to ridge regression, Hoerl and Kennard (1970). There is an extensive statistics literature on the problem of ill-behaved $\mathbf{X}'\mathbf{X}$. Most solutions to this problem that have been proposed are (1) biased (shrunken estimation) or (2) dropping one or more elements of $\boldsymbol{\beta}$ from the model with either backward or forward type of elimination, Draper and Smith (1966). See for example a paper by Dempster et al. (1977) with an extensive list of references. Also Hocking (1976) has many references.

Another type of covariate is involved in fitting polynomials, for example

$$y_i = \mu + x_i\gamma_1 + x_i^2\gamma_2 + x_i^3\gamma_3 + x_i^4\gamma_4 + e_i.$$

As in the case when covariates involve products, the sampling variances of predictors are large when x_i departs far from $\bar{x}$. The numerator mean square with 1 d.f. can be computed easily. For the i^{th} γ_i it is

$$\hat{\gamma}_i^2/c^{i+1},$$

where c^{i+1} is the i+1 diagonal of the inverse of the coefficient matrix. The numerator can also be computed by reduction under the full model minus the reduction when γ_i is dropped from the solution.

Chapter 21
Analysis of Covariance Model

C. R. Henderson

1984 - Guelph

A covariance model is one in which **X** has columns referring to levels of factors or interactions and one or more columns of covariates. The model may or may not contain **Zu**. It usually does not in text book discussions of covariance models, but in animal breeding applications there would be, or at least should be, a **u** vector, usually referring to breeding values.

1 Two Way Fixed Model With Two Covariates

Consider a model

$$y_{ijk} = r_i + c_j + \gamma_{ij} + w_{1ijk}\alpha_1 + w_{2ijk}\alpha_2 + e_{ijk}.$$

All elements of the model are fixed except for **e**, which is assumed to have variance, $\mathbf{I}\sigma_e^2$. The n_{ijk}, $y_{ijk.}$, $w_{1ijk.}$, and $w_{2ijk.}$ are as follows

	n_{ijk}			$y_{ijk.}$			$w_{1ijk.}$			$w_{2ijk.}$		
	1	2	3	1	2	3	1	2	3	1	2	3
1	3	2	1	20	9	4	8	7	2	12	11	4
2	1	3	4	3	20	24	6	10	11	5	15	14
3	2	1	2	13	7	8	7	2	4	9	2	7

and the necessary sums of squares and crossproducts are

$$\begin{aligned}
\textstyle\sum_i \sum_j \sum_k w_{1ijk}^2 &= 209, \\
\textstyle\sum_i \sum_j \sum_k w_{1ijk} w_{2ijk} &= 264, \\
\textstyle\sum_i \sum_j \sum_k w_{2ijk}^2 &= 373, \\
\textstyle\sum_i \sum_j \sum_k w_{1ijk} y_{ijk} &= 321, \\
\textstyle\sum_i \sum_j \sum_k w_{2ijk} y_{ijk} &= 433.
\end{aligned}$$

Then the matrix of coefficients of OLS equations are in (21.1). The right hand side vector is (33, 47, 28, 36, 36, 36, 20, 9, 4, 3, 20, 24, 13, 7, 8, 321, 433)'.

$$
\begin{pmatrix}
6 & 0 & 0 & 3 & 2 & 1 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 17 & 27 \\
 & 8 & 0 & 1 & 3 & 4 & 0 & 0 & 0 & 1 & 3 & 4 & 0 & 0 & 0 & 27 & 34 \\
 & & 5 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 13 & 18 \\
 & & & 6 & 0 & 0 & 3 & 0 & 0 & 1 & 0 & 0 & 2 & 0 & 0 & 21 & 26 \\
 & & & & 6 & 0 & 0 & 2 & 0 & 0 & 3 & 0 & 0 & 1 & 0 & 19 & 28 \\
 & & & & & 7 & 0 & 0 & 1 & 0 & 0 & 4 & 0 & 0 & 2 & 17 & 25 \\
 & & & & & & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 12 \\
 & & & & & & & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 & 11 \\
 & & & & & & & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 4 \\
 & & & & & & & & & 1 & 0 & 0 & 0 & 0 & 0 & 6 & 5 \\
 & & & & & & & & & & 3 & 0 & 0 & 0 & 0 & 10 & 15 \\
 & & & & & & & & & & & 4 & 0 & 0 & 0 & 11 & 14 \\
 & & & & & & & & & & & & 2 & 0 & 0 & 7 & 9 \\
 & & & & & & & & & & & & & 1 & 0 & 2 & 2 \\
 & & & & & & & & & & & & & & 2 & 4 & 7 \\
 & & & & & & & & & & & & & & & 209 & 264 \\
 & & & & & & & & & & & & & & & & 373
\end{pmatrix} \quad (1)
$$

A g-inverse of the coefficient matrix can be obtained by taking a regular inverse with the first 6 rows and columns set to 0. The lower 11 × 11 submatrix of the g-inverse is in (21.2).

$$
10^{-4}
\begin{pmatrix}
8685 & 7311 & 5162 & 7449 & 6690 & 4802 & 6163 & 2866 & 4588 & -285 & -1148 \\
 & 15009 & 7135 & 9843 & 9139 & 6507 & 8357 & 3832 & 6309 & -264 & -1652 \\
 & & 15313 & 5849 & 6452 & 4422 & 5694 & 2430 & 4592 & 226 & -1441 \\
 & & & 25714 & 9312 & 7519 & 9581 & 5325 & 5718 & 2401 & -262 \\
 & & & & 11695 & 6002 & 7704 & 3582 & 5735 & -356 & -1435 \\
 & & & & & 6938 & 5686 & 2780 & 4012 & -569 & -821 \\
 & & & & & & 12286 & 3551 & 5158 & -704 & -1072 \\
 & & & & & & & 11869 & 2290 & -654 & -281 \\
 & & & & & & & & 9016 & 6 & -1151 \\
 & & & & & & & & & 767 & -440 \\
 & & & & & & & & & & 580
\end{pmatrix} \quad (2)
$$

This gives a solution vector (0, 0, 0, 0, 0, 0, 7.9873, 6.4294, 5.7748, 2.8341, 8.3174, 6.8717, 7.6451, 7.2061, 5.3826, .6813, -.7843). One can test an hypothesis concerning interactions by subtracting from the reduction under the full model the reduction when $\boldsymbol{\gamma}$ is dropped from the model. This tests that all $\gamma_{ij} - \bar{\gamma}_{i.} - \bar{\gamma}_{.j} + \bar{\gamma}_{..}$ are 0. The reduction under the full model is 652.441. A solution with $\boldsymbol{\gamma}$ dropped is

$$(6.5808,\ 7.2026,\ 6.5141,\ 1.4134,\ 1.4386,\ 0,\ .1393,\ -.5915).$$

This gives a reduction = 629.353. Then the numerator SS with 4 d.f. is 652.441 - 629.353.

The usual test of hypothesis concerning rows is that all $r_i + \bar{c}_. + \bar{\gamma}_{i.}$ are equal. This is comparable to the test effected by weighted squares of means when there are no covariates. We could define the test as all $r_i + \bar{c}_. + \bar{\gamma}_{i.} + \alpha_1 w_{10} + \alpha_2 w_{20}$ are equal, where w_{10}, w_{20} can have any values. This is not valid, as shown in Section 16.6, when the regressions are not homogeneous. To find the numerator SS with 2 d.f. for rows take the matrix

$$\begin{aligned} \mathbf{K}' &= \begin{pmatrix} 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \end{pmatrix}. \\ \mathbf{K}'\hat{\boldsymbol{\gamma}} &= \begin{pmatrix} 2.1683 \\ -.0424 \end{pmatrix}, \end{aligned}$$

where $\hat{\boldsymbol{\gamma}}$ is the solution under the full model with $\mathbf{r}^o$, $\mathbf{c}^o$ set to $\mathbf{0}$. Next compute $\mathbf{K}'$ [first 9 rows and columns of (21.2)] $\mathbf{K}$ as

$$= \begin{pmatrix} 4.5929 & 2.2362 \\ 2.2362 & 4.3730 \end{pmatrix}.$$

Then

$$\begin{aligned} \text{numerator SS} &= (2.1683 \quad -.0424) \begin{pmatrix} 4.5929 & 2.2362 \\ 2.2362 & 4.3730 \end{pmatrix}^{-1} \begin{pmatrix} 2.1683 \\ -.0424 \end{pmatrix} \\ &= 1.3908. \end{aligned}$$

If we wish to test w_1, compute as the numerator SS, with 1 d.f., .6813 $(.0767)^{-1}$ .6813, where

$$\hat{\alpha}_1 = .6813,\ Var(\hat{\alpha}_1) = .0767\ \sigma_e^2.$$

2 Two Way Fixed Model With Missing Subclasses

We found in Section 17.3 that the two way fixed model with interaction and with one or more missing subclasses precludes obtaining the usual estimates and tests of main effects and interactions. This is true also, of course, in the covariance model with missing subclasses for fixed by fixed classifications. We illustrate with the same example as before except that the (3,3) subclass is missing. The OLS equations are in (21.3). The right hand side vector is (33, 47, 20, 36, 36, 28, 20, 9, 4, 3, 20, 24, 13, 7, 0, 307, 406)$'$. Note that the equation for γ_{33} is included even though the subclass is missing.

$$\left(\begin{array}{ccccccccccccccccc}
6 & 0 & 0 & 3 & 2 & 1 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 17 & 27 \\
 & 8 & 0 & 1 & 3 & 4 & 0 & 0 & 0 & 1 & 3 & 4 & 0 & 0 & 0 & 27 & 34 \\
 & & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 9 & 11 \\
 & & & 6 & 0 & 0 & 3 & 0 & 0 & 1 & 0 & 0 & 2 & 0 & 0 & 21 & 26 \\
 & & & & 6 & 0 & 0 & 2 & 0 & 0 & 3 & 0 & 0 & 1 & 0 & 19 & 28 \\
 & & & & & 5 & 0 & 0 & 1 & 0 & 0 & 4 & 0 & 0 & 0 & 13 & 18 \\
 & & & & & & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 12 \\
 & & & & & & & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 & 11 \\
 & & & & & & & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 4 \\
 & & & & & & & & & 1 & 0 & 0 & 0 & 0 & 0 & 6 & 5 \\
 & & & & & & & & & & 3 & 0 & 0 & 0 & 0 & 10 & 15 \\
 & & & & & & & & & & & 4 & 0 & 0 & 0 & 11 & 14 \\
 & & & & & & & & & & & & 2 & 0 & 0 & 7 & 9 \\
 & & & & & & & & & & & & & 1 & 0 & 2 & 2 \\
 & & & & & & & & & & & & & & 0 & 0 & 0 \\
 & & & & & & & & & & & & & & & 199 & 249 \\
 & & & & & & & & & & & & & & & & 348
\end{array}\right) \tag{3}$$

We use these equations to estimate a pseudo-variance, σ^2_γ to use in biased estimation with priors on $\boldsymbol{\gamma}$. We use Method 3. Reductions and expectations are

$$\begin{aligned}
\mathbf{y}'\mathbf{y} &= 638,\ E(\mathbf{y}'\mathbf{y}) = 17\ \sigma^2_e + 17\ \sigma^2_\gamma + q. \\
\text{Red (full)} &= 622.111,\ E() = 10\ \sigma^2_e + 17\ \sigma^2_\gamma + q. \\
\text{Red } (\mathbf{r}, \mathbf{c}, \gamma) &= 599.534,\ E() = 7\ \sigma^2_e + 12.6121\ \sigma^2_\gamma + q. \\
\text{q} &= \text{a quadratic in } \mathbf{r}, \mathbf{c}, \boldsymbol{\alpha}.
\end{aligned}$$

Solving we get $\hat{\sigma}^2_e = 2.26985$, $\hat{\sigma}^2_\gamma = 3.59328$ or a ratio of .632. Then we add .632 to each of the diagonal coefficients corresponding to γ equations in (21.3). A resulting solution is

$$\begin{gathered}
(6.6338,\ 6.1454,\ 7.3150,\ -.3217,\ .6457,\ 0,\ 1.3247,\ -.7287,\ -.5960, \\
-1.7830, 1.1870, .5960, .4583, -.4583, 0, .6179, -.7242)
\end{gathered}$$

The resulting biased estimates of $r_i + c_j + \gamma_{ij}$ given $w_1 = w_2 = 0$ are

$$\begin{pmatrix}
7.6368 & 6.5509 & 6.0378 \\
4.0407 & 7.9781 & 6.7414 \\
7.4516 & 7.5024 & 7.3150
\end{pmatrix} \tag{4}$$

The matrix of estimated mean squared errors obtained by pre and post multiplying a g-inverse of the coefficient matrix by

$$\mathbf{L}' = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & \vdots & & & & & & & & & & & & \vdots & & \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

is in (21.5).

$$10,000^{-1} \begin{pmatrix} 8778 & 7287 & 5196 & 8100 & 6720 & 4895 & 6151 & 3290 & 2045 \\ & 13449 & 6769 & 8908 & 8661 & 6017 & 7210 & 4733 & 2788 \\ & & 13215 & 4831 & 5606 & 4509 & 4931 & 2927 & 7191 \\ & & & 22170 & 9676 & 7524 & 9408 & 4825 & 1080 \\ & & & & 11201 & 6007 & 7090 & 4505 & 2540 \\ & & & & & 6846 & 5423 & 3214 & 3885 \\ & & & & & & 11244 & 4681 & 5675 \\ & & & & & & & 10880 & 6514 \\ & & & & & & & & 45120 \end{pmatrix} \quad (5)$$

To test that all $r_i + \bar{c}_. + \bar{\gamma}_{i.}$ are equal, use the matrix

$$\begin{pmatrix} 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \end{pmatrix} \text{ with (21.4) and (21.5).}$$

Then the numerator SS is

$$(1.4652 \quad -2.0434) \begin{pmatrix} 4.4081 & 1.6404 \\ & 9.2400 \end{pmatrix}^{-1} \begin{pmatrix} 1.4652 \\ -2.0434 \end{pmatrix} = 1.2636.$$

The test is approximate because the MSE depends upon $\hat{\sigma}_\gamma^2/\hat{\sigma}_e^2 = \sigma_\gamma^2/\sigma_e^2$. Further, the numerator is not distributed as χ^2.

3 Covariates All Equal At The Same Level Of A Factor

In some applications every $w_{ij} = w_i$ in a one-way covariate model,

$$y_{ij} = \mu + t_i + w_{ij}\gamma + e_{ij}$$

with all $w_{ij} = w_i$. For example, t_i might represent an animal in which there are several observations, y_{ij}, but the covariate is measured only once. This idea can be extended to multiple classifications. When the factor associated with the constant covariate is fixed, estimability problems exist, Henderson and Henderson (1979). In the one way case $t_i - t_j$ is not estimable and neither is γ.

We illustrate with a one-way case in which n_i = (3,2,4), w_i = (2,4,5), $\bar{y}_{i.}$ = (6,5,10). The OLS equations are

$$\begin{pmatrix} 9 & 3 & 2 & 4 & 34 \\ 3 & 3 & 0 & 0 & 6 \\ 2 & 0 & 2 & 0 & 8 \\ 4 & 0 & 0 & 4 & 20 \\ 34 & 6 & 8 & 20 & 144 \end{pmatrix} \begin{pmatrix} \mu \\ t_1 \\ t_2 \\ t_3 \\ \gamma \end{pmatrix} = \begin{pmatrix} 68 \\ 18 \\ 10 \\ 40 \\ 276 \end{pmatrix}. \qquad (6)$$

Note that equations 2,3,4 sum to equation 1 and also (2 4 5) times these equations gives the last equation. Accordingly the coefficient matrix has rank only 3, the same as if there were no covariate. A solution is (0,6,5,10,0).

If t is random, there is no problem of estimability for then we need only to look at the rank of

$$\begin{pmatrix} 9 & 34 \\ 34 & 144 \end{pmatrix},$$

and that is 2. Consequently μ and γ are both estimable, and of course $\mathbf{t}$ is predictable. Let us estimate σ_t^2 and σ_e^2 by Method 3 under the assumption $Var(\mathbf{t}) = \mathbf{I}\sigma_t^2$, $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$. For this we need **y**'**y**, reduction under the full model, and Red (μ, γ).

$$\begin{aligned} \mathbf{y}'\mathbf{y} &= 601,\ E(\mathbf{y}'\mathbf{y}) = 9\ \sigma_e^2 + 9\ \sigma_t^2 + q. \\ \text{Red (full)} &= 558,\ E() = 3\ \sigma_e^2 + 9\ \sigma_t^2 + q. \\ \text{Red } (\mu, \gamma) &= 537.257,\ E() = 2\ \sigma_e^2 + 6.6\ \sigma_t^2 + q. \end{aligned}$$

q is a quadratic in μ, γ. This gives estimates $\hat{\sigma}_e^2 = 7.167$, $\hat{\sigma}_t^2 = 5.657$ or a ratio of 1.27.

Let us use 1 as a prior value of σ_e^2/σ_t^2 and estimate σ_t^2 by MIVQUE given that σ_e^2 = 7.167. We solve for $\hat{t}$ having added 1 to the diagonal coefficients of equations 2,3,4 of (21.6). This gives an inverse,

$$\begin{pmatrix} 4.24370 & -1.63866 & -.08403 & .72269 & -1.02941 \\ & .94958 & .15126 & -.10084 & .35294 \\ & & .54622 & .30252 & -.05882 \\ & & & .79832 & -.29412 \\ & & & & .27941 \end{pmatrix}. \qquad (7)$$

The solution is (3.02521, .55462, -1.66387, 1.10924, 1.11765). From this $\hat{\mathbf{t}}'\hat{\mathbf{t}} = 4.30648$. To find its expectation we compute

$$tr(\mathbf{C}_t \text{ [matrix (21.6)] } \mathbf{C}_t') = tr \begin{pmatrix} .01483 & -.04449 & .02966 \\ & .13347 & -.08898 \\ & & .05932 \end{pmatrix} = .20761,$$

which is the coefficient of σ_e^2 in $E(\hat{\mathbf{t}}'\hat{\mathbf{t}})$. $\mathbf{C}_t$ is the submatrix composed of rows 2-4 of (21.7).

$$tr(\mathbf{C}_t\mathbf{W}'\mathbf{Z}_t\mathbf{Z}_t'\mathbf{W}\mathbf{C}_t') = tr\begin{pmatrix} .03559 & -.10677 & .07118 \\ & .32032 & -.21354 \\ & & .14236 \end{pmatrix} = .49827,$$

the coefficient of σ_t^2 in $E(\hat{\mathbf{t}}'\hat{\mathbf{t}})$. $\mathbf{W}'\mathbf{Z}_t$ is the submatrix composed of cols. 2-4 of (21.6). This gives $\hat{\sigma}_t^2 = 5.657$ or $\hat{\sigma}_e^2/\hat{\sigma}_t^2 = 1.27$. If we do another MIVQUE estimation of σ_t^2, given $\sigma_e^2 = 7.167$ using the ratio, 1.27, the same estimate of σ_t^2 is obtained. Accordingly we have REML of σ_t^2, given σ_e^2. Notice also that this is the Method 3 estimate.

If $\mathbf{t}$ were actually fixed, but we use a pseudo-variance in the mixed model equations we obtain biased estimators. Using $\hat{\sigma}_e^2/\hat{\sigma}_t^2 = 1.27$,

$$\begin{aligned} \hat{\mu} + \hat{t}_i &= (3.53,\ 1.48,\ 4.04). \\ \hat{\mu} + \hat{t}_i + w_i\hat{\gamma}_. &= (5.78,\ 5.98,\ 9.67). \end{aligned}$$

Contrast this last with the corresponding OLS estimates of (6,5,10).

4 Random Regressions

It is reasonable to assume that regression coefficients are random in some models. For example, suppose we have a model,

$$y_{ij} = \mu + c_i + w_{ij}\gamma_i + e_{ij},$$

where y_{ij} is a yield observation on the j^{th} day for the i^{th} cow, w_{ij} is the day, and γ_i is a regression coefficient, linear slope of yield on time. Linearity is a reasonable assumption for a relatively short period following peak production. Further, it is obvious that γ_i is different from cow to cow, and if cows are random, γ_i is also random. Consequently we should make use of this assumption. The following example illustrates the method. We have 4 random cows with 3,5,6,4 observations respectively. The OLS equations are in (21.8).

$$\begin{pmatrix} 18 & 3 & 5 & 6 & 4 & 10 & 30 & 19 & 26 \\ & 3 & 0 & 0 & 0 & 10 & 0 & 0 & 0 \\ & & 5 & 0 & 0 & 0 & 30 & 0 & 0 \\ & & & 6 & 0 & 0 & 0 & 19 & 0 \\ & & & & 4 & 0 & 0 & 0 & 26 \\ & & & & & 38 & 0 & 0 & 0 \\ & & & & & & 190 & 0 & 0 \\ & & & & & & & 67 & 0 \\ & & & & & & & & 182 \end{pmatrix} \begin{pmatrix} \mu^o \\ \mathbf{c}^o \\ \boldsymbol{\gamma}^o \end{pmatrix} = \begin{pmatrix} 90 \\ 14 \\ 18 \\ 26 \\ 32 \\ 51 \\ 117 \\ 90 \\ 216 \end{pmatrix} \quad (8)$$

$$
\begin{aligned}
10 &= w_{1.}, \\
30 &= w_{2.}, \text{ etc.} \\
38 &= \textstyle\sum_j w_{ij}^2, \text{ etc.} \\
51 &= \textstyle\sum_j w_{ij} y_{ij}, \text{ etc.}
\end{aligned}
$$

First let us estimate σ_e^2, σ_c^2, σ_γ^2 by Method 3. The necessary reductions and their expectations are

$$
E \begin{pmatrix} \mathbf{y}'\mathbf{y} \\ \text{Red (full)} \\ \text{Red } (\mu, \mathbf{t}) \\ \text{Red } (\mu, \boldsymbol{\gamma}) \end{pmatrix} = \begin{pmatrix} 18 & 18 & 477 \\ 8 & 18 & 477 \\ 4 & 18 & 442.5 \\ 5 & 16.9031 & 477 \end{pmatrix} \begin{pmatrix} \sigma_e^2 \\ \sigma_c^2 \\ \sigma_\gamma^2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} 18\ \mu^2.
$$

The reductions are (538, 524.4485, 498.8, 519.6894). This gives estimates $\hat{\sigma}_e^2 = 1.3552$, $\hat{\sigma}_c^2 = .6324$, $\hat{\sigma}_\gamma^2 = .5863$. Using the resulting ratios, $\hat{\sigma}_e^2/\hat{\sigma}_c^2 = 2.143$ and $\hat{\sigma}_e^2/\hat{\sigma}_\gamma^2 = 2.311$, the mixed model solution is

$$
\begin{gathered}
(2.02339,\ .11180,\ -.36513,\ -.09307, \\
.34639,\ .73548,\ .34970,\ .76934,\ .83764).
\end{gathered}
$$

Covariance models are discussed also in Chapter 16.

Chapter 22
Animal Model, Single Records

C. R. Henderson

1984 - Guelph

We shall describe a number of different genetic models and present methods for BLUE, BLUP, and estimation of variance and covariance components. The simplest situation is one in which we have only one trait of concern, we assume an additive genetic model, and no animal has more than a single record on this trait. The scalar model, that is, the model for an individual record, is

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + \mathbf{z}_i' \mathbf{u} + a_i + e_i.$$

$\boldsymbol{\beta}$ represents fixed effects with $\mathbf{x}_i$ relating the record on the i^{th} animal to this vector.

$\mathbf{u}$ represents random effects other than breeding values and $\mathbf{z}_i$ relates this vector to y_i.

a_i is the additive genetic value of the i^{th} animal.

e_i is a random error associated with the individual record.

The vector representation of the entire set of records is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{Z}_a\mathbf{a} + \mathbf{e}. \tag{1}$$

If $\mathbf{a}$ represents only those animals with records, $\mathbf{Z}_a = \mathbf{I}$. Otherwise it is an identity matrix with rows deleted that correspond to animals without records.

$$\begin{aligned}
Var(\mathbf{u}) &= \mathbf{G}. \\
Var(\mathbf{a}) &= \mathbf{A}\sigma_a^2. \\
Var(\mathbf{e}) &= \mathbf{R}, \text{ usually } \mathbf{I}\sigma_e^2. \\
Cov(\mathbf{u}, \mathbf{a}') &= \mathbf{0}, \\
Cov(\mathbf{u}, \mathbf{e}') &= \mathbf{0}, \\
Cov(\mathbf{a}, \mathbf{e}') &= \mathbf{0}.
\end{aligned}$$

If $\mathbf{Z}_a \neq \mathbf{I}$, the mixed model equations are

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_a \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}_a \\ \mathbf{Z}_a'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_a'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{Z}_a'\mathbf{R}^{-1}\mathbf{Z}_a + \mathbf{A}^{-1}/\sigma_a^2 \end{pmatrix}
\begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}_a'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{2}$$

If $\mathbf{Z}_a = \mathbf{I}$, (22.2) simplifies to

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{R}^{-1} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{Z}'\mathbf{R}^{-1} \\ \mathbf{R}^{-1}\mathbf{X} & \mathbf{R}^{-1}\mathbf{Z} & \mathbf{R}^{-1} + \mathbf{A}^{-1}/\sigma_a^2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{3}$$

If $\mathbf{R} = \mathbf{I}\sigma_e^2$ (22.3) simplifies further to

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}' \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{G}^{-1}\sigma_e^2 & \mathbf{Z}' \\ \mathbf{X} & \mathbf{Z} & \mathbf{I} + \mathbf{A}^{-1}\sigma_e^2/\sigma_a^2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{y} \end{pmatrix}. \tag{4}$$

If the number of animals is large, one should, of course, use Henderson's method (1976) for computing $\mathbf{A}^{-1}$. Because this method requires using a "base" population of non-inbred, unrelated animals, some of these probably do not have records. Also we may wish to evaluate some progeny that have not yet made a record. Both of these circumstances will result in $\mathbf{Z}_a \neq \mathbf{I}$, but $\hat{\mathbf{a}}$ will contain predicted breeding values of these animals without records.

1 Example With Dam-Daughter Pairs

We illustrate the model above with 5 pairs of dams and daughters, the dams' records being made in period 1 and the daughters' in period 2. Ordering the records within periods and with record 1 being made by the dam of the individual making record 6, etc.

$$\begin{aligned} \mathbf{X}' &= \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \\ \mathbf{Z}_a &= \mathbf{I}, \\ \mathbf{y}' &= [5, 4, 3, 2, 6, 6, 7, 3, 5, 4]. \\ \mathbf{A} &= \begin{pmatrix} \mathbf{I}_5 & .5\mathbf{I}_5 \\ .5\mathbf{I}_5 & \mathbf{I}_5 \end{pmatrix}, \\ \mathbf{R} &= \mathbf{I}_{10}\sigma_e^2. \end{aligned}$$

The sires and dams are all unrelated. We write the mixed model equations with σ_e^2/σ_a^2 assumed to be 5. These equations are

$$\begin{pmatrix} 5 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & & \frac{23}{3} & \mathbf{I}_5 & & \frac{-10}{3} & \mathbf{I}_5 & & & & \\ 1 & 0 & & & & & & & & & & \\ 1 & 0 & & & & & & & & & & \\ 1 & 0 & & & & & & & & & & \\ 1 & 0 & & & & & & & & & & \\ 0 & 1 & & & & & & & & & & \\ 0 & 1 & & & & & & & & & & \\ 0 & 1 & & \frac{-10}{3} & \mathbf{I}_5 & & \frac{23}{3} & \mathbf{I}_5 & & & & \\ 0 & 1 & & & & & & & & & & \\ 0 & 1 & & & & & & & & & & \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} 20 \\ 25 \\ 5 \\ 4 \\ 3 \\ 2 \\ 6 \\ 6 \\ 7 \\ 3 \\ 5 \\ 4 \end{pmatrix}. \quad (5)$$

The inverse of the coefficient matrix is

$$\begin{pmatrix} .24 & .02 & -.04 & -.04 & -.04 & -.04 & -.04 & -.02 & -.02 & -.02 & -.02 & -.02 \\ .02 & .24 & -.02 & -.02 & -.02 & -.02 & -.02 & -.04 & -.04 & -.04 & -.04 & -.04 \\ -.04 & -.02 & & & & & & & & & & \\ -.04 & -.02 & & & & & & & & & & \\ -.04 & -.02 & & & \mathbf{P} & & & & & \mathbf{Q} & & \\ -.04 & -.02 & & & & & & & & & & \\ -.04 & -.02 & & & & & & & & & & \\ -.02 & -.04 & & & & & & & & & & \\ -.02 & -.04 & & & & & & & & & & \\ -.02 & -.04 & & & \mathbf{Q} & & & & & \mathbf{P} & & \\ -.02 & -.04 & & & & & & & & & & \\ -.02 & -.04 & & & & & & & & & & \end{pmatrix} \quad (6)$$

P is a 5×5 matrix with .16867 in diagonals and .00783 in all off-diagonals. **Q** is a 5×5 matrix with .07594 in diagonals and .00601 in off-diagonals. The solution is (4, 5, .23077, .13986, -.30070, -.32168, .25175, .23077, .32168, -.39161, -.13986, -.20298).

Let us estimate σ_e^2, σ_a^2 by MIVQUE using the prior on $\sigma_e^2/\sigma_a^2 = 5$ as we did in computing BLUP. The quadratics needed are

$$\hat{\mathbf{e}}'\hat{\mathbf{e}} \text{ and } \hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}}.$$

$$\hat{\mathbf{e}} = \mathbf{y} - (\mathbf{X} \ \ \mathbf{Z}_a) \begin{pmatrix} \hat{\mathbf{p}} \\ \hat{\mathbf{a}} \end{pmatrix} = (.76923, \ -.13986, \ -.69930, \ -1.67832, \ 1.74825,$$
$$.76923, 1.67832, \ -1.60839, \ .13986, -.97902)'.$$

$$
\begin{aligned}
\hat{\mathbf{e}}'\hat{\mathbf{e}} &= 13.94689. \\
Var(\hat{\mathbf{e}}) &= (\mathbf{I} - \mathbf{WCW}')(\mathbf{I} - \mathbf{WCW}')\ \sigma_e^2 \\
&\quad + (\mathbf{I} - \mathbf{WCW}')\mathbf{A}(\mathbf{I} - \mathbf{WCW}')\ \sigma_a^2. \\
\mathbf{W} &= (\mathbf{X}\ \ \mathbf{Z}_a), \\
\mathbf{C} &= \text{matrix of (22.6).} \\
E(\hat{\mathbf{e}}'\hat{\mathbf{e}}) &= tr(Var(\hat{\mathbf{e}})) = 5.67265\ \sigma_e^2 + 5.20319\ \sigma_a^2 \\
\hat{\mathbf{a}}' &= \mathbf{C}_a\mathbf{W}'\mathbf{y}, \\
\text{where } \mathbf{C}_a &= \text{last 10 rows of } \mathbf{C}. \\
Var(\hat{\mathbf{a}}) &= \mathbf{C}_a\mathbf{W}'\mathbf{W}\mathbf{C}_a'\ \sigma_e^2 + \mathbf{C}_a\mathbf{W}'\mathbf{AW}\mathbf{C}_a'\ \sigma_a^2. \\
E(\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}}) &= tr(\mathbf{A}^{-1}\ Var(\hat{\mathbf{a}})) = .20813\ \sigma_e^2 + .24608\ \sigma_a^2. \\
\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}} &= .53929.
\end{aligned}
$$

Using these quadratics and solving for $\hat{\sigma}_e^2, \hat{\sigma}_a^2$ we obtain $\hat{\sigma}_e^2 = 2.00$, $\hat{\sigma}_a^2 = .50$.

The same estimates are obtained for any prior used for σ_e^2/σ_a^2. This is a consequence of the fact that we have a balanced design. Therefore the estimates are truly BQUE and also are REML. Further the traditional method, daughter-dam regression, gives the same estimates. These are

$$
\begin{aligned}
\hat{\sigma}_a^2 &= \text{2 times regression of daughter on dam.} \\
\hat{\sigma}_e^2 &= \text{within period mean square } - \ \hat{\sigma}_a^2.
\end{aligned}
$$

For unbalanced data MIVQUE is not invariant to the prior used, and daughter-dam regression is neither MIVQUE nor REML. We illustrate by assuming that y_{10} was not observed. With σ_e^2/σ_a^2 assumed equal to 2 we obtain

$$
\begin{aligned}
\hat{\mathbf{e}}'\hat{\mathbf{e}} &= 11.99524 \text{ with expectation } = 3.37891\ \sigma_e^2 + 2.90355\ \sigma_a^2. \\
\hat{\mathbf{a}}'\mathbf{A}^{-1}\mathbf{a}' &= 2.79712 \text{ with expectation } = .758125\ \sigma_e^2 + .791316\ \sigma_a^2.
\end{aligned}
$$

This gives

$$
\begin{aligned}
\hat{\sigma}_a^2 &= .75619, \\
\hat{\sigma}_e^2 &= 2.90022.
\end{aligned}
$$

When σ_e^2/σ_a^2 is assumed equal to 5, the results are

$$
\begin{aligned}
\hat{\mathbf{e}}'\hat{\mathbf{e}} &= 16.83398 \text{ with expectation } 4.9865\ \sigma_e^2 + 4.6311\ \sigma_a^2, \\
\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}}' &= .66973 \text{ with expectation } .191075\ \sigma_e^2 + .214215\ \sigma_a^2.
\end{aligned}
$$

Then

$$
\begin{aligned}
\hat{\sigma}_a^2 &= .67132, \\
\hat{\sigma}_e^2 &= 2.7524.
\end{aligned}
$$

Chapter 23
Sire Model, Single Records

C. R. Henderson

1984 - Guelph

A simple sire model is one in which sires, possibly related, are mated to a random sample of unrelated dams, no dam has more than one progeny with a record, and each progeny produces one record. A scalar model for this is

$$y_{ij} \;=\; \mathbf{x}_{ij}^{'}\boldsymbol{\beta} + \mathbf{s}_i + \mathbf{z}_i^{'}\mathbf{u} + \mathbf{e}_{ij}. \tag{1}$$

$\boldsymbol{\beta}$ represents fixed effects with $\mathbf{x}_{ij}$ relating the j^{th} progeny of the i^{th} sire to these effects.

$\mathbf{s}_i$ represents the sire effect on the progeny record.

$\mathbf{u}$ represents other random factors with $\mathbf{z}_{ij}$ relating these to the ij^{th} progeny record.

$\mathbf{e}_{ij}$ is a random "error".

The vector representation is

$$\mathbf{y} \;=\; \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_s\mathbf{s} + \mathbf{Z}\mathbf{u} + \mathbf{e}. \tag{2}$$

$Var(\mathbf{s}) = \mathbf{A}\sigma_s^2$, where $\mathbf{A}$ is the numerator relationship of the sires, and σ_s^2 is the sire variance in the "base" population. If the sires comprise a random sample from this population $\sigma_s^2 = \frac{1}{4}$ additive genetic variance. Some columns of $\mathbf{Z}_s$ will be null if $\mathbf{s}$ contains sires with no progeny, as will usually be the case if the simple method for computation of $\mathbf{A}^{-1}$ requiring base population animals, is used.

$$\begin{aligned} Var(\mathbf{u}) &= \mathbf{G},\; Cov(\mathbf{s}, \mathbf{u}') = \mathbf{0}. \\ Var(\mathbf{e}) &= \mathbf{R}, \text{usually} = \mathbf{I}\sigma_e^2. \\ Cov(\mathbf{s}, \mathbf{e}') &= \mathbf{0}, Cov(\mathbf{u}, \mathbf{e}') = \mathbf{0}. \end{aligned}$$

If sires and dams are truly random,

$$\begin{aligned} \mathbf{I}\sigma_e^2 &= .75\mathbf{I} \text{ (additive genetic variance)} \\ &\quad + \mathbf{I} \text{ (environmental variance)}. \end{aligned}$$

With this model the mixed model equations are

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_s & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}_s'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_s'\mathbf{R}^{-1}\mathbf{Z}_s + \mathbf{A}^{-1}\sigma_s^{-2} & \mathbf{Z}_s'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}_s & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{s}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}_s'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \quad (3)$$

If $\mathbf{R} = \mathbf{I}\sigma_e^2$, (23.3) simplifies to (23.4)

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_s & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}_s'\mathbf{X} & \mathbf{Z}_s'\mathbf{Z}_s + \mathbf{A}^{-1}\sigma_e^2/\sigma_s^2 & \mathbf{Z}_s'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z}_s & \mathbf{Z}'\mathbf{Z} + \sigma_e^2\mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{s}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}_s'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}. \quad (4)$$

We illustrate this model with the following data.

	n_{ij}				$y_{ij.}$			
	Herds				Herds			
Sires	1	2	3	4	1	2	3	4
1	3	5	0	0	25	34	–	–
2	0	8	4	0	–	74	31	–
3	4	2	6	8	23	11	43	73

The model assumed is

$$\begin{aligned} y_{ijk} &= s_i + h_j + e_{ijk}. \\ Var(\mathbf{s}) &= \mathbf{A}\sigma_e^2/12, \\ Var(\mathbf{e}) &= \mathbf{I}\sigma_e^2. \end{aligned}$$

$$\mathbf{A} = \begin{pmatrix} 1.0 & .5 & .5 \\ & 1.0 & .25 \\ & & 1.0 \end{pmatrix}.$$

h is fixed.

The ordinary LS equations are

$$\begin{pmatrix} 8 & 0 & 0 & 3 & 5 & 0 & 0 \\ & 12 & 0 & 0 & 8 & 4 & 0 \\ & & 20 & 4 & 2 & 6 & 8 \\ & & & 7 & 0 & 0 & 0 \\ & & & & 15 & 0 & 0 \\ & & & & & 10 & 0 \\ & & & & & & 8 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{s}} \\ \hat{\mathbf{h}} \end{pmatrix} = \begin{pmatrix} 59 \\ 105 \\ 150 \\ 48 \\ 119 \\ 74 \\ 73 \end{pmatrix}. \quad (5)$$

The mixed model equations are

$$\begin{pmatrix} 28 & -8 & -8 & 3 & 5 & 0 & 0 \\ & 28 & 0 & 0 & 8 & 4 & 0 \\ & & 36 & 4 & 2 & 6 & 8 \\ & & & 7 & 0 & 0 & 0 \\ & & & & 15 & 0 & 0 \\ & & & & & 10 & 0 \\ & & & & & & 8 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{s}} \\ \hat{\mathbf{h}} \end{pmatrix} = \begin{pmatrix} 59 \\ 105 \\ 150 \\ 48 \\ 119 \\ 74 \\ 73 \end{pmatrix}. \qquad (6)$$

The inverse of the matrix of (23.6) is

$$\begin{pmatrix} .0764 & .0436 & .0432 & -.0574 & -.0545 & -.0434 & -.0432 \\ .0436 & .0712 & .0320 & -.0370 & -.0568 & -.0477 & -.0320 \\ .0432 & .0320 & .0714 & -.0593 & -.0410 & -.0556 & -.0714 \\ -.0574 & -.0370 & -.0593 & .2014 & .0468 & .0504 & .0593 \\ -.0545 & -.0568 & -.0410 & .0468 & .1206 & .0473 & .0410 \\ -.0434 & -.0477 & -.0556 & .0504 & .0473 & .1524 & .0556 \\ -.0432 & -.0320 & -.0714 & .0593 & .0410 & .0556 & .1964 \end{pmatrix}. \qquad (7)$$

The solution is

$$\begin{aligned} \hat{\mathbf{s}}' &= (-.036661, .453353, -.435022), \\ \hat{\mathbf{h}}' &= (7.121439, 7.761769, 7.479672, 9.560022). \end{aligned}$$

Let us estimate σ_e^2 from the residual mean square using OLS reduction, and σ_e^2 by MIVQUE type computations. A solution to the OLS equations is

$$[10.14097,\ 11.51238,\ 9.12500,\ -2.70328,\ -2.80359,\ -2.67995,\ 0]$$

This gives a reduction in SS of 2514.166.

$$\mathbf{y}'\mathbf{y} = 2922.$$

Then $\hat{\sigma}_e^2$ = (2922 - 2514.166)/(40-6) = 11.995. MIVQUE requires computation of $\hat{\mathbf{s}}'\mathbf{A}^{-1}\hat{\mathbf{s}}$ and equating to its expectation.

$$\begin{aligned} \mathbf{A}^{-1} &= \frac{1}{3} \begin{pmatrix} 5 & -2 & -2 \\ -2 & 4 & 0 \\ -2 & 0 & 4 \end{pmatrix}. \\ \hat{\mathbf{s}}'\mathbf{A}^{-1}\hat{\mathbf{s}} &= .529500. \end{aligned}$$

Var (RHS of mixed model equations) = [Matrix (23.5)] σ_e^2 +

$$\begin{pmatrix} 8 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 20 \\ 3 & 0 & 4 \\ 5 & 8 & 2 \\ 0 & 4 & 6 \\ 0 & 0 & 8 \end{pmatrix} \mathbf{A} \begin{pmatrix} 8 & 0 & 0 & 3 & 5 & 0 & 0 \\ 0 & 12 & 0 & 0 & 8 & 4 & 0 \\ 0 & 0 & 20 & 4 & 2 & 6 & 8 \end{pmatrix} \sigma_s^2.$$

The second term of this is

$$\begin{pmatrix} 64 & 48 & 80 & 40 & 80 & 40 & 32 \\ & 144 & 60 & 30 & 132 & 66 & 24 \\ & & 400 & 110 & 130 & 140 & 160 \\ & & & 37 & 56 & 43 & 44 \\ & & & & 151 & 83 & 5 \\ & & & & & 64 & 56 \\ & & & & & & 64 \end{pmatrix} \sigma_s^2. \qquad (8)$$

$$\begin{aligned} Var(\hat{\mathbf{s}}) &= \mathbf{C}_s \text{ [matrix (23.8)] } \mathbf{C}_s' \, \sigma_s^2 \;+\; \mathbf{C}_s \text{ [matrix (23.5)] } \mathbf{C}_s' \, \sigma_e^2, \\ \text{where } \mathbf{C}_s &= \text{first 3 rows of (23.7).} \end{aligned}$$

$$\begin{aligned} Var(\hat{\mathbf{s}}) = & \begin{pmatrix} .005492 & -.001451 & -.001295 \\ & .007677 & -.006952 \\ & & .007599 \end{pmatrix} \sigma_e^2 \\ & + \begin{pmatrix} .017338 & -.005622 & -.003047 \\ & .053481 & -.050670 \\ & & .052193 \end{pmatrix} \sigma_s^2. \qquad (9) \end{aligned}$$

Then $E(\hat{\mathbf{s}}'\mathbf{A}^{-1}\hat{\mathbf{s}}) = tr(\mathbf{A}^{-1} \text{ [matrix (23.9)]}) = .033184 \, \sigma_e^2 + .181355 \, \sigma_s^2$. With these results we solve for $\hat{\sigma}_s^2$ and this is .7249 using estimated $\hat{\sigma}_e^2$ as 11.995. This is an approximate MIVQUE solution because $\hat{\sigma}_e^2$ was computed from the residual of ordinary least squares reduction rather than by MIVQUE.

Chapter 24
Animal Model, Repeated Records

C. R. Henderson

1984 - Guelph

In this chapter we deal with a one trait, repeated records model that has been extensively used in animal breeding, and particularly in lactation studies with dairy cattle. The assumptions of this model are not entirely realistic, but may be an adequate approximation. The scalar model is

$$y_{ij} \;=\; \mathbf{x}_{ij}^{'}\boldsymbol{\beta} \;+\; \mathbf{z}_{ij}^{'}\mathbf{u} \;+\; c_i \;+\; e_{ij}. \tag{1}$$

$\boldsymbol{\beta}$ represents fixed effects, and $\mathbf{x}_{ij}^{'}$ relates the j^{th} record of the i^{th} animal to elements of $\boldsymbol{\beta}$.

$\mathbf{u}$ represents other random effects, and $\mathbf{z}_{ij}^{'}$ relates the record to them.

c_i is a "cow" effect. It represents both genetic merit for production and permanent environmental effects.

e_{ij} is a random "error" associated with the individual record.

The vector representation is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{Z}_c\mathbf{c} + \mathbf{e}. \tag{2}$$

$$\begin{aligned}
Var(\mathbf{u}) &= \mathbf{G}, \\
Var(\mathbf{c}) &= \mathbf{I}\ \sigma_c^2 \text{ if cows are unrelated, with } \sigma_c^2 = \sigma_a^2 + \sigma_p^2 \\
&= \mathbf{A}\ \sigma_a^2 + \mathbf{I}\ \sigma_p^2 \text{ if cows are related,}
\end{aligned}$$

where σ_p^2 is the variance of permanent environmental effects, and if there are non-additive genetic effects, it also includes their variances. In that case $\mathbf{I}\ \sigma_p^2$ is only approximate.

$$Var(\mathbf{e}) = \mathbf{I}\ \sigma_e^2.$$

$Cov(\mathbf{u}, \mathbf{a}')$, $Cov(\mathbf{u}, \mathbf{e}')$, and $Cov(\mathbf{a}, \mathbf{e}')$ are all null. For the related cow model let

$$\mathbf{Z}_c\mathbf{c} = \mathbf{Z}_c\mathbf{a} + \mathbf{Z}_c\mathbf{p}. \tag{3}$$

It is advantageous to use this latter model in setting up the mixed model equations, for then the simple method for computing $\mathbf{A}^{-1}$ can be used. There appears to be no simple method for computing directly the inverse of $Var(\mathbf{c})$.

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{Z}_c & \mathbf{X}'\mathbf{Z}_c \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \sigma_e^2\mathbf{G}^{-1} & \mathbf{Z}'\mathbf{Z}_c & \mathbf{Z}'\mathbf{Z}_c \\ \mathbf{Z}_c'\mathbf{X} & \mathbf{Z}_c'\mathbf{Z} & \mathbf{Z}_c'\mathbf{Z}_c + \mathbf{A}^{-1}\frac{\sigma_e^2}{\sigma_a^2} & \mathbf{Z}_c'\mathbf{Z}_c \\ \mathbf{Z}_c'\mathbf{X} & \mathbf{Z}_c'\mathbf{Z} & \mathbf{Z}_c'\mathbf{Z}_c & \mathbf{Z}_c'\mathbf{Z}_c + \mathbf{I}\frac{\sigma_e^2}{\sigma_p^2} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{Z}_c'\mathbf{y} \\ \mathbf{Z}_c'\mathbf{y} \end{pmatrix} \quad (4)$$

These equations are easy to write provided $\mathbf{G}^{-1}$ is easy to compute, $\mathbf{G}$ being diagonal, e.g. as is usually the case. $\mathbf{A}^{-1}$ can be computed by the easy method. Further $\mathbf{Z}_c'\mathbf{Z}_c + \mathbf{I}\sigma_e^2/\sigma_p^2$ is diagonal, so $\hat{\mathbf{p}}$ can be "absorbed" easily. In fact, one would not need to write the $\hat{\mathbf{p}}$ equations. See Henderson (1975b). Also $\mathbf{Z}'\mathbf{Z}+\sigma_e^2\mathbf{G}^{-1}$ is sometimes diagonal and therefore $\hat{\mathbf{u}}$ can be absorbed easily. If predictions of breeding values are of primary interest, $\hat{\mathbf{a}}$ is what is wanted. If, in addition, predictions of real producing abilities are wanted, one needs $\hat{\mathbf{p}}$. Note that by subtracting the 4^{th} equation of (24.4) from the 3^{rd} we obtain

$$\mathbf{A}^{-1}\left(\sigma_e^2/\sigma_a^2\right)\ \hat{\mathbf{a}} - \mathbf{I}\left(\sigma_e^2/\sigma_p^2\right)\ \hat{\mathbf{p}} = \mathbf{0}.$$

Consequently

$$\hat{\mathbf{p}} = \left(\sigma_p^2/\sigma_a^2\right)\ \mathbf{A}^{-1}\hat{\mathbf{a}}, \quad (5)$$

and predictions of real producing abilities are

$$\left(\mathbf{I} + \left(\sigma_p^2/\sigma_a^2\right)\ \mathbf{A}^{-1}\right)\ \hat{\mathbf{a}}. \quad (6)$$

Note that under the model used in this chapter

$$Var(y_{ij}) = Var(y_{ik}),\ j \neq k.$$

$Cov(y_{ij}, y_{ik})$ is identical for all pairs of $j \neq k$. This is not necessarily a realistic model. If we wish a more general model, probably the most logical and easiest one to analyze is that which treats different lactations as separate traits, the methods for which are described in Chapter 26.

We illustrate the simple repeatability model with the following example. Four animals produced records as follows in treatments 1,2,3. The model is

$$y_{ij} = t_i + a_j + p_j + e_{ij}.$$

	Animals			
Treatment	1	2	3	4
1	5	3	-	4
2	6	5	7	-
3	8	-	9	-

The relationship matrix of the 4 animals is

$$\begin{pmatrix} 1 & .5 & .5 & .5 \\ & 1 & .25 & .125 \\ & & 1 & .5 \\ & & & 1 \end{pmatrix}.$$

$$\begin{aligned} Var(\mathbf{a}) &= .25\ \mathbf{A}\sigma_y^2, \\ Var(\mathbf{p}) &= .2\ \mathbf{I}\sigma_y^2, \\ \mathbf{I}\sigma_e^2 &= .55\ \mathbf{I}\sigma_y^2. \end{aligned}$$

These values correspond to $h^2 = .25$ and $r = .45$, where r denotes repeatability. The OLS equations are

$$\begin{pmatrix} 3 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ & 3 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ & & 2 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ & & & 3 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ & & & & 2 & 0 & 0 & 0 & 2 & 0 & 0 \\ & & & & & 2 & 0 & 0 & 0 & 2 & 0 \\ & & & & & & 1 & 0 & 0 & 0 & 1 \\ & & & & & & & 3 & 0 & 0 & 0 \\ & & & & & & & & 2 & 0 & 0 \\ & & & & & & & & & 2 & 0 \\ & & & & & & & & & & 1 \end{pmatrix} \begin{pmatrix} \mathbf{t} \\ \mathbf{a} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} 12 \\ 18 \\ 17 \\ 19 \\ 8 \\ 16 \\ 4 \\ 19 \\ 8 \\ 16 \\ 4 \end{pmatrix}. \tag{7}$$

Note that the last 4 equations are identical to equations 4-7. Thus **a** and **p** are confounded in a fixed model. Now we add 2.2 $\mathbf{A}^{-1}$ to the 4-7 diagonal block of coefficients and 2.75 **I** to the 8-11 diagonal block of coefficients. The resulting coefficient matrix is in (24.8). $2.2 = .55/.25$, and $2.75 = .55/.2$.

$$\begin{pmatrix} 3.0 & 0 & 0 & 1.0 & 1.0 & 0 & 1.0 & 1.0 & 1.0 & 0 & 1.0 \\ & 3.0 & 0 & 1.0 & 1.0 & 1.0 & 0 & 1.0 & 1.0 & 1.0 & 0 \\ & & 2.0 & 1.0 & 0 & 1.0 & 0 & 1.0 & 0 & 1.0 & 0 \\ & & & 7.2581 & -1.7032 & -.9935 & -1.4194 & 3.0 & 0 & 0 & 0 \\ & & & & 5.0280 & -.1892 & .5677 & 0 & 2.0 & 0 & 0 \\ & & & & & 5.3118 & -1.1355 & 0 & 0 & 2.0 & 0 \\ & & & & & & 4.4065 & 0 & 0 & 0 & 1.0 \\ & & & & & & & 5.75 & 0 & 0 & 0 \\ & & & & & & & & 4.75 & 0 & 0 \\ & & & & & & & & & 4.75 & 0 \\ & & & & & & & & & & 3.75 \end{pmatrix} \tag{8}$$

The inverse of (24.8) (times 1000) is

$$
\begin{pmatrix}
693 & 325 & 313 & -280 & -231 & -217 & -247 & -85 & -117 & -43 & -119 \\
 & 709 & 384 & -288 & -246 & -266 & -195 & -96 & -114 & -118 & -34 \\
 & & 943 & -306 & -205 & -303 & -215 & -126 & -60 & -152 & -26 \\
 & & & 414 & 227 & 236 & 225 & -64 & 24 & 26 & 15 \\
 & & & & 390 & 153 & 107 & 0 & -64 & 31 & 33 \\
 & & & & & 410 & 211 & 14 & 37 & -53 & 2 \\
 & & & & & & 406 & -3 & 48 & -2 & -42 \\
 & & & & & & & 261 & 38 & 41 & 24 \\
 & & & & & & & & 286 & 21 & 18 \\
 & & & & & & & & & 290 & 12 \\
 & & & & & & & & & & 310
\end{pmatrix} \quad (9)
$$

The solution is

$$
\begin{aligned}
\hat{\mathbf{t}}' &= (4.123 \quad 5.952 \quad 8.133), \\
\hat{\mathbf{a}}' &= (.065, \ -.263, \ .280, \ .113), \\
\hat{\mathbf{p}}' &= (.104, \ -.326, \ .285, \ -.063).
\end{aligned}
$$

We next estimate σ_e^2, σ_a^2, σ_p^2, by MIVQUE with the priors that were used in the above mixed model solution. The $\mathbf{Z}_c'\mathbf{W}$ submatrix for both $\mathbf{a}$ and $\mathbf{p}$ is

$$
\begin{pmatrix}
1 & 1 & 1 & 3 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1
\end{pmatrix}. \quad (10)
$$

The variance of the right hand sides of the mixed model equations contains $\mathbf{W}'\mathbf{Z}_c\mathbf{A}\mathbf{Z}_c'\mathbf{W} \ \sigma_a^2$, where $\mathbf{W} = (\mathbf{X} \ \mathbf{Z} \ \mathbf{Z}_c \ \mathbf{Z}_c)$. The matrix of coefficients of σ_a^2 is in (24.11). $Var(\mathbf{r})$ also contains $\mathbf{W}'\mathbf{Z}_c\mathbf{Z}_c'\mathbf{W} \ \sigma_p^2$ and this matrix is in (24.12). The coefficients of σ_e^2 are in (24.7).

$$
\begin{pmatrix}
5.25 & 4.88 & 3.25 & 6.0 & 3.25 & 2.5 & 1.65 & 6.0 & 3.25 & 2.5 & 1.63 \\
 & 5.5 & 3.75 & 6.0 & 3.5 & 3.5 & 1.13 & 6.0 & 3.5 & 3.5 & 1.13 \\
 & & 3.0 & 4.5 & 1.5 & 3.0 & 1.0 & 4.5 & 1.5 & 3.0 & 1.0 \\
 & & & 9.0 & 3.0 & 3.0 & 1.5 & 9.0 & 3.0 & 3.0 & 1.5 \\
 & & & & 4.0 & 1.0 & .25 & 3.0 & 4.0 & 1.0 & .25 \\
 & & & & & 4.0 & 1.0 & 3.0 & 1.0 & 4.0 & 1.0 \\
 & & & & & & 1.0 & 1.5 & .25 & 1.0 & 1.0 \\
 & & & & & & & 9.0 & 3.0 & 3.0 & 1.5 \\
 & & & & & & & & 4.0 & 1.0 & .25 \\
 & & & & & & & & & 4.0 & 1.0 \\
 & & & & & & & & & & 1.0
\end{pmatrix} \quad (11)
$$

$$\begin{pmatrix} 3 & 2 & 1 & 3 & 2 & 0 & 1 & 3 & 2 & 0 & 1 \\ & 3 & 2 & 3 & 2 & 2 & 0 & 3 & 2 & 2 & 0 \\ & & 2 & 3 & 0 & 2 & 0 & 3 & 0 & 2 & 0 \\ & & & 9 & 0 & 0 & 0 & 9 & 0 & 0 & 0 \\ & & & & 4 & 0 & 0 & 0 & 4 & 0 & 0 \\ & & & & & 4 & 0 & 0 & 0 & 4 & 0 \\ & & & & & & 1 & 0 & 0 & 0 & 1 \\ & & & & & & & 9 & 0 & 0 & 0 \\ & & & & & & & & 4 & 0 & 0 \\ & & & & & & & & & 4 & 0 \\ & & & & & & & & & & 1 \end{pmatrix} \quad (12)$$

Now $Var(\hat{\mathbf{a}})$ contains $\mathbf{C}_a(Var(\mathbf{r}))\mathbf{C}_a'\sigma_a^2$, where $\mathbf{C}_a$ is the matrix formed by rows 4-9 of the matrix in (24.9). Then $\mathbf{C}_a(Var(\mathbf{r}))\mathbf{C}_a^{'}$ is

$$\begin{pmatrix} .0168 & .0012 & -.0061 & .0012 \\ & .0423 & -.0266 & -.0323 \\ & & .0236 & .0160 \\ & & & .0274 \end{pmatrix} \sigma_a^2 \quad (13)$$

$$+ \begin{pmatrix} .0421 & -.0019 & -.0099 & .0050 \\ & .0460 & -.0298 & -.0342 \\ & & .0331 & .0136 \\ & & & .0310 \end{pmatrix} \sigma_p^2 \quad (14)$$

$$+ \begin{pmatrix} .0172 & .0001 & -.0022 & -.0004 \\ & .0289 & -.0161 & -.0234 \\ & & .0219 & .0042 \\ & & & .0252 \end{pmatrix} \sigma_e^2. \quad (15)$$

We need $\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}}' \;=\; .2067$. The expectation of this is

$$tr\mathbf{A}^{-1} \;[\text{matrix (24.13)} \;+\; \text{matrix (24.14)} \;+\; \text{matrix (24.15)}]$$
$$= .1336\ \sigma_e^2 \;+\; .1423\ \sigma_a^2 \;+\; .2216\ \sigma_p^2.$$

To find $Var(\hat{\mathbf{p}})$ we use $\mathbf{C}_p$, the last 6 rows of (24.9).

$$Var(\hat{\mathbf{p}}) \;=\; \begin{pmatrix} .0429 & -.0135 & -.0223 & -.0071 \\ & .0455 & -.0154 & -.0166 \\ & & .0337 & .0040 \\ & & & .0197 \end{pmatrix} \sigma_a^2 \quad (16)$$

$$+ \begin{pmatrix} .1078 & -.0423 & -.0466 & -.0189 \\ & .0625 & -.0106 & -.0096 \\ & & .0586 & -.0014 \\ & & & .0298 \end{pmatrix} \sigma_p^2 \quad (17)$$

$$
+ \begin{pmatrix} .0441 & -.0167 & -.0139 & -.0135 \\ & .0342 & -.0101 & -.0074 \\ & & .0374 & -.0133 \\ & & & .0341 \end{pmatrix} \sigma_e^2. \qquad (18)
$$

We need $\hat{\mathbf{p}}'\hat{\mathbf{p}} = .2024$ with expectation

$$
\begin{aligned} & tr[\text{matrix } (24.16) + \text{matrix } (24.17) + \text{matrix } (24.18)] \\ & \quad = .1498 \ \sigma_e^2 + .1419 \ \sigma_a^2 + .2588 \ \sigma_p^2. \end{aligned}
$$

We need $\hat{\mathbf{e}}'\hat{\mathbf{e}}$.

$$
\hat{\mathbf{e}} = [\mathbf{I} - \mathbf{WCW}']\mathbf{y},
$$

where $\mathbf{C}$ = matrix (24.9), and $\mathbf{I} - \mathbf{WCW}'$ is

$$
\begin{pmatrix} .4911 & -.2690 & -.2221 & -.1217 & .1183 & .0034 & -.0626 & .0626 \\ & .4548 & -.1858 & .1113 & -.1649 & .0536 & .0289 & -.0289 \\ & & .4079 & .0104 & .0466 & -.0570 & .0337 & -.0337 \\ & & & .5122 & -.2548 & -.2574 & -.1152 & .1152 \\ & & & & .4620 & -.2073 & -.0238 & .0238 \\ & & & & & .4647 & .1390 & -.1390 \\ & & & & & & .3729 & -.3729 \\ & & & & & & & .3729 \end{pmatrix} \qquad (19)
$$

Then

$$
\begin{aligned}
\hat{\mathbf{e}} &= [.7078, \ -.5341, \ -.1736, \ -.1205, \ -.3624, \ .4829, \ -.3017, \ .3017]. \\
\hat{\mathbf{e}}'\hat{\mathbf{e}} &= 1.3774. \\
Var(\hat{\mathbf{e}}) &= (\mathbf{I} - \mathbf{WCW}') \ Var(\mathbf{y}) \ (\mathbf{I} - \mathbf{WCW}'), \\
Var(\mathbf{y}) &= \mathbf{Z}_c\mathbf{A}\mathbf{Z}_c' \ \sigma_a^2 + \mathbf{Z}_c\mathbf{Z}_c' \ \sigma_p^2 \ + \mathbf{I} \ \sigma_e^2. \\
&= \begin{pmatrix} 1 & .5 & .5 & 1. & .5 & .5 & 1. & .5 \\ & 1. & .125 & .5 & 1. & .25 & .5 & .25 \\ & & 1. & .5 & .125 & .5 & .5 & .5 \\ & & & 1. & .5 & .5 & 1. & .5 \\ & & & & 1. & .25 & .5 & .25 \\ & & & & & 1. & .5 & 1. \\ & & & & & & 1. & .5 \\ & & & & & & & 1. \end{pmatrix} \sigma_a^2 \\
& \quad + \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 & 0 & 0 \\ & & & 1 & 0 & 0 & 1 & 0 \\ & & & & 1 & 0 & 0 & 0 \\ & & & & & 1 & 0 & 1 \\ & & & & & & 1 & 0 \\ & & & & & & & 1 \end{pmatrix} \sigma_p^2 + \mathbf{I}\sigma_e^2.
\end{aligned}
$$

Then the diagonals of $Var(\hat{\mathbf{e}})$ are

$$\begin{aligned} & (.0651,\ .1047,\ .1491,\ .0493,\ .0918,\ .0916,\ .0475,\ .0475)\ \sigma_a^2 \\ + & (.1705,\ .1358,\ .2257,\ .1167,\ .1498,\ .1462,\ .0940,\ .0940)\ \sigma_p^2 \\ + & [\text{diagonals of } (24.19)]\ \sigma_e^2. \end{aligned}$$

Then $E(\hat{\mathbf{e}}'\hat{\mathbf{e}})$ is the sum of these diagonals

$$= .6465\ \sigma_a^2 + 1.1327\ \sigma_p^2 + 3.5385\ \sigma_e^2.$$

Chapter 25
Sire Model, Repeated Records

C. R. Henderson

1984 - Guelph

This chapter is a combination of those of Chapters 23 and 24. That is, we are concerned with progeny testing of sires, but some progeny have more than one record. The scalar model is

$$y_{ijk} = \mathbf{x}'_{ijk}\boldsymbol{\beta} + \mathbf{z}'_{ijk}\mathbf{u} + s_i + p_{ij} + e_{ijk}.$$

$\mathbf{u}$ represents random factors other than $\mathbf{s}$ and $\mathbf{p}$. It is assumed that all dams are unrelated and all progeny are non-inbred. Under an additive genetic model the covariance between any record on one progeny and any record on another progeny of the same sire is $\sigma_s^2 = \frac{1}{4}\ h^2\ \sigma_y^2$ if sires are a random sample from the population. The covariance between any pair of records on the same progeny is $\sigma_s^2 + \sigma_p^2 = r\sigma_y^2$. If sires are unselected, $\sigma_p^2 = (r - \frac{1}{4}h^2)\sigma_y^2$, $\sigma_e^2 = (1-r)\sigma_y^2$, $\sigma_s^2 = \frac{1}{4}h^2\sigma_y^2$.

In vector notation the model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zu} + \mathbf{Z}_s\mathbf{s} + \mathbf{Z}_p\mathbf{p} + \mathbf{e}.$$
$$Var(\mathbf{s}) = \mathbf{A}\ \sigma_s^2,\ Var(\mathbf{p}) = \mathbf{I}\ \sigma_p^2,\ Var(\mathbf{e}) = \mathbf{I}\ \sigma_e^2.$$

With field data one might eliminate progeny that do not have a first record in order to reduce bias due to culling, which is usually more intense on first than on later records. Further, if a cow changes herds, the records only in the first herd might be used. In this case useful computing strategies can be employed. The data can be entered by herds, and $\mathbf{p}$ easily absorbed because $\mathbf{Z}_p'\mathbf{Z}_p + \mathbf{I}\sigma_e^2/\sigma_p^2$ is diagonal. Once this has been done, fixed effects pertinent to that particular herd can be absorbed. These methods are described in detail in Ufford *et al*. (1979). They are illustrated also in a simple example which follows.

We have a model in which the fixed effects are herd-years. The observations are displayed in the following table.

		Herd - years						
Sires	Progeny	11	12	13	21	22	23	24
1	1	5	6	4				
	2	5	8	-				
	3	-	9	4				
	4				5	6	7	3
	5				4	5	-	-
	6				-	4	3	-
	7				-	-	2	8
2	8	7	6	-				
	9	-	5	4				
	10	-	9	-				
	11	-	-	4				
	12				3	7	6	-
	13				-	5	6	8
	14				-	-	5	4

We assume $\sigma_e^2/\sigma_s^2 = 8.8$, $\sigma_e^2/\sigma_p^2 = 1.41935$. These correspond to unselected sires, $h^2 = .25$, r $= .45$. Further, we assume that **A** for the 2 sires is

$$\begin{pmatrix} 1 & .25 \\ .25 & 1 \end{pmatrix}.$$

Ordering the solution vector **hy**, **s**, **p** the matrix of coefficients of OLS equations is in (25.1), and the right hand side vector is (17, 43, 16, 12, 27, 29, 23, 88, 79, 15, 13, 13, 21, 9, 7, 10, 13, 9, 9, 4, 16, 19, 9)′.

$$\begin{aligned}
\mathbf{X}'\mathbf{X} &= \text{diag}\ (3, 6, 4, 3, 5, 6, 4) \\
\mathbf{Z}_s'\mathbf{Z}_s &= \text{diag}\ (17, 14) \\
\mathbf{Z}_p'\mathbf{Z}_p &= \text{diag}\ (3, 2, 2, 4, 2, 2, 2, 2, 2, 1, 1, 3, 3, 2) \\
\mathbf{Z}_s'\mathbf{X} &= \begin{pmatrix} 2 & 3 & 2 & 2 & 3 & 3 & 2 \\ 1 & 3 & 2 & 1 & 2 & 3 & 2 \end{pmatrix} \\
\mathbf{Z}_s'\mathbf{Z}_p &= \begin{pmatrix} 3 & 2 & 2 & 4 & 2 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 & 1 & 3 & 3 & 2 \end{pmatrix}
\end{aligned}$$

$$\mathbf{Z}_p'\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \tag{1}$$

Modifying these by adding 8.8 $\mathbf{A}^{-1}$ and 1.41935 $\mathbf{I}$ to appropriate submatrices of the coefficient matrix, the BLUP solution is

$$\begin{aligned} \widehat{\mathbf{hy}} &= [5.83397, 7.14937, 4.18706, 3.76589, 5.29825, 4.83644, 5.64274]', \\ \hat{\mathbf{s}} &= [-.06397, .06397]', \\ \hat{\mathbf{p}} &= [-.44769, .04229, .52394, .31601, .01866, -.87933, -.10272, \\ &\quad -.03255, -.72072, .73848, -.10376, .43162, .68577, -.47001]'. \end{aligned}$$

If one absorbs $\mathbf{p}$ in the mixed model equations, we obtain

$$\begin{pmatrix} 2.189 & -.811 & -.226 & 0 & 0 & 0 & 0 & .736 & .415 \\ & 4.191 & -.811 & 0 & 0 & 0 & 0 & 1.151 & 1.417 \\ & & 2.775 & 0 & 0 & 0 & 0 & .736 & 1.002 \\ & & & 2.297 & -.703 & -.411 & -.184 & .677 & .321 \\ & & & & 3.778 & -.930 & -.411 & 1.092 & .642 \\ & & & & & 4.486 & -.996 & 1.092 & 1.057 \\ & & & & & & 3.004 & .677 & .736 \\ & & & & & & & 15.549 & -2.347 \\ & & & & & & & & 14.978 \end{pmatrix}$$

$$\begin{pmatrix} \widehat{\mathbf{hy}} \\ \hat{\mathbf{s}} \end{pmatrix} = \begin{pmatrix} 6.002 \\ 21.848 \\ 4.518 \\ 1.872 \\ 10.526 \\ 9.602 \\ 9.269 \\ 31.902 \\ 31.735 \end{pmatrix}.$$

The solution for $\widehat{\mathbf{hy}}$ and $\hat{\mathbf{s}}$ are the same as before.

If one chooses, and this would be mandatory in large sets of data, $\widehat{\mathbf{hy}}$ can be absorbed herd by herd. Note that the coefficients of $\widehat{\mathbf{hy}}$ are in block diagonal form. When $\widehat{\mathbf{hy}}$ is absorbed, the equations obtained are

$$\begin{pmatrix} 12.26353 & -5.222353 \\ -5.22353 & 12.26353 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{s}}_1 \\ \hat{\mathbf{s}}_2 \end{pmatrix} = \begin{pmatrix} -1.1870 \\ 1.1870 \end{pmatrix}.$$

The solution is approximately the same for sires as before, the approximation being due to rounding errors.

Chapter 26
Animal Model, Multiple Traits

C. R. Henderson

1984 - Guelph

1 No Missing Data

In this chapter we deal with the same model as in Chapter 22 except now there are 2 or more traits. First we shall discuss the simple situation in which every trait is observed on every animal. There are n animals and t traits. Therefore the record vector has nt elements, which we denote by

$$\mathbf{y}' = [\mathbf{y}_1' \;\; \mathbf{y}_2' \ldots \mathbf{y}_t'].$$

$\mathbf{y}_1'$ is the vector of n records on trait 1, etc. Let the model be

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_t \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & 0 & \ldots & 0 \\ 0 & \mathbf{X}_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \mathbf{X}_t \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_t \end{pmatrix} + \begin{pmatrix} \mathbf{I} & 0 & \ldots & 0 \\ 0 & \mathbf{I} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_t \end{pmatrix} + \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_t \end{pmatrix}. \quad (1)$$

Accordingly the model for records on the first trait is

$$\mathbf{y}_1 = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{a}_1 + \mathbf{e}_1, \text{ etc.} \quad (2)$$

Every $\mathbf{X}_i$ has n rows and p_i columns, the latter corresponding to $\boldsymbol{\beta}_i$ with p_i elements. Every $\mathbf{I}$ has order n x n, and every $\mathbf{e}_i$ has n elements.

$$Var \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_t \end{pmatrix} = \begin{pmatrix} \mathbf{A}g_{11} & \mathbf{A}g_{12} & \ldots & \mathbf{A}g_{1t} \\ \mathbf{A}g_{12} & \mathbf{A}g_{22} & \ldots & \mathbf{A}g_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}g_{1t} & \mathbf{A}g_{2t} & \ldots & \mathbf{A}g_{tt} \end{pmatrix} = \mathbf{G}. \quad (3)$$

g_{ij} represents the elements of the additive genetic variance-covariance matrix in a noninbred population.

$$Var \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_t \end{pmatrix} = \begin{pmatrix} \mathbf{I}r_{11} & \mathbf{I}r_{12} & \ldots & \mathbf{I}r_{1t} \\ \mathbf{I}r_{12} & \mathbf{I}r_{22} & \ldots & \mathbf{I}r_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{I}r_{1t} & \mathbf{I}r_{2t} & \ldots & \mathbf{I}r_{tt} \end{pmatrix} = \mathbf{R}. \tag{4}$$

r_{ij} represents the elements of the environmental variance-covariance matrix. Then

$$\mathbf{G}^{-1} = \begin{pmatrix} \mathbf{A}^{-1}g^{11} & \ldots & \mathbf{A}^{-1}g^{1t} \\ \vdots & & \vdots \\ \mathbf{A}^{-1}g^{1t} & \ldots & \mathbf{A}^{-1}g^{tt} \end{pmatrix}. \tag{5}$$

g^{ij} are the elements of the inverse of the additive genetic variance covariance matrix.

$$\mathbf{R}^{-1} = \begin{pmatrix} \mathbf{I}r^{11} & \ldots & \mathbf{I}r^{1t} \\ \vdots & & \vdots \\ \mathbf{I}r^{1t} & \ldots & \mathbf{I}r^{tt} \end{pmatrix}. \tag{6}$$

r^{ij} are the elements of the inverse of the environmental variance-covariance matrix. Now the GLS equations regarding $\mathbf{a}$ fixed are

$$\begin{pmatrix} \mathbf{X'}_1\mathbf{X}_1r^{11} & \ldots & \mathbf{X'}_1\mathbf{X}_tr^{1t} & \mathbf{X'}_1r^{11} & \ldots & \mathbf{X'}_1r^{1t} \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{X'}_t\mathbf{X}_1r^{1t} & \ldots & \mathbf{X'}_t\mathbf{X}_tr^{tt} & \mathbf{X'}_tr^{1t} & \ldots & \mathbf{X'}_tr^{tt} \\ \mathbf{X}_1r^{11} & \ldots & \mathbf{X}_tr^{1t} & \mathbf{I}r^{11} & \ldots & \mathbf{I}r^{1t} \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{X}_1r^{1t} & \ldots & \mathbf{X}_tr^{tt} & \mathbf{I}r^{1t} & \ldots & \mathbf{I}r^{tt} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o_1 \\ \vdots \\ \boldsymbol{\beta}^o_t \\ \hat{\mathbf{a}}_1 \\ \vdots \\ \hat{\mathbf{a}}_t \end{pmatrix}$$
$$= \begin{pmatrix} \mathbf{X'}_1\mathbf{y}_1r^{11} & + & \ldots & + & \mathbf{X'}_1\mathbf{y}_tr^{1t} \\ \vdots & & & & \vdots \\ \mathbf{X'}_t\mathbf{y}_1r^{1t} & + & \ldots & + & \mathbf{X'}_t\mathbf{y}_tr^{tt} \\ \mathbf{y}_1r^{11} & + & \ldots & + & \mathbf{y}_tr^{1t} \\ \vdots & & & & \vdots \\ \mathbf{y}_1r^{1t} & + & \ldots & + & \mathbf{y}_tr^{tt} \end{pmatrix}. \tag{7}$$

The mixed model equations are formed by adding (26.5) to the lower t^2 blocks of (26.7).

If we wish to estimate the g_{ij} and r_{ij} by MIVQUE we take prior values of g_{ij} and r_{ij} for the mixed model equations and solve. We find that quadratics in $\hat{\mathbf{a}}$ and $\hat{\mathbf{e}}$ needed for MIVQUE are

$$\hat{\mathbf{a}}'_i\mathbf{A}^{-1}\hat{\mathbf{a}}_j \text{ for } i = l, \ \ldots, \ t; \ j = i, \ \ldots, \ t. \tag{8}$$

$$\hat{\mathbf{e}}'_i\hat{\mathbf{e}}_j \text{ for } i = l, \ \ldots, \ t; \ j = i, \ \ldots, \ t. \tag{9}$$

To obtain the expectations of (26.8) we first compute the variance-covariance matrix of the right hand sides of (26.7). This will consist of $t(t-1)/2$ matrices each of the same order as the matrix of (26.7) multiplied by an element of g_{ij}. It will also consist of the same number of matrices with the same order multiplied by an element of r_{ij}. The matrix for g_{kk} is

$$\begin{pmatrix} \mathbf{X'}_1\mathbf{A}\mathbf{X}_1 r^{1k}r^{1k} & \ldots & \mathbf{X'}_1\mathbf{A}\mathbf{X}_t r^{1k}r^{tk} & \mathbf{X'}_1\mathbf{A} r^{1k}r^{1k} & \ldots & \mathbf{X'}_1\mathbf{A} r^{1k}r^{tk} \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{X'}_t\mathbf{A}\mathbf{X}_1 r^{tk}r^{1k} & \ldots & \mathbf{X'}_t\mathbf{A}\mathbf{X}_t r^{tk}r^{tk} & \mathbf{X'}_t\mathbf{A} r^{tk}r^{1k} & \ldots & \mathbf{X'}_t\mathbf{A} r^{tk}r^{tk} \\ \mathbf{A}\mathbf{X}_1 r^{1k}r^{1k} & \ldots & \mathbf{A}\mathbf{X}_t r^{1k}r^{tk} & \mathbf{A} r^{1k}r^{1k} & \ldots & \mathbf{A} r^{1k}r^{tk} \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{A}\mathbf{X}_1 r^{tk}r^{1k} & \ldots & \mathbf{A}\mathbf{X}_t r^{tk}r^{tk} & \mathbf{A} r^{tk}r^{1k} & \ldots & \mathbf{A} r^{tk}r^{tk} \end{pmatrix} \tag{10}$$

The ij^{th} sub-block of the upper left set of $t \times t$ blocks is $\mathbf{X'}_i\mathbf{A}\mathbf{X}_j r^{ik}r^{jk}$. The sub-block of the upper right set of $t \times t$ blocks is $\mathbf{X'}_i\mathbf{A} r^{ik}r^{jk}$. The sub-block of the lower right set of $t \times t$ blocks is $\mathbf{A} r^{ik}r^{jk}$.

The matrix for g_{km} is

$$\begin{pmatrix} \mathbf{P} & \mathbf{T} \\ \mathbf{T'} & \mathbf{S} \end{pmatrix}, \tag{11}$$

where

$$\mathbf{P} = \begin{pmatrix} 2\mathbf{X'}_1\mathbf{A}\mathbf{X}_1 r^{1k}r^{1m} & \ldots & \mathbf{X'}_1\mathbf{A}\mathbf{X}_t(r^{1k}r^{tm} + r^{1m}r^{tk}) \\ \vdots & & \vdots \\ \mathbf{X'}_t\mathbf{A}\mathbf{X}_1(r^{1k}r^{tm} + r^{1m}r^{tk}) & \ldots & 2\mathbf{X'}_t\mathbf{A}\mathbf{X}_t r^{tk}r^{tm} \end{pmatrix},$$

$$\mathbf{T} = \begin{pmatrix} 2\mathbf{X'}_1\mathbf{A} r^{1k}r^{1m} & \ldots & \mathbf{X'}_1\mathbf{A}(r^{1k}r^{tm} + r^{1m}r^{tk}) \\ \vdots & & \vdots \\ \mathbf{X'}_t\mathbf{A}(r^{1k}r^{tm} + r^{1m}r^{tk}) & \ldots & 2\mathbf{X'}_t\mathbf{A} r^{tk}r^{tm} \end{pmatrix},$$

and

$$\mathbf{S} = \begin{pmatrix} 2\mathbf{A} r^{1k}r^{1m} & \ldots & \mathbf{A}(r^{1k}r^{tm} + r^{1m}r^{tk}) \\ \vdots & & \vdots \\ \mathbf{A}(r^{1k}r^{tm} + r^{1m}r^{tk}) & \ldots & 2\mathbf{A} r^{tk}r^{tm} \end{pmatrix}.$$

The ij^{th} sub-block of the upper left set of $t \times t$ blocks is

$$\mathbf{X'}_i\mathbf{A}\mathbf{X}_j(r^{ik}r^{jm} + r^{im}r^{jk}). \tag{12}$$

The ij^{th} sub-block of the upper right set is

$$\mathbf{X'}_i\mathbf{A}(r^{ik}r^{jm} + r^{im}r^{jk}). \tag{13}$$

The ij^{th} sub-block of the lower right set is

$$\mathbf{A}(r^{ik}r^{jm} + r^{im}r^{jk}). \tag{14}$$

The matrix for r_{kk} is the same as (26.10) except that $\mathbf{I}$ replaces $\mathbf{A}$. Thus the 3 types of sub-blocks are $\mathbf{X'}_i\mathbf{X}_j r^{ik}r^{jk}$, $\mathbf{X'}_i r^{ik}r^{jk}$, and $\mathbf{I}r^{ik}r^{jk}$. The matrix for r_{km} is the same as (26.11) except that $\mathbf{I}$ replaces $\mathbf{A}$. Thus the 3 types of blocks are $\mathbf{X'}_i\mathbf{X}_j(r^{ik}r^{jm} + r^{im}r^{jk})$, $\mathbf{X'}_i(r^{ik}r^{jm} + r^{im}r^{jk})$, and $\mathbf{I}(r^{ik}r^{jm} + r^{im}r^{jk})$.

Now define the $p+1, \ \ldots, \ p+n$ rows of a g-inverse of mixed model coefficient matrix as $\mathbf{C}_1$, the next n rows as $\mathbf{C}_2$, etc., with the last n rows being $\mathbf{C}_t$. Then

$$Var(\hat{\mathbf{a}}_i) = \mathbf{C}_i[Var(\mathbf{r})]\mathbf{C}'_i, \tag{15}$$

where $Var(\mathbf{r})$ = variance of right hand sides expressed as matrices multiplied by the g_{ij} and r_{ij} as described above.

$$Cov(\hat{\mathbf{a}}_i, \hat{\mathbf{a}}'_j) = \mathbf{C}_i[Var(\mathbf{r})]\mathbf{C}'_j. \tag{16}$$

Then

$$E(\hat{\mathbf{a}}_i\mathbf{A}^{-1}\hat{\mathbf{a}}'_i) = tr\mathbf{A}^{-1}\ Var(\hat{\mathbf{a}}_i). \tag{17}$$

$$E(\hat{\mathbf{a}}_i\mathbf{A}^{-1}\hat{\mathbf{a}}'_j) = tr\mathbf{A}^{-1}\ Cov(\hat{\mathbf{a}}_i, \hat{\mathbf{a}}'_j). \tag{18}$$

To find the quadratics of (26.9) and their expectations we first compute

$$\mathbf{I} - \mathbf{WCW'}\tilde{\mathbf{R}}^{-1}, \tag{19}$$

where $\mathbf{W} = (\mathbf{X}\ \ \mathbf{Z})$ and $\mathbf{C}$ = g-inverse of mixed model coefficient matrix. Then

$$\hat{\mathbf{e}} = (\mathbf{I} - \mathbf{WCW'}\tilde{\mathbf{R}}^{-1})\mathbf{y}. \tag{20}$$

Let the first n rows of (26.19) be denoted $\mathbf{B}_1$, the next n rows $\mathbf{B}_2$, etc. Also let

$$\mathbf{B}_i \equiv (\mathbf{B}_{i1}\ \ \mathbf{B}_{i2}\ \ \ldots\ \ \mathbf{B}_{it}). \tag{21}$$

Each $\mathbf{B}_{ij}$ has dimension $n \times n$ and is symmetric. Also $\mathbf{I} - \mathbf{WCW'}\tilde{\mathbf{R}}^{-1}$ is symmetric and as a consequence $\mathbf{B}_{ij} = \mathbf{B}_{ji}$. Use can be made of these facts to reduce computing labor. Now

$$\begin{aligned} \hat{\mathbf{e}}_i &= \mathbf{B}_i\mathbf{y}\ (i = 1,\ \ldots,\ t). & (22)\\ Var(\hat{\mathbf{e}}_i) &= \mathbf{B}_i[Var(\mathbf{y})]\mathbf{B}'_i. & (23)\\ Cov(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) &= \mathbf{B}_i[Var(\mathbf{y})]\mathbf{B}'_j. & (24) \end{aligned}$$

By virtue of the form of $Var(\mathbf{y})$,

$$\begin{aligned} Var(\hat{\mathbf{e}}_i) = & \sum_{k=1}^{t}\mathbf{B}_{ik}^2 r_{kk} + \sum_{k=1}^{t-1}\sum_{m=k+1}^{t} 2\mathbf{B}_{ik}\mathbf{B}_{im}r_{km} \\ & + \sum_{k=1}^{t}\mathbf{B}_{ik}\mathbf{A}\mathbf{B}_{ik}g_{kk} + \sum_{k=1}^{t-1}\sum_{m=k+1}^{t} 2\mathbf{B}_{ik}\mathbf{A}\mathbf{B}_{im}g_{km}. \end{aligned} \tag{25}$$

$$
\begin{aligned}
Cov(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) &= \sum_{k=1}^{t} \mathbf{B}_{ik}\mathbf{B}_{jk} r_{kk} \\
&+ \sum_{k=1}^{t-1} \sum_{m=k+1}^{t} (\mathbf{B}_{ik}\mathbf{B}_{jm}\mathbf{B}_{im}\mathbf{B}_{jk}) r_{km} \\
&+ \sum_{k=1}^{t} \mathbf{B}_{ik}\mathbf{A}\mathbf{B}_{jk} g_{kk} \\
&+ \sum_{k=1}^{t-1} \sum_{m=k+1}^{t} (\mathbf{B}_{ik}\mathbf{A}\mathbf{B}_{jm} + \mathbf{B}_{im}\mathbf{A}\mathbf{B}_{jk}) g_{km}. \qquad (26) \\
E(\hat{\mathbf{e}}_i'\hat{\mathbf{e}}_i) &= trVar(\hat{\mathbf{e}}_i). \qquad (27) \\
E(\hat{\mathbf{e}}_i'\hat{\mathbf{e}}_j) &= trCov(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j'). \qquad (28)
\end{aligned}
$$

Note that only the diagonals of the matrices of (26.25) and (26.26) are needed.

2 Missing Data

When data are missing on some traits of some of the animals, the computations are more difficult. An attempt is made in this section to present algorithms that are efficient for computing, including strategies for minimizing data storage requirements. Henderson and Quaas (1976) discuss BLUP techniques for this situation.

The computations for the missing data problem are more easily described and carried out if we order the records, traits within animals. It also is convenient to include missing data as a dummy value = 0. Then $\mathbf{y}$ has nt elements as follows:

$$\mathbf{y}' = (\mathbf{y}_1'\mathbf{y}_2' \ \ldots \ \mathbf{y}_n'),$$

where $\mathbf{y}_i$ is the vector of records on the t traits for the i^{th} animal. With no missing data the model for the nt records is

$$\begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1t} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2t} \\ \vdots \\ y_{n1} \\ y_{n2} \\ \vdots \\ y_{nt} \end{pmatrix} = \begin{pmatrix} \mathbf{x}'_{11} & 0 & \ldots & 0 \\ 0 & \mathbf{x}'_{12} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \mathbf{x}'_{1t} \\ \mathbf{x}'_{21} & 0 & \ldots & 0 \\ 0 & \mathbf{x}'_{22} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \mathbf{x}'_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{x}'_{n1} & 0 & \ldots & 0 \\ 0 & \mathbf{x}'_{n2} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \mathbf{x}'_{nt} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_t \end{pmatrix} + \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1t} \\ a_{21} \\ a_{22} \\ \vdots \\ a_{2t} \\ \vdots \\ a_{n1} \\ a_{n2} \\ \vdots \\ a_{nt} \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{12} \\ \vdots \\ e_{1t} \\ e_{21} \\ e_{22} \\ \vdots \\ e_{2t} \\ \vdots \\ e_{n1} \\ e_{n2} \\ \vdots \\ e_{nt} \end{pmatrix}.$$

$\mathbf{x}'_{ij}$ is a row vector relating the record on the j^{th} trait of the i^{th} animal to $\boldsymbol{\beta}_j$, the fixed effects for the j^{th} trait. $\boldsymbol{\beta}_j$ has p_j elements and $\sum_j p_j = p$. When a record is missing, it is set to 0 and so are the elements of the model for that record. Thus, whether data are missing or not, the incidence matrix has dimension, nt by $(p + nt)$. Now $\mathbf{R}$ has block diagonal form as follows.

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & 0 & \ldots & 0 \\ 0 & \mathbf{R}_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \mathbf{R}_n \end{pmatrix}. \tag{29}$$

For an animal with no missing data, $\mathbf{R}_i$ is the $t \times t$ environmental covariance matrix. For an animal with missing data the rows (and columns) of $\mathbf{R}_i$ pertaining to missing data are set to zero. Then in place of $\mathbf{R}^{-1}$ ordinarily used in the mixed model equations, we use $\mathbf{R}^-$ which is

$$\begin{pmatrix} \mathbf{R}_1^- & & & \mathbf{0} \\ & \mathbf{R}_2^- & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{R}_n^- \end{pmatrix}. \tag{30}$$

$\mathbf{R}_1^-$ is the zeroed type of g-inverse described in Section 3.3. It should be noted that $\mathbf{R}_i$ is the same for every animal that has the same missing data. There are at most $t^2 - 1$ such unique matrices, and in the case of sequential culling only t such matrices corresponding to trait 1 only, traits 1 and 2 only, $\ldots$, all traits. Thus we do not need to store $\mathbf{R}$ and $\mathbf{R}^-$ but only the unique types of $\mathbf{R}_i^-$.

$Var(\mathbf{a})$ has a simple form, which is

$$Var(\mathbf{a}) = \begin{pmatrix} a_{11}\mathbf{G}_0 & a_{12}\mathbf{G}_0 & \ldots & a_{1n}\mathbf{G}_0 \\ a_{12}\mathbf{G}_0 & a_{22}\mathbf{G}_0 & \ldots & a_{2n}\mathbf{G}_0 \\ \vdots & \vdots & & \vdots \\ a_{1n}\mathbf{G}_0 & a_{2n}\mathbf{G}_0 & \ldots & a_{nn}\mathbf{G}_0 \end{pmatrix}, \tag{31}$$

where $\mathbf{G}_0$ is the $t \times t$ covariance matrix of additive effects in an unselected non-inbred population. Then

$$[Var(\mathbf{a})]^{-1} = \begin{pmatrix} a^{11}\mathbf{G}_0^{-1} & a^{12}\mathbf{G}_0^{-1} & \ldots & a^{1n}\mathbf{G}_0^{-1} \\ a^{12}\mathbf{G}_0^{-1} & a^{22}\mathbf{G}_0^{-1} & \ldots & a^{2n}\mathbf{G}_0^{-1} \\ \vdots & \vdots & & \vdots \\ a^{1n}\mathbf{G}_0^{-1} & a^{2n}\mathbf{G}_0^{-1} & \ldots & a^{nn}\mathbf{G}_0^{-1} \end{pmatrix}. \tag{32}$$

a^{ij} are the elements of the inverse of $\mathbf{A}$. Note that all nt of the a_{ij} are included in the mixed model equations even though there are missing data.

We illustrate prediction by the following example that includes 4 animals and 3 traits with the $\boldsymbol{\beta}_j$ vector having 2, 1, 2 elements respectively.

	Trait		
Animal	1	2	3
1	5	3	6
2	2	5	7
3	-	3	4
4	2	-	-

$$\mathbf{X} \text{ for } \boldsymbol{\beta}_1 = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix},$$

$$\text{for } \boldsymbol{\beta}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$\text{and for } \boldsymbol{\beta}_3 = \begin{pmatrix} 1 & 3 \\ 1 & 4 \\ 1 & 2 \end{pmatrix}.$$

Then, with missing records included, the incidence matrix is

$$\begin{pmatrix}
1 & 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 4 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}. \tag{33}$$

We assume that the environmental covariance matrix is

$$\begin{pmatrix} 5 & 3 & 1 \\ & 6 & 4 \\ & & 7 \end{pmatrix}.$$

Then $\mathbf{R}^-$ for animals 1 and 2 is

$$\begin{pmatrix} .3059 & -.2000 & .0706 \\ & .4000 & -.2000 \\ & & .2471 \end{pmatrix},$$

$\mathbf{R}^-$ for animal 3 is

$$\begin{pmatrix} 0 & 0 & 0 \\ & .2692 & -.1538 \\ & & .2308 \end{pmatrix},$$

and for animal 4 is

$$\begin{pmatrix} .2 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{pmatrix}.$$

Suppose that

$$\mathbf{A} = \begin{pmatrix} 1. & 0 & .5 & 0 \\ & 1. & .5 & .5 \\ & & 1. & .25 \\ & & & 1. \end{pmatrix}$$

and

$$\mathbf{G}_0 = \begin{pmatrix} 2 & 1 & 1 \\ & 3 & 2 \\ & & 4 \end{pmatrix}.$$

$Var(\hat{\mathbf{a}})$ is

$$\begin{pmatrix}
2.0 & 1.0 & 1.0 & 0 & 0 & 0 & 1.0 & .5 & .5 & 0 & 0 & 0 \\
 & 3.0 & 2.0 & 0 & 0 & 0 & .5 & 1.5 & 1.0 & 0 & 0 & 0 \\
 & & 4.0 & 0 & 0 & 0 & .5 & 1.0 & 2.0 & 0 & 0 & 0 \\
 & & & 2.0 & 1.0 & 1.0 & 1.0 & .5 & .5 & 1.0 & .5 & .5 \\
 & & & & 3.0 & 2.0 & .5 & 1.5 & 1.0 & .5 & 1.5 & 1.0 \\
 & & & & & 4.0 & .5 & 1.0 & 2.0 & .5 & 1.0 & 2.0 \\
 & & & & & & 2.0 & 1.0 & 1.0 & .5 & .25 & .25 \\
 & & & & & & & 3.0 & 2.0 & .25 & .75 & .5 \\
 & & & & & & & & 4.0 & .25 & .5 & 1.0 \\
 & & & & & & & & & 2.0 & 1.0 & 1.0 \\
 & & & & & & & & & & 3.0 & 2.0 \\
 & & & & & & & & & & & 4.0
\end{pmatrix}. \quad (34)$$

Using the incidence matrix, $\mathbf{R}^-$, $\mathbf{G}^{-1}$, and $\mathbf{y}$ we get the coefficient matrix of mixed model equations in (26.35) $\ldots$ (26.37). The right hand side vector is (1.8588, 4.6235, -.6077, 2.5674, 8.1113, 1.3529, -1.0000, 1.2353, .1059, .2000, .8706, 0, .1923, .4615, .4000, 0, 0)$'$. The solution vector is (8.2451, -1.7723, 3.9145, 3.4054, .8066, .1301, -.4723, .0154, -.2817, .3965, -.0911, -.1459, -.2132, -.2480, .0865, .3119, .0681)'.

Upper left 8 x 8 (times 1000)

$$\begin{pmatrix}
812 & 2329 & -400 & 141 & 494 & 306 & -200 & 71 \\
 & 7176 & -1000 & 353 & 1271 & 612 & -400 & 141 \\
 & & 1069 & -554 & -1708 & -200 & 400 & -200 \\
 & & & 725 & 2191 & 71 & -200 & 247 \\
 & & & & 7100 & 212 & -600 & 741 \\
 & & & & & 1229 & -431 & -045 \\
 & & & & & & 1208 & -546 \\
 & & & & & & & 824
\end{pmatrix}. \quad (35)$$

Upper right 8 x 9 and (lower left 9 $\times$ 8)' (times 1000)

$$\begin{pmatrix}
306 & -200 & 71 & 0 & 0 & 0 & 200 & 0 & 0 \\
918 & -600 & 212 & 0 & 0 & 0 & 800 & 0 & 0 \\
-200 & 400 & -200 & 0 & 269 & -154 & 0 & 0 & 0 \\
71 & -200 & 247 & 0 & -154 & 231 & 0 & 0 & 0 \\
282 & -800 & 988 & 0 & -308 & 462 & 0 & 0 & 0 \\
308 & -77 & -38 & -615 & 154 & 77 & 0 & 0 & 0 \\
-77 & 269 & -115 & 154 & -538 & 231 & 0 & 0 & 0 \\
-38 & -115 & 192 & 77 & 231 & -385 & 0 & 0 & 0
\end{pmatrix}. \quad (36)$$

Lower right 9 × 9 (times 1000)

$$\begin{pmatrix} 1434 & -482 & -70 & -615 & 154 & 77 & -410 & 103 & 51 \\ & 1387 & -623 & 154 & -538 & 231 & 103 & -359 & 154 \\ & & 952 & 77 & 231 & -385 & 51 & 154 & -256 \\ & & & 1231 & -308 & -154 & 0 & 0 & 0 \\ & & & & 1346 & -615 & 0 & 0 & 0 \\ & & & & & 1000 & 0 & 0 & 0 \\ & & & & & & 1020 & -205 & -103 \\ & & & & & & & 718 & -308 \\ & & & & & & & & 513 \end{pmatrix}. \tag{37}$$

3 EM Algorithm

In spite of its possible slow convergence I tend to favor the EM algorithm for REML to estimate variances and covariances. The reason for this preference is its simplicity as compared to iterated MIVQUE and, above all, because the solution remains in the parameter space at each round of iteration.

If the data were stored animals in traits and $\hat{\mathbf{a}}_i$ = BLUP for the n breeding values on the i^{th} trait, g_{ij} would be estimated by iterating on

$$\hat{g}_{ij} = (\hat{\mathbf{a}}_i \mathbf{A}^{-1} \hat{\mathbf{a}}_j + tr\mathbf{A}^{-1}\mathbf{C}_{ij})/n, \tag{38}$$

where $\mathbf{C}_{ij}$ is the submatrix pertaining to $Cov(\hat{a}_i - a_i,\ \hat{a}'_j - a'_j)$ in a g-inverse of the mixed model coefficient matrix. These same computations can be effected from the solution with ordering traits in animals. The following FORTRAN routine accomplishes this.

```
      REAL *8  A( ), C( ), U( ), S
      INTEGER T
      .
      .
      .
      NT=N*T
      DO 7 I=1, T
      DO 7 J=I, T
      S=0.  DO
      DO 6 K=1, N
      DO 6 L=1, N
6     S=S+A(IHMSSF(K,L,N))*U(T*K-T+I)*U(T*L-T+J)
7     Store S
      .
      .
      .
      DO 9 I=1, T
      DO 9 J=I, T
      S=0. DO
      DO 8 K=1, N
      DO 8 L=1, N
8     S=S+A(IHMSSF(K,L,N))*C(IHMSSF(T*K-T+I,T*L-T+J,NT))
9     Store S
```

A is a one dimensional array with N(N+1)/2 elements containing $\mathbf{A}^{-1}$. C is a one dimensional array with NT(NT+1)/2 elements containing the lower $(\mathrm{NT})^2$ submatrix of a g-inverse of the coefficient matrix. This also is half-stored. U is the solution vector for $\mathbf{a}'$. IHMSSF is a half-stored matrix subscripting function. The $t(t+1)/2$ values of S in statement 7 are the values of $\hat{\mathbf{a}}_i'\mathbf{A}^{-1}\hat{\mathbf{a}}_i$. The values of S in statement 9 are the values of $tr\mathbf{A}^{-1}\mathbf{C}_{ij}$.

In our example these are the following for the first round. (.1750, -.1141, .0916, .4589, .0475, .1141), for $\hat{\mathbf{a}}_i'\mathbf{A}^{-1}\mathbf{a}_j$, and (7.4733, 3.5484, 3.4788, 10.5101, 7.0264, 14.4243) for $tr\mathbf{A}^{-1}\mathbf{C}_{ij}$. This gives us as the first estimate of $\mathbf{G}_0$ the following,

$$\begin{pmatrix} 1.912 & .859 & .893 \\ & 2.742 & 1.768 \\ & & 3.635 \end{pmatrix}.$$

Note that the matrix of the quadratics in $\hat{\mathbf{a}}$ remain the same for all rounds of iteration, that is, $\mathbf{A}^{-1}$.

In constrast, the quadratics in $\hat{\mathbf{e}}$ change with each round of iteration. However, they have a simple form since they are all of the type,

$$\sum_{i=1}^{n} \hat{\mathbf{e}}_i' \mathbf{Q}_i \hat{\mathbf{e}}_i,$$

where $\hat{\mathbf{e}}_i$ is the vector of BLUP of errors for the t traits in the i^{th} animal. The $\hat{\mathbf{e}}$ are computed as follows

$$\hat{e}_{ij} = y_{ij} - \mathbf{x}_{ij}' \boldsymbol{\beta}_j^o - \hat{a}_{ij} \tag{39}$$

when y_{ij} is observed. $\hat{e}_{ij}$ is set to 0 for $y_{ij} = 0$. This is not BLUP, but suffices for subsequent computations. At each round we iterate on

$$tr\mathbf{Q}_{ij}\mathbf{R} = (\hat{\mathbf{e}}'\mathbf{Q}_{ij}\hat{\mathbf{e}} + tr\mathbf{Q}_{ij}\mathbf{WCW}') \; i = 1, \; \ldots, \; t_{ij}, \; j = i, \; \ldots, \; t. \tag{40}$$

This gives at each round a set of equations of the form

$$\mathbf{T}\hat{\mathbf{r}} = \mathbf{q}, \tag{41}$$

where $\mathbf{T}$ is a symmetric $t \times t$ matrix, $\mathbf{r} = (r_{11} \; r_{12} \; \ldots \; r_{tt})'$, and $\mathbf{q}$ is a $t \times 1$ vector of numbers. Advantage can be taken of the symmetry of $\mathbf{T}$, so that only $t(t+1)/2$ coefficients need be computed rather than t^2.

Advantage can be taken of the block diagonal form of all $\mathbf{Q}_{ij}$. Each of them has the following form

$$\mathbf{Q}_{ij} = \begin{pmatrix} \mathbf{B}_{1ij} & & & \mathbf{0} \\ & \mathbf{B}_{2ij} & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{B}_{nij} \end{pmatrix}. \tag{42}$$

There are at most $t^2 - 1$ unique $\mathbf{B}_{kij}$ for any $\mathbf{Q}_{ij}$, these corresponding to the same number of unique $\mathbf{R}_k^-$. The $\mathbf{B}$ can be computed easily as follows. Let

$$\mathbf{R}_k^- = \begin{pmatrix} f_{11} & f_{12} & \ldots & f_{1t} \\ f_{12} & f_{22} & \ldots & f_{2t} \\ \vdots & \vdots & & \vdots \\ f_{1t} & f_{2t} & \ldots & f_{tt} \end{pmatrix} \equiv (\mathbf{f}_1 \; \mathbf{f}_2 \; \ldots \; \mathbf{f}_t).$$

Then

$$\mathbf{B}_{kii} = \mathbf{f}_i \mathbf{f}_i' \tag{43}$$

$$\mathbf{B}_{kij} = (\mathbf{f}_i \mathbf{f}_j') + (\mathbf{f}_i \mathbf{f}_j')' \text{ for } i \neq j. \tag{44}$$

In computing $tr\mathbf{Q}_{ij}\mathbf{R}$ remember that $\mathbf{Q}$ and $\mathbf{R}$ have the same block diagonal form. This computation is very easy for each of the n products. Let

$$\mathbf{B}_{kij} = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1t} \\ b_{12} & b_{22} & \ldots & b_{2t} \\ \vdots & \vdots & & \vdots \\ b_{1t} & b_{2t} & \ldots & b_{tt} \end{pmatrix}.$$

Then the coefficient of r_{ii} contributed by the k^{th} animal in the trace is b_{ii}. The coefficient of r_{ij} is $2b_{ij}$.

Finally note that we need only the n blocks of order $t \times t$ down the diagonals of $\mathbf{WCW}'$ for $tr\mathbf{Q}_{ij}\mathbf{WCW}'$. Partition $\mathbf{C}$ as

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{xx} & \mathbf{C}_{x1} & \ldots & \mathbf{C}_{xn} \\ \mathbf{C}'_{x1} & \mathbf{C}_{11} & \ldots & \mathbf{C}_{1n} \\ \vdots & \vdots & & \vdots \\ \mathbf{C}'_{xn} & \mathbf{C}'_{1n} & \ldots & \mathbf{C}_{nn} \end{pmatrix}.$$

Then the block of $\mathbf{WCW}'$ for the i^{th} animal is

$$\mathbf{X}_i\mathbf{C}_{xx}\mathbf{X}'_i + \mathbf{X}_i\mathbf{C}_{xi} + (\mathbf{X}_i\mathbf{C}_{xi})' + \mathbf{C}_{ii} \tag{45}$$

and then zeroed for missing rows and columns, although this is not really necessary since the $\mathbf{Q}_{kij}$ are correspondingly zeroed. $\mathbf{X}_i$ is the submatrix of $\mathbf{X}$ pertaining to the i^{th} animal. This submatrix has order $t \times p$.

We illustrate some of these computations for $\hat{\mathbf{r}}$. First, consider computation of $\mathbf{Q}_{ij}$. Let us look at $\mathbf{B}_{211}$, that is, the block for the second animal in $\mathbf{Q}_{11}$.

$$\mathbf{R}_2^- = \begin{pmatrix} .3059 & -.2000 & .0706 \\ -.2000 & .4000 & -.2000 \\ .0706 & -.2000 & .2471 \end{pmatrix}.$$

Then

$$\mathbf{B}_{211} = \begin{pmatrix} .3059 \\ -.2000 \\ .0706 \end{pmatrix} (.3059 \quad -.2000 \quad .0706) = \begin{pmatrix} .0936 & -.0612 & .0216 \\ & .0400 & -.0141 \\ & & .0050 \end{pmatrix}.$$

Look at $\mathbf{B}_{323}$, that is, the block for the third animal in $\mathbf{Q}_{23}$.

$$\mathbf{R}_3^- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & .2692 & -.1538 \\ 0 & -.1538 & .2308 \end{pmatrix}.$$

Then

$$\begin{aligned} \mathbf{B}_{323} &= \begin{pmatrix} 0 \\ .2692 \\ -.1538 \end{pmatrix} (0 \quad -.1538 \quad .2308) + \text{ transpose of this product} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -.0414 & .0621 \\ 0 & .0237 & -.03500 \end{pmatrix} + (\ )' \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -.0828 & .0858 \\ 0 & .0858 & -.0710 \end{pmatrix}. \end{aligned}$$

Next we compute $tr\mathbf{Q}_{ij}\mathbf{R}$. Consider the contribution of the first animal to $tr\mathbf{Q}_{12}\mathbf{R}$.

$$\mathbf{B}_{112} = \begin{pmatrix} -.1224 & .1624 & -.0753 \\ & -.1600 & .0682 \\ & & -.0282 \end{pmatrix}.$$

Then this animal contributes

$$-.1224\, r_{11} + .2(.1624)\, r_{12} - 2(.0753)\, r_{13} - .1600\, r_{22} + 2(.0682)\, r_{23} - .0282\, r_{33}.$$

Finally we illustrate computing a block of $\mathbf{WCW}'$ by (26.45). We use the third animal.

$$\mathbf{X}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

$$\mathbf{C}_{xx} = \begin{pmatrix} 29.4578 & -9.2266 & 3.0591 & 2.7324 & -.3894 \\ & 3.1506 & -.6444 & -.5254 & .0750 \\ & & 3.5851 & 2.3081 & .0305 \\ & & & 30.8055 & -8.5380 \\ & & & & 2.7720 \end{pmatrix}.$$

$$\mathbf{C}_{x3} = \begin{pmatrix} -1.7046 & -1.2370 & -1.0037 \\ .2759 & .2264 & .1692 \\ -.6174 & -1.9083 & -1.2645 \\ -.8627 & -1.2434 & -4.7111 \\ .0755 & -.0107 & .7006 \end{pmatrix}.$$

$$\mathbf{C}_{33} = \begin{pmatrix} 1.9642 & .9196 & .9374 \\ & 2.7786 & 1.8497 \\ & & 3.8518 \end{pmatrix}.$$

Then the computations of (26.45) give

$$\begin{pmatrix} 1.9642 & .3022 & .2257 \\ & 2.5471 & 1.6895 \\ & & 4.9735 \end{pmatrix}.$$

Since the first trait was missing on animal 3, the block of $\mathbf{WCW}'$ becomes

$$\begin{pmatrix} 0 & 0 & 0 \\ & 2.5471 & 1.6895 \\ & & 4.9735 \end{pmatrix}.$$

Combining these results, $\hat{\mathbf{r}}$ for the first round is the solution to

$$\begin{pmatrix} .227128 & -.244706 & .086367 & .080000 & -.056471 & .009965 \\ & .649412 & -.301176 & -.320000 & .272941 & -.056471 \\ & & .322215 & .160000 & -.254118 & .069758 \\ & & & .392485 & -.402840 & .103669 \\ & & & & .726892 & -.268653 \\ & & & & & .175331 \end{pmatrix} \hat{\mathbf{r}}$$

$$
\begin{aligned}
= \quad & (.137802,\ -.263298,\ .084767,\ .161811, -.101820,\ .029331)' \\
& + (.613393,\ -.656211,\ .263861,\ .713786,\ -.895139,\ .571375)'.
\end{aligned}
$$

This gives the solution

$$
\hat{\mathbf{R}} = \begin{pmatrix} 3.727 & 1.295 & .311 \\ & 3.419 & 2.270 \\ & & 4.965 \end{pmatrix} .
$$

Chapter 27
Sire Model, Multiple Traits

C. R. Henderson

1984 - Guelph

1 Only One Trait Observed On A Progeny

This section deals with a rather simple model in which there are t traits measured on the progeny of a set of sires. But the design is such that only one trait is measured on any progeny. This results in $\mathbf{R}$ being diagonal. It is assumed that each dam has only one recorded progeny, and the dams are non-inbred and unrelated. An additive genetic model is assumed. Order the observations by progeny within traits. There are t traits and k sires. Then the model is

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_t \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{X}_t \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_t \end{pmatrix} + \begin{pmatrix} \mathbf{Z}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{Z}_t \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_t \end{pmatrix} + \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_t \end{pmatrix} \quad (1)$$

$\mathbf{y}_i$ represents n_i progeny records on trait i, $\boldsymbol{\beta}_i$ is the vector of fixed effects influencing the records on the i^{th} trait, $\mathbf{X}_i$ relates $\boldsymbol{\beta}_i$ to elements of $\mathbf{y}_i$, and $\mathbf{s}_i$ is the vector of sire effects for the i^{th} trait. It has k has a null column corresponding to such a sire.

$$Var \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_t \end{pmatrix} = \begin{pmatrix} \mathbf{A}b_{11} & \mathbf{A}b_{12} & \ldots & \mathbf{A}b_{1t} \\ \mathbf{A}b_{12} & \mathbf{A}b_{22} & \ldots & \mathbf{A}b_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}b_{1t} & \mathbf{A}b_{2t} & \ldots & \mathbf{A}b_{tt} \end{pmatrix} = \mathbf{G}. \quad (2)$$

$\mathbf{A}$ is the $k \times k$ numerator relationship matrix for the sires. If the sires were unselected, $b_{ij} = g_{ij}/4$, where g_{ij} is the additive genetic covariance between traits i and j.

$$Var \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_t \end{pmatrix} = \begin{pmatrix} \mathbf{I}d_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{I}d_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{I}d_t \end{pmatrix} = \mathbf{R}. \quad (3)$$

Under the assumption of unselected sires

$$d_i = .75\ g_{ii} + r_{ii},$$

where r_{ii} is the i^{th} diagonal of the error covariance matrix of the usual multiple trait model. Then the GLS equations for fixed $\mathbf{s}$ are

$$\begin{pmatrix} d_1^{-1}\mathbf{X'}_1\mathbf{X}_1 & \ldots & \mathbf{0} & d_1^{-1}\mathbf{X'}_1\mathbf{Z}_1 & \ldots & \mathbf{0} \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{0} & \ldots & d_t^{-1}\mathbf{X'}_t\mathbf{X}_t & \mathbf{0} & \ldots & d_t^{-1}\mathbf{X'}_t\mathbf{Z}_t \\ d_1^{-1}\mathbf{Z'}_1\mathbf{X}_1 & \ldots & \mathbf{0} & d_1^{-1}\mathbf{Z'}_1\mathbf{Z}_1 & \ldots & \mathbf{0} \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{0} & \ldots & d_t^{-1}\mathbf{Z'}_t\mathbf{X}_t & \mathbf{0} & \ldots & d_t^{-1}\mathbf{Z'}_t\mathbf{Z}_t \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{\beta}_1^o \\ \vdots \\ \boldsymbol{\beta}_t^o \\ \hat{\mathbf{s}}_1 \\ \vdots \\ \hat{\mathbf{s}}_t \end{pmatrix} = \begin{pmatrix} d_1^{-1}\mathbf{X'}_1\mathbf{y}_1 \\ \vdots \\ d_t^{-1}\mathbf{X'}_t\mathbf{y}_t \\ d_1^{-1}\mathbf{Z'}_1\mathbf{y}_1 \\ \vdots \\ d_1^{-1}\mathbf{Z'}_t\mathbf{y}_t \end{pmatrix} \quad (4)$$

The mixed model equations are formed by adding $\mathbf{G}^{-1}$ to the lower right $(kt)^2$ submatrix of (27.4), where

$$\mathbf{G}^{-1} = \begin{pmatrix} \mathbf{A}^{-1}b^{11} & \ldots & \mathbf{A}^{-1}b^{1t} \\ \vdots & & \vdots \\ \mathbf{A}^{-1}b^{1t} & \ldots & \mathbf{A}^{-1}b^{tt} \end{pmatrix}, \quad (5)$$

and b^{ij} is the ij^{th} element of the inverse of

$$\begin{pmatrix} b_{11} & \ldots & b_{1t} \\ \vdots & & \vdots \\ b_{1t} & \ldots & b_{tt} \end{pmatrix}.$$

With this model it seems logical to estimate d_i by

$$[\mathbf{y'}_i\mathbf{y}_i - (\boldsymbol{\beta}_i^o)'\mathbf{X'}_i\mathbf{y}_i - (\mathbf{u}_i^o)'\mathbf{Z'}_i\mathbf{y}_i]/[n_i - \text{rank}(\mathbf{X}_i\ \ \mathbf{Z}_i)]. \quad (6)$$

$\boldsymbol{\beta}_i^o$ and $\mathbf{u}_i^o$ are some solution to (27.7)

$$\begin{pmatrix} \mathbf{X'}_i\mathbf{X}_i & \mathbf{X'}_i\mathbf{Z}_i \\ \mathbf{Z'}_i\mathbf{X}_i & \mathbf{Z'}_i\mathbf{Z}_i \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_i^o \\ \mathbf{u}_i^o \end{pmatrix} = \begin{pmatrix} \mathbf{X'}_i\mathbf{y}_i \\ \mathbf{Z'}_i\mathbf{y}_i \end{pmatrix}. \quad (7)$$

Then using these $\hat{d}_i$, estimate the b_{ij} by quadratics in $\hat{\mathbf{s}}$, the solution to (27.4). The quadratics needed are

$$\hat{\mathbf{s}}_i'\mathbf{A}^{-1}\hat{\mathbf{s}}_j; \quad i = 1,\ \ldots,\ t; \quad j = i,\ \ldots,\ t.$$

These are computed and equated to their expectations. We illustrate this section with a small example. The observations on progeny of three sires and two traits are

Sire	Trait	Progeny Records
1	1	5,3,6
2	1	7,4
3	1	5,3,8,6
1	2	5,7
2	2	9,8,6,5

Suppose $\mathbf{X}_1' = [1 \ldots 1]$ with 9 elements, and $\mathbf{X}_2' = [1 \ldots 1]$ with 6 elements.

$$\mathbf{Z}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{Z}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Suppose that

$$\mathbf{R} = \begin{pmatrix} 30\,\mathbf{I}_9 & \mathbf{0} \\ \mathbf{0} & 25\,\mathbf{I}_6 \end{pmatrix},$$

$$\mathbf{A} = \begin{pmatrix} 1. & .5 & .5 \\ .5 & 1. & .25 \\ .5 & .25 & 1. \end{pmatrix},$$

and

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}.$$

Then

$$\mathbf{G} = \begin{pmatrix} 3. & 1.5 & 1.5 & 1. & .5 & .5 \\ & 3. & .75 & .5 & 1. & .25 \\ & & 3. & .5 & .25 & 1. \\ & & & 2. & 1. & 1. \\ & & & & 2. & .5 \\ & & & & & 2. \end{pmatrix},$$

$$\mathbf{G}^{-1} = \begin{pmatrix} 10 & -4 & -4 & -5 & 2 & 2 \\ & 8 & 0 & 2 & -4 & 0 \\ & & 8 & 2 & 0 & -4 \\ & & & 15 & -6 & -6 \\ & & & & 12 & 0 \\ & & & & & 12 \end{pmatrix} \frac{1}{15}.$$

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1} \\ \mathbf{Z}'\mathbf{R}^{-1} \end{pmatrix} (\mathbf{X} \quad \mathbf{Z}) = \frac{1}{150} \begin{pmatrix} 45 & 0 & 15 & 10 & 20 & 0 & 0 & 0 \\ & 36 & 0 & 0 & 0 & 12 & 24 & 0 \\ & & 15 & 0 & 0 & 0 & 0 & 0 \\ & & & 10 & 0 & 0 & 0 & 0 \\ & & & & 20 & 0 & 0 & 0 \\ & & & & & 12 & 0 & 0 \\ & & & & & & 24 & 0 \\ & & & & & & & 0 \end{pmatrix} \quad (8)$$

Adding $\mathbf{G}^{-1}$ to the lower 6 × 6 submatrix of (27.8) gives the mixed model coefficient matrix. The right hand sides are [1.5667, 1.6, .4667, .3667, .7333, .48, 1.12, 0]. The inverse of the mixed model coefficient matrix is

$$\begin{pmatrix} 5.210 & 0.566 & -1.981 & -1.545 & -1.964 & -0.654 & -0.521 & -0.652 \\ 0.566 & 5.706 & -0.660 & -0.785 & -0.384 & -1.344 & -1.638 & -0.690 \\ -1.981 & -0.660 & 2.858 & 1.515 & 1.556 & 0.934 & 0.523 & 0.510 \\ -1.545 & -0.785 & 1.515 & 2.803 & 0.939 & 0.522 & 0.917 & 0.322 \\ -1.964 & -0.384 & 1.556 & 0.939 & 2.783 & 0.510 & 0.322 & 0.923 \\ -0.654 & -1.344 & 0.934 & 0.522 & 0.510 & 1.939 & 1.047 & 0.984 \\ -0.521 & -1.638 & 0.523 & 0.917 & 0.322 & 1.047 & 1.933 & 0.544 \\ -0.652 & -0.690 & 0.510 & 0.322 & 0.923 & 0.984 & 0.544 & 1.965 \end{pmatrix} \quad (9)$$

The solution to the MME is (5.2380, 6.6589, -.0950, .0236, .0239, -.0709, .0471, -.0116).

2 Multiple Traits Recorded On A Progeny

When multiple traits are observed on individual progeny, $\mathbf{R}$ is no longer diagonal. The linear model can still be written as (27.1). Now, however, the $\mathbf{y}_i$ do not have the same number of elements, and $\mathbf{X}_i$ and $\mathbf{Z}_i$ have varying numbers of rows. Further,

$$\mathbf{R} = \begin{pmatrix} \mathbf{I} \ r_{11} & \mathbf{P}_{12} r_{12} & \ldots & \mathbf{P}_{1t} r_{1t} \\ \mathbf{P}'_{12} r_{12} & \mathbf{I} \ r_{22} & \ldots & \mathbf{P}_{2t} r_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{P}'_{1t} r_{1t} & \mathbf{P}'_{2t} r_{2t} & \ldots & \mathbf{I} \ r_{tt} \end{pmatrix}. \quad (10)$$

The $\mathbf{I}$ matrices have order equal to the number of progeny with that trait recorded.

$$\begin{pmatrix} r_{11} & \ldots & r_{1t} \\ \vdots & & \vdots \\ r_{1t} & \ldots & r_{tt} \end{pmatrix}$$

is the error variance-covariance matrix. We can use the same strategy as in Chapter 25 for missing data. That is, each $\mathbf{y}_i$ is the same length with 0's inserted for missing data.

Accordingly, all $\mathbf{X}_i$ and $\mathbf{Z}_i$ have the same number of rows with rows pertaining to missing observations set to 0. Further, $\mathbf{R}$ is the same as for no missing data except that rows corresponding to missing observations are set to 0. Then the zeroed type of g-inverse of $\mathbf{R}$ is

$$\begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} & \ldots & \mathbf{D}_{1t} \\ \mathbf{D}_{12} & \mathbf{D}_{22} & \ldots & \mathbf{D}_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{D}_{1t} & \mathbf{D}_{2t} & \ldots & \mathbf{D}_{tt} \end{pmatrix}. \tag{11}$$

Each of the $\mathbf{D}_{ij}$ is diagonal with order, n. Now the GLS equations for fixed $\mathbf{s}$ are

$$\begin{pmatrix} \mathbf{X'}_1\mathbf{D}_{11}\mathbf{X}_1 & \ldots & \mathbf{X'}_1\mathbf{D}_{1t}\mathbf{X}_t & \mathbf{X'}_1\mathbf{D}_{11}\mathbf{Z}_1 & \ldots & \mathbf{X'}_1\mathbf{D}_{1t}\mathbf{Z}_t \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{X'}_t\mathbf{D}_{1t}\mathbf{X}_1 & \ldots & \mathbf{X'}_t\mathbf{D}_{tt}\mathbf{X}_t & \mathbf{X'}_t\mathbf{D}_{1t}\mathbf{Z}_1 & \ldots & \mathbf{X'}_t\mathbf{D}_{tt}\mathbf{Z}_t \\ \mathbf{Z'}_1\mathbf{D}_{11}\mathbf{X}_1 & \ldots & \mathbf{Z'}_1\mathbf{D}_{1t}\mathbf{X}_t & \mathbf{Z'}_1\mathbf{D}_{11}\mathbf{Z}_1 & \ldots & \mathbf{Z'}_1\mathbf{D}_{1t}\mathbf{Z}_t \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{Z'}_t\mathbf{D}_{1t}\mathbf{X}_1 & \ldots & \mathbf{Z'}_t\mathbf{D}_{tt}\mathbf{X}_t & \mathbf{Z'}_t\mathbf{D}_{1t}\mathbf{Z}_1 & \ldots & \mathbf{Z'}_t\mathbf{D}_{tt}\mathbf{Z}_t \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1^o \\ \vdots \\ \boldsymbol{\beta}_t^o \\ \hat{s}_1 \\ \vdots \\ \hat{s}_t \end{pmatrix} =$$

$$\begin{pmatrix} \mathbf{X'}_1\mathbf{D}_{11}\mathbf{y}_1 & + & \ldots & + & \mathbf{X'}_1\mathbf{D}_{1t}\mathbf{y}_t \\ \vdots & & & & \vdots \\ \mathbf{X'}_t\mathbf{D}_{1t}\mathbf{y}_1 & + & \ldots & + & \mathbf{X'}_t\mathbf{D}_{tt}\mathbf{y}_t \\ \mathbf{Z'}_1\mathbf{D}_{11}\mathbf{y}_1 & + & \ldots & + & \mathbf{Z'}_1\mathbf{D}_{1t}\mathbf{y}_t \\ \vdots & & & & \vdots \\ \mathbf{Z'}_t\mathbf{D}_{1t}\mathbf{y}_1 & + & \ldots & + & \mathbf{Z'}_t\mathbf{D}_{tt}\mathbf{y}_t \end{pmatrix}. \tag{12}$$

With $\mathbf{G}^{-1}$ added to the lower part of (27.12) we have the mixed model equations.

We illustrate with the following example.

		Trait	
Sire	Progeny	1	2
1	1	6	5
	2	3	5
	3	-	7
	4	8	-
2	5	4	6
	6	-	7
	7	3	-
3	8	5	4
	9	8	-

We assume the same $\mathbf{G}$ as in the illustration of Section 27.1, and

$$\begin{pmatrix} r_{11} & r_{12} \\ r_{12} & r_{22} \end{pmatrix} = \begin{pmatrix} 30 & 10 \\ 10 & 25 \end{pmatrix}.$$

We assume that the only fixed effects are μ_1 and μ_2 . Then using the data vector with length 13, ordered progeny in sire in trait,

$$\mathbf{X}_1' = (1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1), \quad \mathbf{X}_2' = (1 \ 1 \ 1 \ 1 \ 1 \ 1),$$

$$\mathbf{Z}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{Z}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{y}_1' = (6 \ 3 \ 8 \ 4 \ 3 \ 5 \ 8), \quad \mathbf{y}_2' = (5 \ 5 \ 7 \ 6 \ 7 \ 4),$$

and

$$\mathbf{R} = \begin{pmatrix} 30 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 \\ & 30 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0 & 0 & 0 & 0 \\ & & 30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 30 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0 & 0 \\ & & & & 30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & 30 & 0 & 0 & 0 & 0 & 0 & 0 & 10 \\ & & & & & & 30 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & & 25 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & & & 25 & 0 & 0 & 0 & 0 \\ & & & & & & & & & 25 & 0 & 0 & 0 \\ & & & & & & & & & & 25 & 0 & 0 \\ & & & & & & & & & & & 25 & 0 \\ & & & & & & & & & & & & 25 \end{pmatrix}.$$

Then the GLS coefficient matrix for fixed **s** is in (27.13).

$$\begin{pmatrix} 0.254 & -0.061 & 0.110 & 0.072 & 0.072 & -0.031 & -0.015 & -0.015 \\ -0.061 & 0.265 & -0.031 & -0.015 & -0.015 & 0.132 & 0.086 & 0.046 \\ 0.110 & -0.031 & 0.110 & 0.0 & 0.0 & -0.031 & 0.0 & 0.0 \\ 0.072 & -0.015 & 0.0 & 0.072 & 0.0 & 0.0 & -0.015 & 0.0 \\ 0.071 & -0.015 & 0.0 & 0.0 & 0.072 & 0.0 & 0.0 & -0.015 \\ -0.031 & 0.132 & -0.031 & 0.0 & 0.0 & 0.132 & 0.0 & 0.0 \\ -0.015 & 0.086 & 0.0 & -0.015 & 0.0 & 0.0 & 0.086 & 0.0 \\ -0.015 & 0.046 & 0.0 & 0.0 & -0.015 & 0.0 & 0.0 & 0.046 \end{pmatrix} \tag{13}$$

$\mathbf{G}^{-1}$ is added to the lower 6×6 submatrix to form the mixed model coefficient matrix. The right hand sides are (1.0179, 1.2062, .4590, .1615, .3974, .6031, .4954, .1077). The inverse of the coefficient matrix is

$$\begin{pmatrix} 6.065 & 1.607 & -2.111 & -1.735 & -1.709 & -0.702 & -0.603 & -0.546 \\ 1.607 & 5.323 & -0.711 & -0.625 & -0.519 & -1.472 & -1.246 & -1.017 \\ -2.111 & -0.711 & 2.880 & 1.533 & 1.527 & 0.953 & 0.517 & 0.506 \\ -1.735 & -0.625 & 1.533 & 2.841 & 0.893 & 0.517 & 0.938 & 0.303 \\ -1.709 & -0.519 & 1.527 & 0.893 & 2.844 & 0.506 & 0.303 & 0.944 \\ -0.702 & -1.472 & 0.953 & 0.517 & 0.506 & 1.939 & 1.028 & 1.003 \\ -0.602 & -1.246 & 0.517 & 0.938 & 0.303 & 1.028 & 1.924 & 0.562 \\ -0.546 & -1.017 & 0.506 & 0.303 & 0.943 & 1.002 & 0.562 & 1.936 \end{pmatrix}. \tag{14}$$

The solution is [5.4038, 5.8080, .0547, -.1941, .1668, .0184, .0264, -.0356].

If we use the technique of including in $\mathbf{y}$, $\mathbf{X}$, $\mathbf{Z}$, $\mathbf{R}$, $\mathbf{G}$ the missing data we have

$$\mathbf{X}_1' = (1\ \ 1\ \ 0\ \ 1\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1), \quad \mathbf{X}_2' = (1\ \ 1\ \ 1\ \ 0\ \ 1\ \ 1\ \ 0\ \ 1\ \ 0),$$

$$\mathbf{Z}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{Z}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\mathbf{y}_1' = (6,\ 3,\ 0,\ 8,\ 4,\ 0,\ 3,\ 5,\ 8),$$

and

$$\mathbf{y}_2' = (5,\ 5,\ 7,\ 0,\ 6,\ 7,\ 0,\ 4,\ 0).$$

$$\begin{aligned} \mathbf{D}_{11} &= \text{diag}[.0385,\ .0385,\ 0,\ .0333,\ .0385,\ 0,\ .0333,\ .0385,\ .0333] \\ \mathbf{D}_{12} &= \text{diag}[-.0154,\ -.0154,\ 0,\ 0,\ -.0154,\ 0,\ 0,\ -.0154,\ 0] \\ \mathbf{D}_{22} &= \text{diag}[.0462,\ .0462,\ .04,\ 0,\ .0462,\ .04,\ 0,\ .0462,\ 0]. \end{aligned}$$

This leads to the same set of equations and solution as when $\mathbf{y}$ has 13 elements.

3 Relationship To Sire Model With Repeated Records On Progeny

The methods of Section 27.2 could be used for sire evaluation using progeny with repeated records (lactations, e.g.), but we do not wish to invoke the simple repeatability model. Then lactation 1 is trait 1, lactation 2 is trait 2, etc.

Chapter 28
Joint Cow and Sire Evaluation

C. R. Henderson

1984 - Guelph

At the present time, (1984), agencies evaluating dairy sires and dairy females have designed separate programs for each. Sires usually have been evaluated solely on the production records of their progeny. With the development of an easy method for computing $\mathbf{A}^{-1}$ this matrix has been incorporated by some agencies, and that results in the evaluation being a combination of the progeny test of the individual in question as well as progeny tests of his relatives, eg., sire and paternal brothers. In addition, the method also takes into account the predictions of the merits of the sires of the mates of the bulls being evaluated. This is an approximation to the merits of the mates without using their records.

In theory one could utilize all records available in a production testing program and could compute $\mathbf{A}^{-1}$ for all animals that have produced these records as well as additional related animals without records that are to be evaluated. Then these could be incorporated into a single set of prediction equations. This, of course, could result in a set of equations that would be much too large to solve with existing computers. Nevertheless, if we are willing to sacrifice some accuracy by ignoring the fact that animals change herds, we can set up equations that are block diagonal in form that may be feasible to solve.

1 Block Diagonality Of Mixed Model Equations

Henderson (1976) presented a method for rapid calculation of $\mathbf{A}^{-1}$ without computing $\mathbf{A}$. A remarkable property of $\mathbf{A}^{-1}$ is that the only non-zero off-diagonal elements are those pertaining to a pair of mates, and those pertaining to parent - progeny pairs. These non-zero elements can be built up by entering the data in any order, with each piece of data incorporating the individual identification number, the sire number, and the dam number. At the same time one could enter with this information the production record and elements of the incidence matrix of the individual. Now when the dam and her progeny are in different herds, we pretend that we do not know the dam of the progeny and if, when a natural service sire has progeny in more than one herd, we treat him as a different sire in each herd, there are no non-zero elements of $\mathbf{A}^{-1}$ between herds. This strategy, along with the fact that most if not all elements of $\boldsymbol{\beta}$ are peculiar to the individual herd, results in the mixed model coefficient matrix having a block diagonal form. The elements of the model are ordered as follows

$\boldsymbol{\beta}_0$: a subvector of $\boldsymbol{\beta}$ common to all elements of $\mathbf{y}$.

$\mathbf{a}_0$: a subvector of $\mathbf{a}$, additive genetic values, pertaining to sires used in several herds.

$\boldsymbol{\beta}_i$ ($i = 1, \ldots,$ number of herds): a subvector of $\boldsymbol{\beta}$ pertaining only to records in the i^{th} herd.

$\mathbf{a}_i$: a subvector of $\mathbf{a}$ pertaining to animals in the i^{th} herd. $\mathbf{a}_i$ can represent cows with records, or dams and non-AI sires of the cows with records. In computing $\mathbf{A}^{-1}$ for the animals in the i^{th} herd the dam is assumed unknown if it is in a different herd. When Sectiom 28.3 method is used (multiple records) no records of a cow should be used in a herd unless the first lactation record is available. This restriction prevents using records of a cow that moves to another herd subsequent to first lactation. With this ordering and with these restrictions in computing $\mathbf{A}^{-1}$ the BLUP equations have the following form

$$\begin{pmatrix} \mathbf{C}_{00} & \mathbf{C}_{01} & \mathbf{C}_{02} & \cdots & \mathbf{C}_{0k} \\ \mathbf{C}'_{01} & \mathbf{C}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}'_{02} & \mathbf{0} & \mathbf{C}_{22} & \cdots & \mathbf{0} \\ \vdots & & & & \vdots \\ \mathbf{C}'_{0k} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_{kk} \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_k \end{pmatrix} = \begin{pmatrix} \mathbf{r}_0 \\ \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_k \end{pmatrix}.$$

$$\begin{aligned} \gamma'_0 &= (\boldsymbol{\beta}'_0 \ \ \mathbf{a}'_0), \\ \gamma'_i &= (\boldsymbol{\beta}'_i \ \ \mathbf{a}'_i). \end{aligned}$$

Then with this form of the equations the herd unknowns can be "absorbed" into the $\boldsymbol{\beta}_0$ and $\mathbf{a}_0$ equations provided the $\mathbf{C}_{ii}$ blocks can be readily inverted. Otherwise one would need to solve iteratively. For example, one might first solve iteratively for $\boldsymbol{\beta}_0$ and $\mathbf{a}_0$ sires ignoring $\boldsymbol{\beta}_i$, $\mathbf{a}_i$. Then with these values one would solve iteratively for the herd values. Having obtained these one would re-solve for $\boldsymbol{\beta}_0$ and the $\mathbf{a}_0$ values, adjusting the right hand sides for the previously estimated herd values.

The AI sire equations would also contain values for the "base population" sires. A base population dam with records would be included with the herd in which its records were made. Any base population dam that has no records, has only one AI son, and has no female progeny can be ignored without changing the solution.

2 Single Record On Single Trait

The simplest example of joint cow and sire evaluation with multiple herds involves a single trait and with only one record per tested animal. We illustrate this with the following example.

Base population animals

1 male
2 female with record in herd 1
3 female with record in herd 2

AI Sires

4 with parents 1 and 2
5 with parents 1 and 3

Other Females With Records

6 with unknown parents, record in herd 1
7 with unknown parents, record in herd 2
8 with parents 4 and 6, record in herd 1
9 with parents 4 and 3, record in herd 2
10 with parents 5 and 7, record in herd 2
11 with parents 5 and 2, record in herd 1

Ordering these animals (1,4,5,2,6,8,11,3,7,9,10) the **A** matrix is in (28.1).

$$\begin{pmatrix}
1 & .5 & .5 & 0 & 0 & .25 & .25 & 0 & 0 & .25 & .25 \\
 & 1 & .25 & .5 & 0 & .5 & .375 & 0 & 0 & .5 & .125 \\
 & & 1 & 0 & 0 & .125 & .5 & .5 & 0 & .375 & .5 \\
 & & & 1 & 0 & .25 & .5 & 0 & 0 & .25 & 0 \\
 & & & & 1 & .5 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 1 & .1875 & 0 & 0 & .25 & .0625 \\
 & & & & & & 1 & .25 & 0 & .3125 & .25 \\
 & & & & & & & 1 & 0 & .5 & .25 \\
 & & & & & & & & 1 & 0 & .5 \\
 & & & & & & & & & 1 & .1875 \\
 & & & & & & & & & & 1
\end{pmatrix} \tag{1}$$

$\mathbf{A}^{-1}$ shown in (28.2)

$$\begin{pmatrix}
2 & -1 & -1 & .5 & 0 & 0 & 0 & .5 & 0 & 0 & 0 \\
 & 3 & 0 & -1 & .5 & -1 & 0 & .5 & 0 & -1 & 0 \\
 & & 3 & .5 & 0 & 0 & -1 & -1 & .5 & 0 & -1 \\
 & & & 2 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
 & & & & 1.5 & -1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 2 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 2 & 0 & 0 & 0 & 0 \\
 & & & & & & & 2 & 0 & -1 & 0 \\
 & & & & & & & & 1.5 & 0 & -1 \\
 & & & & & & & & & 2 & 0 \\
 & & & & & & & & & & 2
\end{pmatrix} \tag{2}$$

Note that the lower 8×8 submatrix is block diagonal with two blocks of order 4×4 down the diagonal and 4×4 null off-diagonal blocks. The model assumed for our illustration is

$$y_{ij} = \mu_i + a_{ij} + e_{ij},$$

where i refers to herd and j to individual within herd. Then with ordering

$$(a_1,\ a_4,\ a_5,\ \mu_1,\ a_2\ a_6,\ a_8,\ a_{11},\ \mu_2,\ a_3\ a_7,\ a_9,\ a_{10})$$

the incidence matrix is as shown in (28.3). Note that $\boldsymbol{\beta}^o$ does not exist.

$$\begin{pmatrix}
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1
\end{pmatrix} \tag{3}$$

Suppose $\mathbf{y}$' = [3,2,5,6,7,9,2,3] corresponding to animals 2,6,8,11,3,7,9,10. We assume that $h^2 = .25$ which implies $\sigma_e^2/\sigma_a^2 = 3$. Then adding 3 $\mathbf{A}^{-1}$ to appropriate elements of the OLS equations we obtain mixed model equations displayed in (28.4).

$$
\begin{pmatrix}
6 & -3 & -3 & 0 & 1.5 & 0 & 0 & 0 & 0 & 1.5 & 0 & 0 & 0 \\
 & 9 & 0 & 0 & -3 & 1.5 & -3 & 0 & 0 & 1.5 & 0 & -3 & 0 \\
 & & 9 & 0 & 1.5 & 0 & 0 & -3 & 0 & -3 & 1.5 & 0 & -3 \\
 & & & 4 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 7 & 0 & 0 & -3 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 5.5 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 7 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 7 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & 4 & 1 & 1 & 1 & 1 \\
 & & & & & & & & & 7 & 0 & -3 & 0 \\
 & & & & & & & & & & 5.5 & 0 & -3 \\
 & & & & & & & & & & & 7 & 0 \\
 & & & & & & & & & & & & 7
\end{pmatrix}
\begin{pmatrix}
\hat{a}_1 \\ \hat{a}_4 \\ \hat{a}_5 \\ \hat{\mu}_1 \\ \hat{a}_2 \\ \hat{a}_6 \\ \hat{a}_8 \\ \hat{a}_{11} \\ \hat{\mu}_2 \\ \hat{a}_3 \\ \hat{a}_7 \\ \hat{a}_9 \\ \hat{a}_{10}
\end{pmatrix}
=
\begin{pmatrix}
0 \\ 0 \\ 0 \\ 16 \\ 3 \\ 2 \\ 5 \\ 6 \\ 21 \\ 7 \\ 9 \\ 2 \\ 3
\end{pmatrix}
\qquad (4)
$$

Note that the lower 10×10 block of the coefficient matrix is block diagonal with 5×5 blocks down the diagonal and 5×5 null blocks off-diagonal. The solution to (28.4) is

[-.1738, -.3120, -.0824, 4.1568, -.1793, -.4102, -.1890, .1512, 5.2135, .0857, .6776, -.5560, -.0611].

Note that the solution to $(a_1,\ a_4,\ a_5)$ could be found by absorbing the other equations as follows.

$$
\left\{ \begin{pmatrix} 6 & -3 & -3 \\ & 9 & 0 \\ & & 9 \end{pmatrix} - \begin{pmatrix} 0 & 1.5 & 0 & 0 & 0 \\ 0 & -3 & 1.5 & -3 & 0 \\ 0 & 1.5 & 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} 4 & 1 & 1 & 1 & 1 \\ & 7 & 0 & 0 & -3 \\ & & 5.5 & -3 & 0 \\ & & & 7 & 0 \\ & & & & 7 \end{pmatrix}^{-1} \right.
$$

$$
\begin{pmatrix} 0 & 0 & 0 \\ 1.5 & -3 & 1.5 \\ 0 & 1.5 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}
-
\begin{pmatrix} 0 & 1.5 & 0 & 0 & 0 \\ 0 & 1.5 & 0 & -3 & 0 \\ 0 & -3 & 1.5 & 0 & -3 \end{pmatrix}
\begin{pmatrix} 4 & 1 & 1 & 1 & 1 \\ & 7 & 0 & -3 & 0 \\ & & 5.5 & 0 & -3 \\ & & & 7 & 0 \\ & & & & 7 \end{pmatrix}^{-1}
$$

$$
\left.\begin{pmatrix} 0 & 0 & 0 \\ 1.5 & 1.5 & -3 \\ 0 & 0 & 1.5 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}\right\}
\begin{pmatrix} a_1 \\ a_4 \\ a_5 \end{pmatrix}
=
\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
-
\begin{pmatrix} 0 & 1.5 & 0 & 0 & 0 \\ 0 & -3 & 1.5 & -3 & 0 \\ 0 & 1.5 & 0 & 0 & -3 \end{pmatrix}
$$

$$
\begin{pmatrix} 4 & 1 & 1 & 1 & 1 \\ & 7 & 0 & 0 & -3 \\ & & 5.5 & -3 & 0 \\ & & & 7 & 0 \\ & & & & 7 \end{pmatrix}^{-1}
\begin{pmatrix} 16 \\ 3 \\ 2 \\ 5 \\ 6 \end{pmatrix}
-
\begin{pmatrix} 0 & 1.5 & 0 & 0 & 0 \\ 0 & 1.5 & 0 & -3 & 0 \\ 0 & -3 & 1.5 & 0 & -3 \end{pmatrix}
$$

$$
\begin{pmatrix} 4 & 1 & 1 & 1 & 1 \\ & 7 & 0 & -3 & 0 \\ & & 5.5 & 0 & -3 \\ & & & 7 & 0 \\ & & & & 7 \end{pmatrix}^{-1}
\begin{pmatrix} 21 \\ 7 \\ 9 \\ 2 \\ 3 \end{pmatrix}.
$$

Iterations on these equations were carried out by two different methods. First, the herd equations were iterated 5 rounds with AI sire values fixed. Then the AI sire equations were iterated 5 rounds with the herd values fixed and so on. It required 17 cycles (85 rounds) to converge to the direct solution previously reported. Regular Gauss-Seidel iteration produced conversion in 33 rounds. The latter procedure would require more retrieval of data from external storage devices.

3 Simple Repeatability Model

As our next example we use the same animals as before but now we have records as follows.

	Herd 1 Years				Herd 2 Years		
Cow	1	2	3	Cow	1	2	3
2	5	6	-	3	8	-	-
6	4	5	3	7	9	8	7
8	-	7	6	9	-	8	8
11	-	-	8	10	-	-	7

We assume a model,

$$y_{ijk} = \mu_{ij} + a_{ik} + p_{ik} + e_{ijk}.$$

i refers to herd, j to year, and k to cow. It is assumed that $h^2 = .25$, $r = .45$. Then

$$\begin{aligned} Var(\mathbf{a}) &= \mathbf{A}\ \sigma_e^2/2.2. \\ Var(\mathbf{p}) &= \mathbf{I}\ \sigma_e^2/2.75. \\ \sigma_e^2/\sigma_a^2 &= 2.2,\ \sigma_e^2/\sigma_p^2 = 2.75 \end{aligned}$$

The diagonal coefficients of the $\mathbf{p}$ equations of OLS have added to them 2.75. Then $\hat{\mathbf{p}}$ can be absorbed easily. This can be done without writing the complete equations by weighting each observation by

$$\frac{2.75}{n_{ik} + \sigma_e^2/2.75}$$

where n_{ik} is the number of records on the ik^{th} cow. These weights are .733, .579, .478 for 1,2,3 records respectively. Once these equations are derived, we then add 2.2 $\mathbf{A}^{-1}$ to appropriate coefficients to obtain the mixed model equations. The coefficient matrix is in (28.5) ... (28.7), and the right hand side vector is (0, 0, 0, 4.807, 9.917, 10.772, 6.369, 5.736, 7.527, 5.864, 10.166, 8.456, 13.109, 5.864, 11.472, 9.264, 5.131)'. The unknowns are in this order (a_1, a_4, a_5, μ_{11}, μ_{12}, μ_{13}, a_2, a_6, a_8, a_{11}, μ_{21}, μ_{22}, μ_{23}, a_3, a_7, a_9, a_{10}). Note that block diagonality has been retained. The solution is

[.1956, .3217, .2214, 4.6512, 6.127, 5.9586, .2660, -.5509, .0045, .5004, 8.3515, 7.8439, 7.1948, .0377, .0516, .2424, .0892].

Upper 8 × 8

$$\begin{pmatrix} 4.4 & -2.2 & -2.2 & 0 & 0 & 0 & 1.1 & 0 \\ & 6.6 & 0 & 0 & 0 & 0 & -2.2 & 1.1 \\ & & 6.6 & 0 & 0 & 0 & 1.1 & 0 \\ & & & 1.057 & 0 & 0 & .579 & .478 \\ & & & & 1.636 & 0 & .579 & .478 \\ & & & & & 1.79 & 0 & .478 \\ & & & & & & 5.558 & 0 \\ & & & & & & & 4.734 \end{pmatrix} \quad (5)$$

Upper right 8 × 9 and (lower left 9 × 8)'

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1.1 & 0 & 0 & 0 \\ -2.2 & 0 & 0 & 0 & 0 & 1.1 & 0 & -2.2 & 0 \\ 0 & -2.2 & 0 & 0 & 0 & -2.2 & 1.1 & 0 & -2.2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .579 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .579 & .733 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (6)$$

Lower right 9 $\times$ 9

$$\begin{pmatrix}
5.558 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & 5.133 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & 1.211 & 0 & 0 & .733 & .478 & 0 & 0 \\
 & & & 1.057 & 0 & 0 & .478 & .579 & 0 \\
 & & & & 1.79 & 0 & .478 & .579 & .733 \\
 & & & & & 5.133 & 0 & -2.2 & 0 \\
 & & & & & & 4.734 & 0 & -2.2 \\
 & & & & & & & 5.558 & 0 \\
 & & & & & & & & 5.133
\end{pmatrix} \quad (7)$$

4 Multiple Traits

As a final example of joint cow and sire evaluation we evaluate on two traits. Using the same animals as before the records are as follows.

	Herd 1 Trait			Herd 2 Trait	
Cow	1	2	Cow	1	2
2	6	8	3	7	-
6	4	6	7	-	2
8	9	-	9	-	8
11	-	3	10	6	9

We assume a model,

$$y_{ijk} = \mu_{ij} + a_{ijk} + e_{ijk},$$

where i refers to herd, j to trait, and k to cow. We assume that the error variance-covariance matrix for a cow and the additive genetic variance-covariance matrix for a non-inbred individual are

$$\begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix} \text{ and } \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix},$$

respectively. Then $\mathbf{R}$ is

$$\begin{pmatrix} 5 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 5 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & 8 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & 5 & 0 & 0 & 0 & 0 \\ & & & & & & & 8 & 0 & 0 & 0 \\ & & & & & & & & 8 & 0 & 0 \\ & & & & & & & & & 5 & 2 \\ & & & & & & & & & & 8 \end{pmatrix}.$$

Ordering traits within animals, $\mathbf{G}$ is composed of 2×2 blocks as follows

$$\begin{pmatrix} 2\ a_{ij} & a_{ij} \\ a_{ij} & 3\ a_{ij} \end{pmatrix}.$$

The right hand sides of the mixed model equations are (0, 0, 0, 0, 0, 0, 3.244, 1.7639, .8889, .7778, .5556, .6111, 1.8, 0, 0, .375, 2.3333, 2.1167, 1.4, 0, 0, .25, 0, 1., .8333, .9167) corresponding to ordering of equations,

$[a_{11}, a_{12}, a_{41}, a_{42}, a_{51}, a_{52}, \mu_{11}, \mu_{12}, a_{21}, a_{22}, a_{61}, a_{62}, a_{81}, a_{82}, a_{11,1}, a_{11,2}, \mu_{21}, \mu_{22}, a_{31}, a_{32}, a_{71}, a_{72}, a_{91}, a_{92}, a_{10,1}, a_{10,2}]$.

The coefficient matrix is block diagonal with two 10×10 blocks in the lower diagonal and with two 10×10 null blocks off-diagonal. The solution is

(.2087, .1766, .4469, .4665, .1661, .1912, 5.9188, 5.9184, .1351, .3356, -.1843, .1314, .6230, .5448, -.0168, -.2390, 6.0830, 6.5215, .2563, .2734, -.4158, -.9170, .4099, .5450, -.0900, .0718).

5 Summary Of Methods

The model to be used contains the following elements.

1. $\mathbf{X}_0\boldsymbol{\beta}_0$: pertaining to all records

2. $\mathbf{X}_i\boldsymbol{\beta}_i$: pertaining to records only on the i^{th} herd.

3. $\mathbf{Z}_0\mathbf{a}_0$: additive genetic values of sires used in several herds, AI sires in particular, but could include natural service sires used in several herds.

4. $\mathbf{Z}_i\mathbf{a}_i$: additive genetic values of all females that have made records in the i^{th} herd. Some of these may be dams of AI sires. Others will be daughters of AI sires, and some will be both dams and daughters of different AI sires. $\mathbf{Z}_i\mathbf{a}_i$ will also contain any sire with daughters only in the i^{th} herd or with daughters in so few other herds that this is ignored, and he is regarded as a different sire in each of the other herds. One will need to decide how to handle such sires, that is, how many to include with AI sires and how many to treat as a separate sire in each of the herds in which he has progeny.

5. $\mathbf{A}^{-1}$ should be computed by Henderson's simple method, possibly ignoring inbreeding in large data sets, since this reduces computations markedly. In order to generate block diagonality in the mixed model equations the elements of $\mathbf{A}^{-1}$ for animals in $\mathbf{Z}_i\mathbf{a}_i$ should be derived only from sires in $\mathbf{a}_0$ and from dams and sires in $\mathbf{a}_i$ (same herd). This insures that there will be no non-zero elements of $\mathbf{A}^{-1}$ between any pair of herds, provided ordering is done according to the following

$$\begin{array}{ll}
(1) & \mathbf{X}_0\boldsymbol{\beta}_0 \\
(2) & \mathbf{Z}_0\mathbf{a}_0 \\
(3) & \mathbf{X}_1\boldsymbol{\beta}_1 \\
(4) & \mathbf{Z}_1\mathbf{a}_1 \\
(5) & \mathbf{X}_2\boldsymbol{\beta}_2 \\
(6) & \mathbf{Z}_2\mathbf{a}_2 \\
 & \vdots \\
 & \text{etc.}
\end{array}$$

6 Gametic Model To Reduce The Number Of Equations

Quaas and Pollak (1980) described a gametic additive genetic model that reduces the number of equations needed for computing BLUP. The only breeding values appearing in the equations are those of animals having tested progeny. Then individuals with no progeny can be evaluated by taking appropriate linear functions of the solution vector. The paper cited above dealt with multiple traits. We shall consider two situations, (1) single traits with one or no record per trait and (2) single traits with multiple records and the usual repeatability model assumed. If one does not choose to assume the repeatability model, the different records in a trait can be regarded as multiple traits and the Quaas and Pollak method used.

6.1 Single record model

Let the model be

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_a\mathbf{a} + \text{other possible random factors} + \mathbf{e}.$$

There are b animals with tested progeny, and c $\leq$ b of these parents are tested. There are d tested animals with no progeny. Thus $\mathbf{y}$ has $c + d$ elements. In the regular mixed model $\mathbf{a}$ has $b + d$ elements. $\mathbf{Z}_a$ is formed from an identity matrix of order $b + d$ and then deleting $b - c$ rows corresponding to parents with no record.

$$\begin{aligned} Var(\mathbf{a}) &= \mathbf{A}\sigma_a^2, \\ Var(\mathbf{e}) &= \mathbf{I}\sigma_e^2, \\ Cov(\mathbf{a}, \mathbf{e}') &= \mathbf{0}. \end{aligned}$$

Now in the gametic model, which is linearly equivalent to the model above, $\mathbf{a}$ has only b elements corresponding to the animals with tested progeny. As before $\mathbf{y}$ has $c + d$ elements, and is ordered such that records of animals with progeny appear first.

$$\mathbf{Z}_a = \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}.$$

$\mathbf{P}$ is a $c \times b$ matrix formed from an identity matrix of order, b, by deleting $b - c$ rows corresponding to parents without a record. $\mathbf{Q}$ is a $d \times b$ matrix with all null elements except the following. For the i^{th} individual .5 is inserted in the i^{th} row of $\mathbf{Q}$ in columns corresponding to its parents in the $\mathbf{a}$ vector. Thus if both parents are present, the row contains two ".5's". If only one parent is present, the row contains one ".5". If neither parent is present, the row is null. Now, of course, $\mathbf{A}$ has order, b, referring to those animals with tested progeny. $Var(\mathbf{e})$ is no longer $\mathbf{I}\sigma_e^2$. It is diagonal with diagonal elements as follows for noninbred animals.

(1) σ_e^2 for parents.

(2) $\sigma_e^2 + .5\ \sigma_a^2$ for progeny with both parents in $\mathbf{a}$.

(3) $\sigma_e^2 + .75\ \sigma_a^2$ for progeny with one parents in $\mathbf{a}$.

(4) $\sigma_e^2 + \sigma_a^2$ for progeny with no parent in $\mathbf{a}$.

This model results in d less equations than in the usual model and a possible large reduction in time required for a solution to the mixed model equations.

Computation of $\hat{a}_i$, BLUP of a tested individual not in the solution for $\mathbf{a}$ but providing data in $\mathbf{y}$, is simple.

$$\hat{e}_i = y_i - \mathbf{x}_i'\boldsymbol{\beta}^o - \mathbf{z}_i'\hat{\mathbf{u}} - .5\ (\text{Sum of parental } \hat{\mathbf{a}}).$$

$\mathbf{x}_i'$ is the incidence matrix for the i^{th} animal with respect to $\boldsymbol{\beta}$.

$\mathbf{z}_i'$ is the incidence matrix for the i^{th} animal with respect to $\hat{\mathbf{u}}$, other random factors in the model. Then

$$\begin{aligned} \hat{a}_i &= .5 \text{ (sum of parental } \hat{\mathbf{a}}) + k_i \hat{e}_i, \\ \text{where } k_i &= .5\ \sigma_a^2/(.5\ \sigma_a^2 + \sigma_e^2) \text{ if both parents known,} \\ &= .75\ \sigma_a^2/(.75\ \sigma_a^2 + \sigma_e^2) \text{ if one parent known,} \\ &= \sigma_a^2/(\sigma_a^2 + \sigma_e^2) \text{ if neither parent known.} \end{aligned}$$

The solution for an animal with no record and no progeny is .5 (sum of parental $\hat{\mathbf{a}}$), provided these parents, if known, are included in the b elements of $\hat{\mathbf{a}}$ in the solution.

A simple sire model for single traits can be considered a special case of this model. The incidence matrix for sires is the same as in Chapter 23 except that it is multipled by .5. The "error" variance is $\mathbf{I}(\sigma_e^2 + .75\ \sigma_a^2)$. The $\mathbf{G}$ submatrix for sires is $\mathbf{A}\sigma_a^2$ rather than $.25\ \sigma_a^2\mathbf{A}$. Then the evaluations from this model for sires are exactly twice those of Chapter 23.

A sire model containing sires of the mates but not the mates' records can be formulated by the gametic model. Then $\mathbf{a}$ would include both sires and grandsires. The incidence matrix for a progeny would contain elements .5 associated with sire and .25 associated with grandsire. Then the "error" variance would contain $\sigma_e^2 + .6875\ \sigma_a^2$, $\sigma_e^2 + .75\ \sigma_a^2$, or $\sigma_e^2 + .9375\ \sigma_a^2$ for progeny with both sire and grandsire, sire only, or grandsire only respectively.

We illustrate the methods of this section with a very simple example. Animals $1, \ldots, 4$ have records (5,3,2,8). $\mathbf{X}' = (1\ 2\ 1\ 3)$. Animals 1 and 2 are the parents of 3, and animal 1 is the parent of 4. The error variance is $\sigma_e^2 = 10$ and $\sigma_a^2 = 4$. We first treat this as an individual animal model where

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & .5 & .5 \\ & 1 & .5 & 0 \\ & & 1 & .25 \\ & & & 1 \end{pmatrix}.$$

The mixed model equations are

$$\begin{pmatrix} 1.5 & .1 & .2 & .1 & .3 \\ & .558333 & .125 & -.25 & -.166667 \\ & & .475 & -.25 & 0 \\ & & & .6 & 0 \\ & & & & .433333 \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{u}_4 \end{pmatrix} = \begin{pmatrix} 3.7 \\ .5 \\ .3 \\ .2 \\ .8 \end{pmatrix}. \tag{8}$$

The solution is

$$(2.40096,\ .73264,\ \text{-}.57212,\ .00006,\ .46574).$$

Now in the gametic model the incidence matrix is

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & .5 & .5 \\ 3 & .5 & 0 \end{pmatrix},\ \mathbf{G} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix},\ \mathbf{R} = \text{ dg } (10, 10, 12, 13).$$

$$12 = 10 + .5(4),\ 13 = 10 + .75(4).$$

Then the mixed model equations are

$$\begin{pmatrix} 1.275641 & .257051 & .241667 \\ & .390064 & .020833 \\ & & .370833 \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} = \begin{pmatrix} 3.112821 \\ .891026 \\ .383333 \end{pmatrix}. \quad (9)$$

The solution is

$$(2.40096,\ .73264,\ -.57212). \quad (10)$$

This is the same as the first 3 elements of (28.8).

$$\begin{aligned} \hat{e}_3 &= 2 - 2.40096\ - .5(.73264 - .57212)\ = \ -.48122. \\ \hat{u}_3 &= .5(.73264 - .57212) + 2(-.48122)/12 = .00006. \\ \hat{e}_4 &= 8 - 3(2.40096) - .5(.73264) = .43080. \\ \hat{u}_4 &= .5(.73264) + 3(.43080)/13 = .46574. \end{aligned}$$

$\hat{u}_3, \hat{u}_4$ are the same as in (28.8).

6.2 Repeated records model

This section is concerned with multiple records in a single trait and under the assumption that

$$Var \begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & r & r & \cdots \\ r & 1 & r & \cdots \\ r & r & 1 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \sigma_y^2,$$

where $\mathbf{y}$ has been adjusted for random factors other than producing ability and random error. The subscript i refers to a particular animal. The model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_p\mathbf{p}\ +\ \text{possibly other random factors}\ +\ \mathbf{e}.$$

$$Var\begin{pmatrix}\mathbf{a}\\ \mathbf{p}\\ \mathbf{e}\end{pmatrix} = \begin{pmatrix}\mathbf{A}\sigma_a^2 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\sigma_p^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2\end{pmatrix}.$$

In an unselected population $\sigma_a^2 = h^2\sigma_y^2, \sigma_p^2 = (r - h^2)\sigma_y^2$, $\sigma_e^2 = (1 - r)\sigma_y^2$, after adjusting $\mathbf{y}$ for other random factors. As before b animals have progeny; $c \leq b$ of these have records. These records number n_1. Also as before d animals with records have no progeny. The number of records made by these animals is n_2.

First we state the model as described in Chapter 24. $\mathbf{X}$, $\mathbf{Z}_a$, $\mathbf{Z}_b$ all have $n_1 + n_2$ rows. The number of elements in $\mathbf{a}$ is $b + d$. The $\mathbf{Z}_a$ matrix is the same as in the conventional model of Section 28.6.1 except that the row pertaining to an individual with records is repeated as many times as there are records on that animal. The number of elements in $\mathbf{p}$ is $c+d$ corresponding to these animals with records. $\mathbf{Z}_p$ would be an identity matrix with order $c + d$ if the $c + d$ animals with records had made only one record each. Then the row of this matrix corresponding to an animal is repeated as many times as the number of records in that animal. Since $\mathbf{Z}_p'\mathbf{Z}_p' + (\mathbf{I}\sigma_p^2)^{-1}$ is diagonal, $\hat{\mathbf{p}}$ can be "absorbed" easily to reduce the number of equations to $b+d$ plus the number of elements in $\boldsymbol{\beta}$. The predicted real producing ability of the i^{th} animal is $\hat{a}_i + \hat{p}_i$, with $\hat{p}_i = 0$ for animals with no records.

Now we state the gametic model for repeated records. As for single records, $\mathbf{a}$ now has b elements corresponding to the b animals with progeny. $\mathbf{Z}_a$ is exactly the same as in the gametic model for single records except that the row pertaining to an animal is repeated as many times as the number of records for that animal. As in the conventional method for repeated records, $\mathbf{p}$ has $c + d$ elements and $\mathbf{Z}_p$ is the same as in that model.

Now Mendelian sampling is taken care of in this model by altering $Var(\mathbf{p})$ rather than $Var(\mathbf{e})$ as was done in the single record gametic model. For the parents $Var(\mathbf{p})$ remains diagonal with the first c diagonals being σ_p^2. The remaining $\mathbf{d}$ have the following possible values.

(1) $\sigma_p^2 + .5\sigma_a^2$ if both parents are in $\mathbf{a}$,

(2) $\sigma_p^2 + .75\sigma_a^2$ if one parent is in $\mathbf{a}$,

(3) $\sigma_p^2 + \sigma_a^2$ if no parent is in $\mathbf{a}$.

Again we can absorb "p" to obtain a set of equations numbering b plus the number of elements in $\boldsymbol{\beta}$, a reduction of c from the conventional equations. The computation of $\hat{a}$ for the d animals with no progeny is simple.

$$\begin{aligned}\hat{a}_i &= .5(\text{sum of parental } \hat{\mathbf{a}}) + k_i\hat{p}_i.\\ \text{where } k_i &= .5\ \sigma_a^2/(\sigma_p^2 + .5\sigma_a^2) \text{ for animals with 2 parents in } \mathbf{a}.\\ &= .75\sigma_a^2/(\sigma_p^2 + .75\sigma_a^2) \text{ for those with one parent in } \mathbf{a}.\\ &= \sigma_a^2/(\sigma_p^2 + \sigma_a^2)\ \text{ for those with no parent in } \mathbf{a}.\end{aligned}$$

These two methods for repeated records are illustrated with the same animals as in Section 28.8 except now there are repeated records. The 4 animals have 2,3,1,2 records respectively. These are (5,3,4,2,3,6,7,8). $\mathbf{X}' = (1\ 2\ 3\ 1\ 2\ 2\ 3\ 2)$. Let

$$\begin{aligned}\sigma_a^2 &= .25, \\ \sigma_p^2 &= .20, \\ \sigma_e^2 &= .55.\end{aligned}$$

Then the regular mixed model equations are in (28.11).

$$\begin{pmatrix} 65.455 & 5.455 & 10.909 & 3.636 & 9.091 & 5.455 & 10.909 & 3.636 & 9.091 \\ & 10.970 & 2.0 & -4.0 & -2.667 & 3.636 & 0 & 0 & 0 \\ & & 11.455 & -4.0 & 0 & 0 & 5.455 & 0 & 0 \\ & & & 9.818 & 0 & 0 & 0 & 1.818 & 0 \\ & & & & 8.970 & 0 & 0 & 0 & 3.636 \\ & & & & & 8.636 & 0 & 0 & 0 \\ & & & & & & 10.455 & 0 & 0 \\ & & & & & & & 6.818 & 0 \\ & & & & & & & & 8.636 \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} 145.455 \\ 14.546 \\ 16.364 \\ 10.909 \\ 27.273 \\ 14.546 \\ 16.364 \\ 10.909 \\ 27.273 \end{pmatrix}. \quad (11)$$

Note that the right hand sides for $\hat{\mathbf{a}}$ and $\hat{\mathbf{p}}$ are identical. The solution is

$$(1.9467,\ .8158,\ -.1972,\ .5660,\ 1.0377,\ .1113,\ -.3632,\ .4108,\ .6718). \quad (12)$$

Next the solution for the gametic model is illustrated with the same data. The incidence matrix is

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 & 0 \\ 2 & .5 & .5 & 0 & 0 & 1 & 0 \\ 3 & .5 & 0 & 0 & 0 & 0 & 1 \\ 2 & .5 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

corresponding to $\beta, a_1, a_2, p_1, p_2, p_3, p_4$.

$$\begin{aligned}
Var(\mathbf{e}) &= .55\ \mathbf{I}, \\
Var(\mathbf{a}) &= \begin{pmatrix} .25 & 0 \\ 0 & .25 \end{pmatrix}, \\
Var(\mathbf{p}) &= \text{diag}\ (.2,\ .2,\ .325,\ .3875).
\end{aligned}$$

Then the mixed model equations are

$$\begin{pmatrix}
65.454 & 11.818 & 12.727 & 5.454 & 10.909 & 3.636 & 9.091 \\
 & 9.0 & .454 & 3.636 & 0 & .909 & 1.818 \\
 & & 9.909 & 0 & 5.454 & .909 & 0 \\
 & & & 8.636 & 0 & 0 & 0 \\
 & & & & 10.454 & 0 & 0 \\
 & & & & & 4.895 & 0 \\
 & & & & & & 6.217
\end{pmatrix}
\begin{pmatrix} \hat{\beta} \\ a_1 \\ a_2 \\ p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix}
=
\begin{pmatrix} 145.454 \\ 33.636 \\ 21.818 \\ 14.546 \\ 16.364 \\ 10.909 \\ 27.273 \end{pmatrix} \tag{13}$$

Note that the coefficient submatrix for $\hat{\mathbf{p}}$ is diagonal. The solution is

$$(1.9467,\ .8158,\ -.1972,\ .1113,\ -.3632,\ .6676,\ 1.3017). \tag{14}$$

Note that $\hat{\beta}, \hat{a}_1,\ \hat{a}_2,\ \hat{p}_1,\ \hat{p}_2$ are the same as in (28.12). Now

$$\begin{aligned}
\hat{a}_3 &= .5(.8158\ - .1972) + .125\ (.6676)/.325 = .5660. \\
\hat{a}_4 &= .5(.8158) + .1875(1.3017)/.3875 = 1.0377.
\end{aligned}$$

These are the same results for $\hat{a}_3$ and $\hat{a}_3$ as (28.12).

Chapter 29
Non-Additive Genetic Merit

C. R. Henderson

1984 - Guelph

1 Model for Genetic Components

All of the applications in previous chapters have been concerned entirely with additive genetic models. This may be a suitable approximation, but theory exists that enables consideration to be given to more complicated genetic models. This theory is simple for non-inbred populations, for then we can formulate genetic merit of the animals in a sample as

$$\mathbf{g} = \sum_i \mathbf{g}_i.$$

$\mathbf{g}$ is the vector of total genetic values for the animals in the sample. $\mathbf{g}_i$ is a vector describing values for a specific type of genetic merit. For example, $\mathbf{g}_1$ represents additive values, $\mathbf{g}_2$ dominance values, $\mathbf{g}_3$ additive $\times$ additive, $\mathbf{g}_4$ additive by dominance, etc. In a non-inbred, unselected population and ignoring linkage

$$Cov(\mathbf{g}_i, \mathbf{g}_j') = \mathbf{0}$$

for all pairs of $i \neq j$.

$$Var(\text{additive}) = \mathbf{A}\sigma_a^2,$$

$$Var(\text{dominance}) = \mathbf{D}\sigma_d^2,$$

$$Var(\text{additive} \times \text{additive}) = \mathbf{A}\#\mathbf{A}\sigma_{aa}^2,$$

$$Var(\text{additive} \times \text{dominance}) = \mathbf{A}\#\mathbf{D}\sigma_{ad}^2,$$

$$Var(\text{additive} \times \text{additive} \times \text{dominance}) = \mathbf{A}\#\mathbf{A}\#\mathbf{D}\sigma_{aad}^2, \text{ etc.}$$

The # operation on $\mathbf{A}$ and $\mathbf{D}$ is described below. These results are due mostly to Cockerham (1954). $\mathbf{D}$ is computed as follows. All diagonals are 1. $d_{km} (k \neq m)$ is computed from certain elements of $\mathbf{A}$. Let the parents of k and m be g, h and i, j respectively. Then

$$d_{km} = .25(a_{gi}a_{hj} + a_{gj}a_{hi}). \tag{1}$$

In a non-inbred population only one at most of the products in this expression can be greater than 0. To illustrate suppose k and m are full sibs. Then $g = i$ and $h = j$.

Consequently

$$d_{km} = .25[(1)(1) + 0] = .25.$$

Suppose k and m are double first cousins. Then

$$d_{km} = .25[(.5)(.5) + 0] = .0625.$$

For non-inbred paternal sibs from unrelated dams is

$$d_{km} = .25[1(0) + 0(0)] = 0,$$

and for parent-progeny $d_{km} = 0$.

The # operation on two matrices means that the new matrix is formed from the products of the corresponding elements of the 2 matrices. Thus the ij^{th} element of $\mathbf{A}\#\mathbf{A}$ is a_{ij}^2, and the ij^{th} element of $\mathbf{A}\#\mathbf{D}$ is $a_{ij}d_{ij}$. These are called Hadamard products. Accordingly, we see that all matrices for $Var(\mathbf{g}_i)$ are derived from $\mathbf{A}$.

2 Single Record on Every Animal

We shall describe BLUP procedures and estimation of variances in this and subsequent sections of Chapter 29 by a model with additive and dominance components. The extension to more components is straightforward. The model for $\mathbf{y}$ with no data missing is

$$\mathbf{y} = (\mathbf{X}\ \ \mathbf{I}\ \ \mathbf{I}) \begin{pmatrix} \boldsymbol{\beta} \\ \mathbf{a} \\ \mathbf{d} \end{pmatrix} + \mathbf{e}\,.$$

$\mathbf{y}$ is n x 1, $\mathbf{X}$ is n x p, both $\mathbf{I}$ are n x n, and $\mathbf{e}$ is n x 1, $\boldsymbol{\beta}$ is p x 1, $\mathbf{a}$ and $\mathbf{d}$ are n x 1.

$$Var(\mathbf{a}) = \mathbf{A}\sigma_a^2,$$

$$Var(\mathbf{d}) = \mathbf{D}\sigma_d^2,$$

$$Var(\mathbf{e}) = \mathbf{I}\sigma_e^2.$$

$Cov(\mathbf{a}, \mathbf{d}')$, $Cov(\mathbf{a}, \mathbf{e}')$, and $Cov(\mathbf{d}, \mathbf{e}')$ are all $n \times n$ null matrices. Now the mixed model equations are

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}' & \mathbf{X}' \\ \mathbf{X} & \mathbf{I} + \mathbf{A}^{-1}\sigma_e^2/\sigma_a^2 & \mathbf{I} \\ \mathbf{X} & \mathbf{I} & \mathbf{I} + \mathbf{D}^{-1}\sigma_e^2/\sigma_d^2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{a}} \\ \hat{\mathbf{d}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{y} \\ \mathbf{y} \end{pmatrix}. \tag{2}$$

Note that if $\mathbf{a}, \mathbf{d}$ were regarded as fixed, the last n equations would be identical to the $p+1,\ \ldots,\ p+n$ equations, and we could estimate only differences among elements of

$\mathbf{a}+\mathbf{d}$. An interesting relationship exists between $\hat{\mathbf{a}}$ and $\hat{\mathbf{d}}$. Subtracting the third equation of (29.2) from the second,

$$\mathbf{A}^{-1}\sigma_e^2/\sigma_a^2\hat{\mathbf{a}} - \mathbf{D}^{-1}\sigma_e^2/\sigma_d^2\hat{\mathbf{d}} = \mathbf{0}.$$

Therefore

$$\hat{\mathbf{d}} = \mathbf{D}\mathbf{A}^{-1}\sigma_d^2/\sigma_a^2\hat{\mathbf{a}}. \tag{3}$$

This identity can be used to reduce (29.2) to

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'(\mathbf{I}+\mathbf{D}\mathbf{A}^{-1}\sigma_d^2/\sigma_a^2) \\ \mathbf{X} & \mathbf{I}+\mathbf{A}^{-1}\sigma_e^2/\sigma_a^2+\mathbf{D}\mathbf{A}^{-1}\sigma_d^2/\sigma_a^2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{y} \end{pmatrix}. \tag{4}$$

Note that the coefficient matrix of (29.4) is not symmetric. Having solved for $\hat{\mathbf{a}}$ in (29.4) compute $\hat{\mathbf{d}}$ by (29.3).

σ_a^2, σ_d^2, σ_e^2 can be estimated by MIVQUE. Quadratics needed to be computed and equated to their expectations are

$$\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}},\ \hat{\mathbf{d}}'\mathbf{D}^{-1}\hat{\mathbf{d}},\ \text{and } \hat{\mathbf{e}}'\hat{\mathbf{e}}. \tag{5}$$

To obtain expectations of the first two of these we need $Var(\mathbf{r})$, where $\mathbf{r}$ is the vector of right hand sides of (29.2). This is

$$\begin{pmatrix} \mathbf{X}'\mathbf{A}\mathbf{X} & \mathbf{X}'\mathbf{A} & \mathbf{X}'\mathbf{A} \\ \mathbf{A}\mathbf{X} & \mathbf{A} & \mathbf{A} \\ \mathbf{A}\mathbf{X} & \mathbf{A} & \mathbf{A} \end{pmatrix}\sigma_a^2 + \begin{pmatrix} \mathbf{X}'\mathbf{D}\mathbf{X} & \mathbf{X}'\mathbf{D} & \mathbf{X}'\mathbf{D} \\ \mathbf{D}\mathbf{X} & \mathbf{D} & \mathbf{D} \\ \mathbf{D}\mathbf{X} & \mathbf{D} & \mathbf{D} \end{pmatrix}\sigma_d^2 + \begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}' & \mathbf{X}' \\ \mathbf{X} & \mathbf{I} & \mathbf{I} \\ \mathbf{X} & \mathbf{I} & \mathbf{I} \end{pmatrix}\sigma_e^2. \tag{6}$$

From (29.6) we can compute $Var(\hat{\mathbf{a}})$, $Var(\hat{\mathbf{d}})$ as follows. Let some g-inverse of the matrix of (29.2) be

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_\beta \\ \mathbf{C}_a \\ \mathbf{C}_d \end{pmatrix}.$$

$\mathbf{C}_a$ and $\mathbf{C}_d$ each have n rows. Then

$$\hat{\mathbf{a}} = \mathbf{C}_a\mathbf{r},$$

and

$$\hat{\mathbf{d}} = \mathbf{C}_d\mathbf{r}.$$

$$Var(\hat{\mathbf{a}}) = \mathbf{C}_a\,Var(\mathbf{r})\mathbf{C}_a', \tag{7}$$

and

$$Var(\hat{\mathbf{d}}) = \mathbf{C}_d\,Var(\mathbf{r})\mathbf{C}_d'. \tag{8}$$

$$E(\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}}) = tr\mathbf{A}^{-1}\,Var(\hat{\mathbf{a}}). \tag{9}$$

$$E(\hat{\mathbf{d}}'\mathbf{D}^{-1}\hat{\mathbf{d}}) = tr\mathbf{D}^{-1}\,Var(\hat{\mathbf{d}}). \tag{10}$$

For the expectation of $\hat{\mathbf{e}}'\hat{\mathbf{e}}$ we compute $tr(Var(\hat{\mathbf{e}}))$. Note that

$$\begin{aligned}
\hat{\mathbf{e}} &= \left(\mathbf{I} - (\mathbf{X}\ \ \mathbf{I}\ \ \mathbf{I})\ \mathbf{C}\begin{pmatrix}\mathbf{X}'\\ \mathbf{I}\\ \mathbf{I}\end{pmatrix}\right)\mathbf{y}\\
&= (\mathbf{I} - \mathbf{X}\mathbf{C}_{11}\mathbf{X}' - \mathbf{X}\mathbf{C}_{12} - \mathbf{C}_{12}'\mathbf{X}' - \mathbf{X}\mathbf{C}_{13} - \mathbf{C}_{13}'\mathbf{X}'\\
&\quad -\mathbf{C}_{22} - \mathbf{C}_{23} - \mathbf{C}_{23}' - \mathbf{C}_{33})\mathbf{y}\\
&\equiv \mathbf{T}\mathbf{y}
\end{aligned} \tag{11}$$

where

$$\mathbf{C} = \begin{pmatrix}\mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13}\\ \mathbf{C}_{12}' & \mathbf{C}_{22} & \mathbf{C}_{23}\\ \mathbf{C}_{13}' & \mathbf{C}_{23}' & \mathbf{C}_{33}\end{pmatrix}. \tag{12}$$

Then

$$Var(\hat{\mathbf{e}}) = \mathbf{T}\,Var(\mathbf{y})\mathbf{T}'. \tag{13}$$

$$Var(\mathbf{y}) = \mathbf{A}\sigma_a^2 + \mathbf{D}\sigma_d^2 + \mathbf{I}\sigma_e^2. \tag{14}$$

REML by the EM type algorithm is quite simple to state. At each round of iteration we need the same quadratics as in (29.5). Now we pretend that $Var(\hat{\mathbf{a}})$, $Var(\hat{\mathbf{d}})$, $Var(\hat{\mathbf{e}})$ are represented by the mixed model result with true variance ratios employed. These are

$$\begin{aligned}
Var(\hat{\mathbf{a}}) &= \mathbf{A}\sigma_a^2 - \mathbf{C}_{22}.\\
Var(\hat{\mathbf{d}}) &= \mathbf{D}\sigma_d^2 - \mathbf{C}_{33}.\\
Var(\hat{\mathbf{e}}) &= \mathbf{I}\sigma_e^2 - \mathbf{W}\mathbf{C}\mathbf{W}'.
\end{aligned}$$

$\mathbf{C}_{22}$, $\mathbf{C}_{33}$, $\mathbf{C}$ are defined in (29.12).

$$\mathbf{W} = (\mathbf{X}\ \ \mathbf{I}\ \ \mathbf{I}).$$

$\mathbf{W}\mathbf{C}\mathbf{W}'$ can be written as $\mathbf{I} - \mathbf{T} = \mathbf{X}\mathbf{C}_{11}\mathbf{X}' + \mathbf{X}\mathbf{C}_{12} +$ etc. From these "variances" we iterate on

$$\hat{\sigma}_a^2 = (\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}} + tr\mathbf{A}^{-1}\mathbf{C}_{22})/n, \tag{15}$$

$$\hat{\sigma}_d^2 = (\hat{\mathbf{d}}'\mathbf{D}^{-1}\hat{\mathbf{d}} + tr\mathbf{D}^{-1}\mathbf{C}_{33})/n, \tag{16}$$

and

$$\hat{\sigma}_e^2 = (\hat{\mathbf{e}}'\hat{\mathbf{e}} + tr\mathbf{W}\mathbf{C}\mathbf{W}')/n. \tag{17}$$

This algorithm guarantees that at each round of iteration all estimates are non-negative provided the starting values of σ_e^2/σ_a^2, σ_e^2/σ_d^2 are positive.

3 Single or No Record on Each Animal

In this section we use the same model as in Section 29.2, except now some animals have no record but we wish to evaluate them in the mixed model solution. Let us order the animals by the set of animals with no record followed by the set with records.

$$\mathbf{y} = (\mathbf{X}\ \ \mathbf{0}\ \ \mathbf{I}\ \ \mathbf{0}\ \ \mathbf{I}) \begin{pmatrix} \boldsymbol{\beta} \\ \mathbf{a}_m \\ \mathbf{a}_p \\ \mathbf{d}_m \\ \mathbf{d}_p \end{pmatrix} + \mathbf{e}\,. \tag{18}$$

The subscript, m, denotes animals with no record, and the subscript, p, denotes animals with a record. Let there be n_p animals with a record and n_m animals with no record. Then $\mathbf{y}$ is n_p x 1, $\mathbf{X}$ is n_p x p, the $\mathbf{0}$ submatrices are both n_p x n_m, and the $\mathbf{I}$ submatrices are both n_p x n_p. The OLS equations are

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{0} & \mathbf{X}' & \mathbf{0} & \mathbf{X}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{X} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{X}' & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{a}}_m \\ \hat{\mathbf{a}}_p \\ \hat{\mathbf{d}}_m \\ \hat{\mathbf{d}}_p \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{0} \\ \mathbf{y} \\ \mathbf{0} \\ \mathbf{y} \end{pmatrix}. \tag{19}$$

The mixed model equations are formed by adding $\mathbf{A}^{-1}\sigma_e^2/\sigma_a^2$ and $\mathbf{D}^{-1}\sigma_e^2/\sigma_d^2$ to the appropriate submatrices of matrix (29.19).

We illustrate these equations with a simple example. We have 10 animals with animals 1,3,5,7 not having records. 1,2,3,4 are unrelated, non-inbred animals. The parents of 5 and 6 are 1,2. The parents of 7 and 8 are 3,4. The parents of 9 are 6,7. The parents of 10 are 5,8. This gives

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & .5 & .5 & 0 & 0 & .25 & .25 \\ & 1 & 0 & 0 & .5 & .5 & 0 & 0 & .25 & .25 \\ & & 1 & 0 & 0 & 0 & .5 & .5 & .25 & .25 \\ & & & 1 & 0 & 0 & .5 & .5 & .25 & .25 \\ & & & & 1 & .5 & 0 & 0 & .25 & .5 \\ & & & & & 1 & 0 & 0 & .5 & .25 \\ & & & & & & 1 & .5 & .5 & .25 \\ & & & & & & & 1 & .25 & .5 \\ & & & & & & & & 1 & .25 \\ & & & & & & & & & 1 \end{pmatrix}.$$

$\mathbf{D}$ = matrix with all 1's in diagonal,

$$d_{56} = d_{65} = d_{78} = d_{87} = .25,$$

$$d_{9,10} = d_{10,9} = .0625,$$

and all other elements = 0.

$$\mathbf{y}' = [6, 9, 6, 7, 4, 6].$$

$$\mathbf{X}' = (1\ 1\ 1\ 1\ 1\ 1).$$

Assuming that $\sigma_e^2/\sigma_a^2 = 2.25$ and $\sigma_e^2/\sigma_d^2 = 5$, the mixed model coefficient matrix with animals ordered as in the **A** and **D** matrices is in (29.20) . . . (29.22). The right hand side vector is [38, 0, 6, 0, 9, 0, 6, 0, 7, 4, 6, 0, 6, 0, 9, 0, 6, 0, 7, 4, 6]′. The solution is

$$\boldsymbol{\beta}^o = 6.400,$$

$$\hat{\mathbf{a}}' = [-.203,\ -.256,\ -.141,\ .600,\ -.259,\ -.403,\ .056,\ .262,\ -.521,\ -.058],$$

and

$$\hat{\mathbf{d}}' = (0,\ -.024,\ 0,\ .333,\ 0,\ 0,\ .014,\ .056,\ -.316,\ -.073).$$

Upper left 11 × 11

$$\left(\begin{array}{rrrrrrrrrrr}
6. & 0 & 1. & 0 & 1. & 0 & 1. & 0 & 1. & 1. & 1. \\
 & 4.5 & 2.25 & 0 & 0 & -2.25 & -2.25 & 0 & 0 & 0 & 0 \\
 & & 5.5 & 0 & 0 & -2.25 & -2.25 & 0 & 0 & 0 & 0 \\
 & & & 4.5 & 2.25 & 0 & 0 & -2.25 & -2.25 & 0 & 0 \\
 & & & & 5.5 & 0 & 0 & -2.25 & -2.25 & 0 & 0 \\
 & & & & & 5.625 & 0 & 0 & 1.125 & 0 & -2.25 \\
 & & & & & & 6.625 & 1.125 & 0 & -2.25 & 0 \\
 & & & & & & & 5.625 & 0 & -2.25 & 0 \\
 & & & & & & & & 6.625 & 0 & -2.25 \\
 & & & & & & & & & 5.5 & 0 \\
 & & & & & & & & & & 5.5
\end{array}\right). \quad (20)$$

Upper right 10 × 10 and (lower left 10 × 11)′

$$\left(\begin{array}{cccccccccc}
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right). \quad (21)$$

Lower right 10 × 10

$$\begin{pmatrix} 5.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 6.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 5.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 6.0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & 5.333 & -1.333 & 0 & 0 & 0 & 0 \\ & & & & & 6.333 & 0 & 0 & 0 & 0 \\ & & & & & & 5.333 & -1.333 & 0 & 0 \\ & & & & & & & 6.333 & 0 & 0 \\ & & & & & & & & 6.02 & -.314 \\ & & & & & & & & & 6.02 \end{pmatrix}. \quad (22)$$

If we wish EM type estimation of variances we iterate on

$$\begin{aligned} \hat{\sigma}_e^2 &= (\mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta}^o - \mathbf{y}'\hat{\mathbf{a}}_p - \mathbf{y}'\hat{\mathbf{d}}_p)/[n - \text{rank}\ (\mathbf{X})], \\ \hat{\sigma}_a^2 &= (\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}} + tr\hat{\sigma}_e^2\mathbf{C}_{aa})/n, \end{aligned}$$

and

$$\hat{\sigma}_d^2 = (\hat{\mathbf{d}}'\mathbf{D}^{-1}\hat{\mathbf{d}} + tr\hat{\sigma}_e^2\mathbf{C}_{dd})/n,$$

for

$$\hat{\mathbf{a}}' = (\hat{\mathbf{a}}_m' \;\; \hat{\mathbf{a}}_p'),$$

$$\hat{\mathbf{d}}' = (\hat{\mathbf{d}}_m' \;\; \hat{\mathbf{d}}_p'),$$

and n = number of animals. A g-inverse of (29.19) is

$$\begin{pmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xa} & \mathbf{C}_{xd} \\ \mathbf{C}_{xa}' & \mathbf{C}_{aa} & \mathbf{C}_{ad} \\ \mathbf{C}_{xd}' & \mathbf{C}_{ad}' & \mathbf{C}_{dd} \end{pmatrix}.$$

Remember that in these computations $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$ and the equations are set up with scaling, $Var(\mathbf{e}) = \mathbf{I}$, $Var(\mathbf{a}) = \mathbf{A}\sigma_a^2/\sigma_e^2$, $Var(\mathbf{d}) = \mathbf{D}\sigma_d^2/\sigma_e^2$.

4 A Reduced Set of Equations

When there are several genetic components in the model, a much more efficient computing strategy can be employed than that of Section 29.3. Let $\mathbf{m}$ be total genetic value of the members of a population, and this is

$$\mathbf{m} = \sum_i \mathbf{g}_i,$$

where $\mathbf{g}_i$ is the merit for a particular type of genetic component, additive for example. Then in a non-inbred population and ignoring linkage

$$Var(\mathbf{m}) = \sum_i Var(\mathbf{g}_i)$$

since

$$Cov(\mathbf{g}_i, \mathbf{g}_j') = \mathbf{0}$$

for all $i \neq j$. Then a model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_m\mathbf{m} + \mathbf{e}. \tag{23}$$

We could, if we choose, add a term for other random components. Now mixed model equations for BLUE and BLUP are

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_m \\ \mathbf{Z}_m'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_m'\mathbf{R}^{-1}\mathbf{Z}_m + [Var(\mathbf{m})]^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{m}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}_m'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \tag{24}$$

If we are interested in BLUP of certain genetic components this is simply

$$\hat{\mathbf{g}}_i = Var(\mathbf{g}_i)[Var(\mathbf{m})]^{-1}\hat{\mathbf{m}}. \tag{25}$$

This method is illustrated by the example of Section 29.2. Except for scaling

$$Var(\mathbf{e}) = \mathbf{I},$$

$$Var(\mathbf{a}) = 2.25^{-1}\mathbf{A},$$

$$Var(\mathbf{d}) = 5^{-1}\mathbf{D}.$$

Then

$$Var(\mathbf{m}) = 2.25^{-1}\mathbf{A} + 5^{-1}\mathbf{D}$$

$$= \begin{pmatrix} .6444 & 0 & 0 & 0 & .2222 & .2222 & 0 & 0 & .1111 & .1111 \\ & .6444 & 0 & 0 & .2222 & .2222 & 0 & 0 & .1111 & .1111 \\ & & .6444 & 0 & 0 & 0 & .2222 & .2222 & .1111 & .1111 \\ & & & .6444 & 0 & 0 & .2222 & .2222 & .1111 & .1111 \\ & & & & .6444 & .2722 & 0 & 0 & .1111 & .2222 \\ & & & & & .6444 & 0 & 0 & .2222 & .1111 \\ & & & & & & .6444 & .2722 & .2222 & .1111 \\ & & & & & & & .6444 & .1111 & .2222 \\ & & & & & & & & .6444 & .1236 \\ & & & & & & & & & .6444 \end{pmatrix}. \tag{26}$$

Adding the inverse of this to the lower 10 $\times$ 10 block of the OLS equations of (29.27) we obtain the mixed model equations. The OLS equations including animals with missing records are

$$
\begin{pmatrix}
6 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\
 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 0 & 0 & 0 & 0 \\
 & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 0 \\
 & & & & & & & & & & 1
\end{pmatrix}
\begin{pmatrix} \hat{\mu} \\ \hat{\mathbf{m}} \end{pmatrix}
=
\begin{pmatrix} 38 \\ 0 \\ 6 \\ 0 \\ 9 \\ 0 \\ 6 \\ 0 \\ 7 \\ 4 \\ 6 \end{pmatrix}. \tag{27}
$$

The resulting solution is

$$\hat{\mu} = 6.400 \text{ as before, and}$$

$$\hat{\mathbf{m}} = [-.203,\ -.280,\ -.141,\ .933,\ -.259,\ -.402,\ .070,\ .319,\ -.837,\ -.131]'.$$

From $\hat{\mathbf{m}}$ and using the method of (29.25) we obtain the same solution to $\hat{\mathbf{a}}$ and $\hat{\mathbf{d}}$ as before.

To obtain REML estimates identical to those of Section 29.4 compute the same quantities except $\hat{\sigma}_e^2$ can be computed by

$$\hat{\sigma}_e^2 = (\mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta}^o - \mathbf{y}'\mathbf{Z}_m\hat{\mathbf{m}})/[n - \text{rank } (\mathbf{X})].$$

Then $\hat{\mathbf{a}}$ and $\hat{\mathbf{d}}$ are computed from $\hat{\mathbf{m}}$ as described in this section. With the scaling done

$$\mathbf{G} = \mathbf{A}\sigma_a^2/\sigma_e^2 + \mathbf{D}\sigma_d^2/\sigma_e^2,$$

$$\mathbf{C}_{aa} = (\sigma_a^2/\sigma_e^2)\mathbf{A} - (\sigma_a^2/\sigma_e^2)\mathbf{A}\mathbf{G}^{-1}(\mathbf{G} - \mathbf{C}_{mm})\mathbf{G}^{-1}\mathbf{A}(\sigma_a^2/\sigma_e^2),$$

$$\mathbf{C}_{dd} = (\sigma_d^2/\sigma_e^2)\mathbf{D} - (\sigma_d^2/\sigma_e^2)\mathbf{D}\mathbf{G}^{-1}(\mathbf{G} - \mathbf{C}_{mm})\mathbf{G}^{-1}\mathbf{D}(\sigma_d^2/\sigma_e^2),$$

where a g-inverse of the reduced coefficient matrix is

$$\begin{pmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xm} \\ \mathbf{C}'_{xm} & \mathbf{C}_{mm} \end{pmatrix}.$$

In our example $\mathbf{C}_{aa}$ for both the extended and the reduced equations is

$$
\begin{pmatrix}
.4179 & -.0001 & .0042 & .0224 & .2057 & .1856 & .0112 & .0196 & .0942 & .1062 \\
 & .3651 & .0224 & .0571 & .1847 & .1802 & .0428 & .0590 & .1176 & .1264 \\
 & & .4179 & -.0001 & .0112 & .0196 & .2057 & .1856 & .1062 & .0942 \\
 & & & .3651 & .0428 & .0590 & .1847 & .1802 & .1264 & .1176 \\
 & & & & .4100 & .1862 & .0287 & .0365 & .1108 & .2084 \\
 & & & & & .3653 & .0365 & .0618 & .1953 & .1305 \\
 & & & & & & .4100 & .1862 & .2084 & .1108 \\
 & & & & & & & .3653 & .1305 & .1953 \\
 & & & & & & & & .3859 & .1304 \\
 & & & & & & & & & .3859
\end{pmatrix}.
$$

Similarly $\mathbf{C}_{dd}$ is

$$\begin{pmatrix} .2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & .1786 & 0 & .0034 & .0016 & .0064 & 0 & .0030 & .0045 & .0047 \\ & & .2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & .1786 & .0008 & .0030 & 0 & .0064 & .0047 & .0045 \\ & & & & .1986 & .0444 & 0 & .0007 & .0016 & .0011 \\ & & & & & .1778 & 0 & .0027 & .0062 & .0045 \\ & & & & & & .2 & 0 & 0 & 0 \\ & & & & & & & .1778 & .0047 & .0062 \\ & & & & & & & & .1778 & .0136 \\ & & & & & & & & & .1778 \end{pmatrix}.$$

$\mathbf{C}_{aa}$ and $\mathbf{C}_{dd}$ have rather large rounding errors.

5 Multiple or No Records

Next consider a model with repeated records and the traditional repeatability model. That is, all records have the same variance and all pairs of records on the same animal have the same covariance. Ordering the animals with no records first the model is

$$\mathbf{y} = [\mathbf{X}\ \ \mathbf{0}\ \ \mathbf{Z}\ \ \mathbf{0}\ \ \mathbf{Z}\ \ \mathbf{Z}](\boldsymbol{\beta} : \mathbf{a}_m : \mathbf{a}_p : \mathbf{d}_m : \mathbf{d}_p\ \ \mathbf{t})' + \mathbf{e}. \tag{28}$$

$\mathbf{y}$ is n x 1, $\mathbf{X}$ is n x p, the null matrices are n x n_m, $\mathbf{Z}$ is n x n_p. n is the number of records, n_m the number of animals with no record, and n_p the number of animals with 1 or more records. $\mathbf{a}_m$, $\mathbf{a}_p$ refer to $\mathbf{a}$ for animals with no records and with records respectively, and similarly for $\mathbf{d}_m$ and $\mathbf{d}_p$. $\mathbf{t}$ refers to permanent environmental effects for animals with records.

$$Var(\mathbf{a}) = \mathbf{A}\sigma_a^2,$$
$$Var(\mathbf{d}) = \mathbf{D}\sigma_d^2,$$
$$Var(\mathbf{t}) = \mathbf{I}\sigma_t^2,$$
$$Var(\mathbf{e}) = \mathbf{I}\sigma_e^2.$$

These 4 vectors are uncorrelated. The OLS equations are

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{0} & \mathbf{X}'\mathbf{Z} & \mathbf{0} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{Z} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{Z}'\mathbf{X} & \mathbf{0} & \mathbf{Z}'\mathbf{Z} & \mathbf{0} & \mathbf{Z}'\mathbf{Z} & \mathbf{Z}'\mathbf{Z} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{Z}'\mathbf{X} & \mathbf{0} & \mathbf{Z}'\mathbf{Z} & \mathbf{0} & \mathbf{Z}'\mathbf{Z} & \mathbf{Z}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{0} & \mathbf{Z}'\mathbf{Z} & \mathbf{0} & \mathbf{Z}'\mathbf{Z} & \mathbf{Z}'\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{a}}_m \\ \hat{\mathbf{a}}_p \\ \hat{\mathbf{d}}_m \\ \hat{\mathbf{d}}_p \\ \hat{\mathbf{t}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{0} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{0} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix} \tag{29}$$

The mixed model equations are formed by adding $\mathbf{A}^{-1}\sigma_e^2/\sigma_a^2$, $\mathbf{D}^{-1}\sigma_e^2/\sigma_d^2$, and $\mathbf{I}\sigma_e^2/\sigma_t^2$ to appropriate blocks in (29.29).

We illustrate with the same 10 animals as in the preceding section, but now there are multiple records as follows.

	Records		
Animals	1	2	3
1	X	X	X
2	6	5	4
3	X	X	X
4	9	8	X
5	X	X	X
6	6	5	6
7	X	X	X
8	7	3	X
9	4	5	X
10	6	X	X

X denotes no record. We assume that the first records have a common mean β_1, the second a common mean β_2, and the third a common mean β_3. It is assumed that $\sigma_e^2/\sigma_a^2 = 1.8$, $\sigma_e^2/\sigma_d^2 = 4$, $= \sigma_e^2/\sigma_t^2 = 4$. Then the mixed model coefficient matrix is in (29.30) $\ldots$ (29.32). The right hand side vector is (38, 26, 10, 0, 15, 0, 17, 0, 17, 0, 10, 9, 6, 0, 15, 0, 17, 0, 17, 0, 10, 9, 6, 15, 17, 17, 10, 9, 6)'. The solution is

$$
\begin{aligned}
\boldsymbol{\beta}^{o'} &= (6.398,\ 5.226,\ 5.287), \\
\hat{\mathbf{a}}' &= (-.067,\ -.295,\ -.364,\ .726,\ -.201,\ -.228,\ .048,\ -.051, \\
&\quad -.355,\ -.166), \\
\hat{\mathbf{d}}' &= (0,\ -.103,\ 0,\ .491,\ .019,\ .077,\ -.048,\ -.190,\ -.241, \\
&\quad -.051), \\
\hat{\mathbf{t}}' &= (-.103,\ .491,\ .077,\ -.190,\ -.239,\ -.036).
\end{aligned}
$$

$\hat{\mathbf{t}}$ refers to the six animals with records. BLUP of the others is $\mathbf{0}$.

Upper left 13 × 13

$$
\begin{pmatrix}
6.0 & 0 & 0 & 0 & 1.0 & 0 & 1.0 & 0 & 1.0 & 0 & 1.0 & 1.0 & 1.0 \\
 & 5.0 & 0 & 0 & 1.0 & 0 & 1.0 & 0 & 1.0 & 0 & 1.0 & 1.0 & 0 \\
 & & 2.0 & 0 & 1.0 & 0 & 0 & 0 & 1.0 & 0 & 0 & 0 & 0 \\
 & & & 3.6 & 1.8 & 0 & 0 & -1.8 & -1.8 & 0 & 0 & 0 & 0 \\
 & & & & 6.6 & 0 & 0 & -1.8 & -1.8 & 0 & 0 & 0 & 0 \\
 & & & & & 3.6 & 1.8 & 0 & 0 & -1.8 & -1.8 & 0 & 0 \\
 & & & & & & 5.6 & 0 & 0 & -1.8 & -1.8 & 0 & 0 \\
 & & & & & & & 4.5 & 0 & 0 & .9 & 0 & -1.8 \\
 & & & & & & & & 7.5 & .9 & 0 & -1.8 & 0 \\
 & & & & & & & & & 4.5 & 0 & -1.8 & 0 \\
 & & & & & & & & & & 6.5 & 0 & -1.8 \\
 & & & & & & & & & & & 5.6 & 0 \\
 & & & & & & & & & & & & 4.6
\end{pmatrix}
\tag{30}
$$

Upper right 13 × 16 and (lower left 16 × 13)’

$$
\begin{pmatrix}
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\tag{31}
$$

Lower right 16 $\times$ 16

$$\begin{pmatrix}
4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\
 & & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\
 & & & & 4.267 & -1.067 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 7.267 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 4.267 & -1.067 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 6.267 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\
 & & & & & & & & 6.016 & -.251 & 0 & 0 & 0 & 0 & 2 & 0 \\
 & & & & & & & & & 5.016 & 0 & 0 & 0 & 0 & 0 & 1 \\
 & & & & & & & & & & 7 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & 6 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & 7 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & 6 & 0 & 0 \\
 & & & & & & & & & & & & & & 6 & 0 \\
 & & & & & & & & & & & & & & & 5
\end{pmatrix} \quad (32)$$

6 A Reduced Set of Equations for Multiple Records

As in Section 29.4 we can reduce the equations by now letting

$$\mathbf{m} = \sum_i \mathbf{g}_i + \mathbf{t},$$

where $\mathbf{g}_i$ have the same meaning as before, and $\mathbf{t}$ is permanent environmental effect with $Var(\mathbf{t}) = \mathbf{I}\sigma_t^2$. Then the mixed model equations are like those of (29.24) and from $\hat{\mathbf{m}}$ one can compute $\hat{\mathbf{g}}_i$ and $\hat{\mathbf{t}}$.

Using the same example as in Section 29.5 the OLS equations are

$$\begin{pmatrix}
6 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\
 & 5 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
 & & 2 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & 3 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 2 & 0 & 0 \\
 & & & & & & & & & & & 2 & 0 \\
 & & & & & & & & & & & & 1
\end{pmatrix}
\begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\mu}_3 \\ \hat{\mathbf{m}} \end{pmatrix}
=
\begin{pmatrix} 38 \\ 26 \\ 10 \\ 0 \\ 15 \\ 0 \\ 17 \\ 0 \\ 17 \\ 0 \\ 10 \\ 9 \\ 6 \end{pmatrix}.$$

Now with scaling $Var(\mathbf{e}) = \mathbf{I}$.

$$Var(\mathbf{m}) = 1.8^{-1}\mathbf{A} + .25\mathbf{D} + .25\mathbf{I}$$

$$= \frac{1}{576}\begin{pmatrix}
608 & 0 & 0 & 0 & 160 & 160 & 0 & 0 & 80 & 80 \\
 & 608 & 0 & 0 & 160 & 160 & 0 & 0 & .80 & 80 \\
 & & 608 & 0 & 0 & 0 & 160 & 160 & 80 & 80 \\
 & & & 608 & 0 & 0 & 160 & 160 & 80 & 80 \\
 & & & & 608 & 196 & 0 & 0 & 80 & 160 \\
 & & & & & 608 & 0 & 0 & 160 & 80 \\
 & & & & & & 608 & 196 & 160 & 80 \\
 & & & & & & & 608 & 80 & 160 \\
 & & & & & & & & 608 & 89 \\
 & & & & & & & & & 608
\end{pmatrix}.$$

Adding the inverse of this to the lower 10×10 block of the OLS equations we obtain the mixed model equations. The solution is

$$(\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3) = (6.398,\ 5.226,\ 5.287),$$

the same as before, and

$$\begin{aligned} \hat{\mathbf{m}} \ &= \ (-.067,\ -.500,\ -.364,\ 1.707,\ -.182,\ -.073,\ .001, \\ &\quad\ -.431,\ -.835,\ -.253)'. \end{aligned}$$

Then

$$\hat{\mathbf{a}} = Var(\mathbf{a})[Var(\mathbf{m})]^{-1}\hat{\mathbf{m}} = \text{same as before.}$$

$$\hat{\mathbf{d}} = Var(\mathbf{d})[Var(\mathbf{m})]^{-1}\hat{\mathbf{m}} = \text{same as before.}$$

$$\hat{\mathbf{t}} = Var(\mathbf{t})[Var(\mathbf{m})]^{-1}\hat{\mathbf{m}} = \text{same as before}$$

recognizing that $\hat{t}_i$ for an animal with no record is 0.

To compute EM type REML iterate on

$$\sigma_e^2 = [\mathbf{y}'\mathbf{y} - (\text{soln. vector})'\text{rhs}]/[n - \text{rank}(\mathbf{X})].$$

Compute $\mathbf{C}_{aa}$, $\mathbf{C}_{dd}$, $\mathbf{C}_{tt}$ as in Section 29.4. Now, however, $\mathbf{C}_{tt}$ will have dimension, 10, rather than 6 in order that the matrix of the quadratic in $\hat{\mathbf{t}}$ at each round of iteration will be $\mathbf{I}$. If we did not include missing t_i, a new matrix would need to be computed at each round of iteration.

Chapter 30
Line Cross and Breed Cross Analyses

C. R. Henderson

1984 - Guelph

This chapter is concerned with a genetic model for line crosses, BLUP of crosses, and estimation of variances. It is assumed that a set of unselected inbred lines is derived from some base population. Therefore the lines are assumed to be uncorrelated.

1 Genetic Model

We make the assumption that the total genetic variance of a population can be partitioned into additive + dominance + (additive × additive) + (additive × dominance), etc. Further, in a non-inbred population these different sets of effects are mutually uncorrelated, e.g., Cov (additive, dominance) = **0**. The covariance among sets of effects can be computed from the **A** matrix. Methods for computing **A** are well known. **D** can be computed as described in Chapter 29.

Var(additive effects) $= \mathbf{A}\sigma_a^2$.

Var(dominance effects) $= \mathbf{D}\sigma_d^2$.

Var(additive × dominance) $= \mathbf{A}\#\mathbf{D}\sigma_{ad}^2$.

Var(additive × additive × dominance) $= \mathbf{A}\#\mathbf{A}\#\mathbf{D}\sigma_{aad}^2$, etc.

# denotes the operation of taking the product of corresponding elements of 2 matrices. Thus the ij^{th} element of $\mathbf{A}\#\mathbf{D}$ is $a_{ij}d_{ij}$.

2 Covariances Between Crosses

If lines are unrelated, the progeny resulting from line crosses are non-inbred and consequently the covariance matrices for the different genetic components can be computed for the progeny. Then one can calculate BLUP for these individual animals by the method described in Chapter 29. With animals as contrasted to plants it would seem wise to include a maternal influence of line of dam in the model as described below. Now in order to reduce computational labor we shall make some simplifying assumptions as follows.

1. All members of all lines have inbreeding coefficient $= f$.

2. The lines are large enough that two random individuals from the same line are unrelated except for the fact that they are members of the same line.

Consequently, the **A** matrix for members of the same line is

$$\begin{pmatrix} 1+f & & 2f \\ & \ddots & \\ 2f & & 1+f \end{pmatrix}.$$

From this result we can calculate the covariance between any random pair of individuals from the same cross or a random individual of one cross with a random individual of another cross. We illustrate first with single crosses. Consider line cross, 1×2, line 1 being used as the sire line. Two random progeny pedigrees can be visualized as

$$\begin{array}{ccccccc} 1a & & & & 1b & & \\ & \searrow & & & & \searrow & \\ & & pa & & & & pb \\ & \nearrow & & & & \nearrow & \\ 2a & & & & 2b & & \end{array}$$

Therefore

$$\begin{aligned} a_{1a,1b} &= a_{2a,2b} = 2f. \\ a_{pa,pb} &= .25(2f + 2f) = f. \\ d_{pa,pb} &= .25[2f(2f) + 0(0)] = f^2. \end{aligned}$$

Then the genetic covariance between 2 random members of any single cross is equal to the genetic variance of single cross means

$$= f\sigma_a^2 + f^2\sigma_d^2 + f^2\sigma_{aa}^2 + f^3\sigma_{ad}^2 + \text{ etc.}$$

Note that if $f = 1$, this simplifies to the total genetic variance of individuals in the population from which the lines were derived.

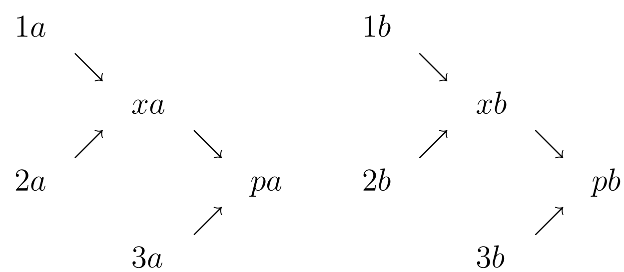
Next consider the covariance between crosses with one common parental line, say 1×2 with 1×3.

$$\begin{array}{ccccccc} 1a & & & & 1b & & \\ & \searrow & & & & \searrow & \\ & & p_a & & & & p_b \\ & \nearrow & & & & \nearrow & \\ 2a & & & & 3b & & \end{array}$$

As before, $a_{1a,1b} = 2f$, but all other relationships among parental pairs are zero. Then

$$\begin{aligned} a_{pa,pb} &= .25(2f) = .5f. \\ d_{pa,pb} &= 0. \\ \text{Covariance} &= .5f\sigma_a^2 + .25f^2\sigma_{aa}^2 + ..., \text{etc.} \end{aligned}$$

Next we consider 3 way crosses. Represent 2 random members of a 3 way cross $(1 \times 2) \times 3$ by



Non-zero additive relationships are

$$\begin{aligned} (1a, 1b) &= (2a, 2b) = (3a, 3b) = 2f, \quad \text{and} \\ (xa, xb) &= f, \\ (pa, pb) &= .25(f + 2f) = .75f, \end{aligned}$$

and the dominance relationship is

$$(pa, pb) = .25[f(2f) + 0(0)] = .5f^2.$$

Thus the genetic variance of 3 way crosses is

$$\frac{3}{4}\,f\sigma_a^2 + \frac{1}{2}\,f^2\sigma_d^2 + \frac{3}{8}\,f^3\sigma_{ad}^2 + \frac{9}{16}\,f^2\sigma_{aa}^2 + \ldots \text{ etc.}$$

The covariance between a single cross and a 3 way cross depends upon the way the crosses are made.

For a $(1 \times 2) \times 3$ with 1×2 is $f/2$, and d is 0.

For a $(1 \times 2) \times 3$ with 1×3 is $.75f$, and d is $.5f^2$.

The variance of 4 way crosses is $.5\ f\sigma_a^2 + .25\ f^2\sigma_d^2 + \ldots$ etc. The variance of top crosses with an inbred line as a parent is $.5\ f\sigma_a^2 + (0)\sigma_d^2 +$ etc.

If we know the magnitude of the various components of genetic variance, we can derive the variance of any line cross or the covariance between any pair of line crosses. Then these can be used to set up mixed model equations. One must be alert to the possibility that some of the variance-covariance matrices of genetic components may be singular.

3 Reciprocal Crosses Assumed Equal

This section is concerned with a model in which the cross, line $i \times$ line j, is assumed the same as the cross, line $j \times$ line i. The model is

$$y_{ijk} = \mathbf{x}'_{ijk}\boldsymbol{\beta} + c_{ij} + e_{ijk}.$$

Var($\mathbf{c}$) has this form

$$\begin{aligned}
Var(c_{ij}) &= f\sigma_a^2 + f^2\sigma_d^2 + f^3\sigma_{ad}^2 + \text{ etc.} \\
&= Cov(c_{ij},\ c_{ji}). \\
Cov(c_{ij}, c_{ij'}) &= Cov(c_{ij}, c_{ji'}) = .5\ f\sigma_a^2 + .25\ f^2\sigma_{aa}^2 + \ldots \text{ etc.} \\
Cov(c_{ij}, c_{i'j'}) &= 0.
\end{aligned}$$

We illustrate BLUP with single crosses among 4 lines with $f = .6$, $\sigma_a^2 = .4$, $\sigma_d^2 = .3$, $\sigma_e^2 = 1$. All other genetic covariances are ignored. $\beta = \mu$. The number of observations per cross and $\overline{y}_{ij}$, are

n_{ij}				$\overline{y}_{ij.}$			
X	5	3	2	X	6	4	7
4	X	6	3	5	X	3	8
4	2	X	5	6	7	X	3
2	3	9	X	5	6	4	X

X denotes no observation. The OLS equations are in (30.1). Note that a_{ij} is combined with a_{ji} to form the variable a_{ij} and similarly for $\mathbf{d}$.

$$\begin{pmatrix}
48 & 9 & 7 & 4 & 8 & 6 & 14 & 9 & 7 & 4 & 8 & 6 & 14 \\
 & 9 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 \\
 & & 7 & 0 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 \\
 & & & 4 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\
 & & & & 8 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\
 & & & & & 6 & 0 & 0 & 0 & 0 & 0 & 6 & 0 \\
 & & & & & & 14 & 0 & 0 & 0 & 0 & 0 & 14 \\
 & & & & & & & 9 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & 7 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 4 & 0 & 0 & 0 \\
 & & & & & & & & & & 8 & 0 & 0 \\
 & & & & & & & & & & & 6 & 0 \\
 & & & & & & & & & & & & 14
\end{pmatrix}
\begin{pmatrix} \mu \\ a_{12} \\ a_{13} \\ a_{14} \\ a_{23} \\ a_{24} \\ a_{34} \\ d_{12} \\ d_{13} \\ d_{14} \\ d_{23} \\ d_{24} \\ d_{34} \end{pmatrix}
=
\begin{pmatrix} 235 \\ 50 \\ 36 \\ 24 \\ 32 \\ 42 \\ 51 \\ 50 \\ 36 \\ 24 \\ 32 \\ 42 \\ 51 \end{pmatrix} \quad (1)$$

$$Var(\mathbf{a}) = \begin{pmatrix} .24 & .12 & .12 & .12 & .12 & 0 \\ & .24 & .12 & .12 & 0 & .12 \\ & & .24 & 0 & .12 & .12 \\ & & & .24 & .12 & .12 \\ & & & & .24 & .12 \\ & & & & & .24 \end{pmatrix}, \; Var(\mathbf{d}) = .108 \, \mathbf{I}.$$

$Var(\mathbf{a})$ is singular. Consequently we pre-multiply equation (30.1) by

$$\begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & Var(\mathbf{a}) & \mathbf{0} \\ \mathbf{0} & 0 & \mathbf{I} \end{pmatrix}$$

and add

$$\begin{pmatrix} 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & [Var(\mathbf{d})]^{-1} \end{pmatrix}$$

to the resulting coefficient matrix. The solution to these equations is

$$\begin{aligned} \hat{\mu} &= 5.1100, \\ \hat{\mathbf{a}}' &= (.5528, \ -.4229, \ .4702, \ -.4702, \ .4229, \ -.5528), \\ \hat{\mathbf{d}}' &= (-.0528, \ .1962, \ .1266, \ -.2965, \ .5769, \ -.5504). \end{aligned}$$

Note that $\sum \hat{d} = 0$. Now the predicted future progeny average of the ij^{th} and ji^{th} cross is

$$\mu^* + \hat{a}_{ij} + \hat{d}_{ij},$$

where μ^* is the fixed part of the model for future progeny.

If we want to predict the future progeny mean of a cross between $i \times k$ or between $k \times i$, where k is not in the sample, we can do this by selection index methods using $\hat{\mathbf{a}}$, $\hat{\mathbf{d}}$ as the "data" with variances and covariances applying to $\mathbf{a}+\mathbf{d}$ rather than $\mathbf{a}$. See Section 5.9. For example the prediction of the 1 × 5 cross is

$$(.12 \;\; .12 \;\; .12 \;\; 0 \;\; 0 \;\; 0) \begin{pmatrix} .348 & .12 & .12 & .12 & .12 & 0 \\ & .348 & .12 & .12 & 0 & .12 \\ & & .348 & 0 & .12 & .12 \\ & & & .348 & .12 & .12 \\ & & & & .348 & .12 \\ & & & & & .348 \end{pmatrix}^{-1} (\hat{\mathbf{a}} + \hat{\mathbf{d}}). \qquad (2)$$

If we were interested only in prediction of crosses among the lines 1 2, 3, 4, we could reduce the mixed model equations to solve for $\hat{\mathbf{a}} + \hat{\mathbf{d}}$ jointly. Then there would be only 7 equations. The 6 × 6 matrix of (30.2) would be $\mathbf{G}^{-1}$ to add to the lower 6 × 6 submatrix of the least squares coefficient matrix.

4 Reciprocal Crosses With Maternal Effects

In most animal breeding models one would assume that because of maternal effects the ij^{th} cross would be different from the ji^{th}. Now the genetic model for maternal effects involves the genetic merit with respect to maternal of the female line in a single cross. This complicates statements of the variances and covariances contributed by different genetic components since the lines are inbred. The statement of σ_a^2 is possible but not the others. The contribution of σ_a^2 is

$$\text{Covariance between 2 progeny of the same cross} \quad = 2f\sigma_a^2 \ ,$$

$$\text{Covariance between progeny of } i \times j \text{ with } k \times j \quad = .5f\sigma_a^2,$$

where the second subscript denotes the female line. Consequently if we ignore other components, we need only to add m_j to the model with $Var(\mathbf{m}) = \mathbf{I}\sigma_m^2$. We illustrate with the same data as in Section 30.3 with $Var(\mathbf{m}) = .5\,\mathbf{I}$. The OLS equations now are in (30.3). Now we pre-multiply these equations by

$$\begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & Var(\mathbf{a}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

Then add to the resulting coefficient matrix

$$\begin{pmatrix} 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & [Var(\mathbf{d})]^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & [Var(\mathbf{m})]^{-1} \end{pmatrix}.$$

The resulting solution is

$$\begin{aligned} \hat{\mu} &= 5.1999, \\ \hat{\mathbf{a}}' &= (.2988,\ -.2413,\ .3217,\ -.3217,\ .2413,\ -.2988), \\ \hat{\mathbf{d}}' &= (-.1737,\ .2307,\ .1136,\ -.1759,\ .4479,\ -.4426), \end{aligned}$$

and

$$\hat{\mathbf{m}}' = (.0560,\ .6920,\ -.8954,\ .1475).$$

$$
\begin{pmatrix}
48 & 9 & 7 & 4 & 8 & 6 & 14 & 9 & 7 & 4 & 8 & 6 & 14 & 10 & 10 & 18 & 10 \\
 & 9 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 4 & 5 & 0 & 0 \\
 & & 7 & 0 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 4 & 0 & 3 & 0 \\
 & & & 4 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 2 & 0 & 0 & 2 \\
 & & & & 8 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 2 & 6 & 0 \\
 & & & & & 6 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 3 & 0 & 3 \\
 & & & & & & 14 & 0 & 0 & 0 & 0 & 0 & 14 & 0 & 0 & 9 & 5 \\
 & & & & & & & 9 & 0 & 0 & 0 & 0 & 0 & 4 & 5 & 0 & 0 \\
 & & & & & & & & 7 & 0 & 0 & 0 & 0 & 4 & 0 & 3 & 0 \\
 & & & & & & & & & 4 & 0 & 0 & 0 & 2 & 0 & 0 & 2 \\
 & & & & & & & & & & 8 & 0 & 0 & 0 & 2 & 6 & 0 \\
 & & & & & & & & & & & 6 & 0 & 0 & 3 & 0 & 3 \\
 & & & & & & & & & & & & 14 & 0 & 0 & 9 & 5 \\
 & & & & & & & & & & & & & 10 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & & 10 & 0 & 0 \\
 & & & & & & & & & & & & & & & 18 & 0 \\
 & & & & & & & & & & & & & & & & 10
\end{pmatrix}
\begin{pmatrix} \hat{\mu} \\ \mathbf{a} \\ \mathbf{d} \\ \mathbf{m} \end{pmatrix}
$$

$$
= (235, 50, 36, 24, 32, 42, 51, 50, 36, 24, 32, 42, 51, 54, 62, 66, 53)'. \qquad (3)
$$

5 Single Crosses As The Maternal Parent

If single crosses are used as the maternal parent in crossing, we can utilize various components of genetic variation with respect to maternal effects, for then the maternal parents are non-inbred.

6 Breed Crosses

If one set of breeds is used as males and a second different set is used as females in a breed cross, the problem is the same as for any two way fixed cross-classified design with interaction and possible missing subclasses. If there is no missing subclass, the weighted squares of means analysis would seem appropriate, but with small numbers of progeny per cross, $\overline{y}_{ij}$ may not be the optimum criterion for choosing the best cross. Rather, we might choose to treat the interaction vector as a pseudo-random variable and proceed to a biased estimation that might well have smaller mean squared error than the $\overline{y}_{ij}$. If subclasses are missing, this biased procedure enables finding a biased estimator of such crosses.

7 Same Breeds Used As Sires And Dams

If the same breeds are used as sires and as dams and with progeny of some or all of the pure breeds included in the design, the analysis can be more complicated. Again one possibility is to evaluate a cross or pure line simply by the subclass mean. However, most breeders have attempted a more complicated analysis involving, for example, the following model for μ_{ij} the true mean of the cross between the i^{th} sire breed and the j^{th} dam breed.

$$\begin{aligned} \mu_{ij} &= \mu + s_i + d_j + \gamma_{ij} + p \text{ if } i = j \\ &= \mu + s_i + d_j + \gamma_{ij} \text{ if } i \neq j. \end{aligned}$$

From the standpoint of ranking crosses by BLUE, this model is of no particular value, for even with filled subclasses the rank of the coefficient matrix is only b^2, where b is the number of breeds. A solution to the OLS equations is

$$\mu^o = \mathbf{s}^o = \mathbf{d}^o = p^o = \mathbf{0}$$

$$\hat{\gamma}_{ij} = \overline{y}_{ij}.$$

Thus BLUE of a breed cross is simply $\overline{y}_{ij}$, provided $n_{ij} > 0$. The extended model provides no estimate of a missing cross since that is not estimable. In contrast, if one is prepared to use biased estimation, a variety of estimates of missing crosses can be derived, and these same biased estimators may, in fact, be better estimators of filled subclasses than $\overline{y}_{ij}$. Let us restrict ourselves to estimators of μ_{ij} that have expectation, $\mu + s_i + d_j + p$ + linear function of γ if $i = j$, or $\mu + s_i + d_j +$ linear function of γ if $i \neq j$. Assume that the γ_{ii} are different from the γ_{ij} $(i \neq j)$. Accordingly, let us assume for convenience that

$$\begin{aligned} \sum_{j=1}^{b} \gamma_{ij} &= 0 \text{ for } i = 1, \ldots, b, \\ \sum_{i=1}^{b} \gamma_{ij} &= 0 \text{ for } j = 1, \ldots, b, \text{ and} \\ \sum_{i=1}^{b} \gamma_{ii} &= 0. \end{aligned}$$

Next permute all labelling of breeds and compute the average squares and products of the γ_{ij}. These have the following form:

$$\begin{aligned} \text{Av.}(\gamma_{ii})^2 &= d. \\ \text{Av.}(\gamma_{ij})^2 &= c. \\ \text{Av.}(\gamma_{ii}\, \gamma_{jj}) &= -d/(b-1). \end{aligned}$$

$$
\begin{aligned}
\text{Av.}(\gamma_{ij}\,\gamma_{ik}) &= \text{Av.}(\gamma_{ij}\,\gamma_{kj}) = \frac{d - c(b-1)}{(b-1)(b-2)}. \\
\text{Av.}(\gamma_{ii}\,\gamma_{ij}) &= -d/(b-1). \\
\text{Av.}(\gamma_{ii}\,\gamma_{ji}) &= -d/(b-1). \\
\text{Av.}(\gamma_{ij}\,\gamma_{ji}) &= r. \\
\text{Av.}(\gamma_{ii}\,\gamma_{jk}) &= 2d/(b-1)(b-2). \\
\text{Av.}(\gamma_{ij}\,\gamma_{ki}) &= \frac{d - r(b-1)}{(b-1)(b-2)}. \\
\text{Av.}(\gamma_{ij}\,\gamma_{kl}) &= \frac{(c+r)(b-1) - 4d}{(b-1)(b-2)(b-3)}. \\
\text{Av.}(\gamma_{ij}\,\gamma_{jk}) &= \text{Av.}(\gamma_{ij}\,\gamma_{ki}).
\end{aligned}
$$

These squares and products comprise a singular **P** matrix which could then be used in pseudo-mixed model equations. This would, of course, require estimating d, c, r from the data. Solving the resulting mixed model type equations,

$$
\begin{aligned}
\hat{\mu}_{ii} &= \mu^o + s_i^o + d_i^o + \hat{\gamma}_{ii} + p^o, \\
\hat{\mu}_{ij} &= \mu^o + s_i^o + d_i^o + \hat{\gamma}_{ij},
\end{aligned}
$$

when $i \neq j$.

A simpler method is to pretend that the model for μ_{ij} is

$$\mu_{ij} = \mu + s_i + d_j + \gamma_{ij} + r_{(i,j)},$$

when $i \neq j$, and

$$\mu_{ii} = \mu + s_i + d_j + \gamma_{ii} + p.$$

$\mathbf{r}$ has $b(b-1)/2$ elements and (ij) denotes $i < j$. Thus the element of $\mathbf{r}$ for μ_{ij} is the same as for μ_{ji}. Then partition $\boldsymbol{\gamma}$ into the γ_{ii} elements and the γ_{ij} elements and pretend that $\boldsymbol{\gamma}$ and $\mathbf{r}$ are random variables with

$$Var\begin{pmatrix} \gamma_{11} \\ \gamma_{22} \\ \vdots \end{pmatrix} = \mathbf{I}\sigma_1^2,\ Var\begin{pmatrix} \gamma_{12} \\ \gamma_{13} \\ \vdots \end{pmatrix} = \mathbf{I}\sigma_2^2,\ Var(\mathbf{r}) = \mathbf{I}\sigma_3^2.$$

The covariances between these three vectors are all null. Then set up and solve the mixed model equations. With proper choices of values of σ_1^2, σ_2^2, σ_3^2 relative to b, d, c, r the estimates of the breed crosses are identical to the previous method using singular **P**. The latter method is easier to compute and it is also much easier to estimate σ_1^2, σ_2^2, σ_3^2 than the parameters of **P**. For example, we could use Method 3 by computing appropriate reductions and equating to their expectations.

We illustrate these two methods with a 4 breed cross. The n_{ij} and $y_{ij.}$ were as follows.

n_{ij}				$y_{ij.}$			
5	2	3	1	6	3	2	7
4	2	6	7	8	3	5	6
3	5	2	8	9	4	7	3
4	2	3	4	2	6	8	6

Assume that $\mathbf{P}$ is the following matrix, (30.4) … (30.6). $Var(\mathbf{e}) = \mathbf{I}$. Then we premultiply the OLS equations by

$$\begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{P} \end{pmatrix}$$

and add $\mathbf{I}$ to the last 16 diagonal coefficients.

Upper 8×8

$$\begin{pmatrix}
1.8 & -.6 & -.6 & -.6 & -.6 & -.6 & .6 & .6 \\
 & 4.48 & -1.94 & -1.94 & .88 & -.6 & -.14 & -.14 \\
 & & 4.48 & -1.94 & -.14 & .6 & -1.94 & 1.48 \\
 & & & 4.48 & -.14 & .6 & 1.48 & -1.94 \\
 & & & & 4.48 & -.6 & -1.94 & -1.94 \\
 & & & & & 1.8 & -.6 & -.6 \\
 & & & & & & 4.48 & -1.94 \\
 & & & & & & & 4.48
\end{pmatrix} \quad (4)$$

Upper right 8×8 and (lower left 8×8)'

$$\begin{pmatrix}
-.6 & .6 & -.6 & .6 & -.6 & .6 & .6 & -.6 \\
-.14 & -1.94 & .6 & 1.48 & -.14 & -1.94 & 1.48 & .6 \\
.88 & -.14 & -.6 & -.14 & -.14 & 1.48 & -1.94 & .6 \\
-.14 & 1.48 & .6 & -1.94 & .88 & -.14 & -.14 & -.6 \\
-1.94 & -.14 & .6 & 1.48 & -1.94 & -.14 & 1.48 & .6 \\
.6 & -.6 & -.6 & .6 & .6 & -.6 & .6 & -.6 \\
-.14 & .88 & -.6 & -.14 & 1.48 & -.14 & -1.94 & .6 \\
1.48 & -.14 & .6 & -1.94 & -.14 & .88 & -.14 & -.6
\end{pmatrix} \quad (5)$$

Lower right 8×8

$$\begin{pmatrix}
4.48 & -1.94 & -.6 & -1.94 & -1.94 & 1.48 & -.14 & .6 \\
 & 4.48 & -.6 & -1.94 & 1.48 & -1.94 & -.14 & .6 \\
 & & 1.8 & -.6 & .6 & .6 & -.6 & -.6 \\
 & & & 4.48 & -.14 & -.14 & .88 & -.6 \\
 & & & & 4.48 & -1.94 & -1.94 & -.6 \\
 & & & & & 4.48 & -1.94 & -.6 \\
 & & & & & & 4.48 & -.6 \\
 & & & & & & & 1.8
\end{pmatrix} \quad (6)$$

A solution to these equations is

$$
\begin{aligned}
\mu^o &= 0, \\
\mathbf{s}^o &= (2.923,\ 1.713,\ 2.311,\ 2.329)', \\
\mathbf{d}^o &= (-.652,\ -.636,\ -.423,\ 0)', \\
p^o &= .007.
\end{aligned}
$$

$\hat{\boldsymbol{\gamma}}$ in tabular form is

$$
\begin{pmatrix}
-1.035 & -.754 & -1.749 & 3.538 \\
.898 & .377 & -.434 & -.841 \\
1.286 & -.836 & 1.453 & -1.902 \\
-1.149 & 1.214 & .729 & -.795
\end{pmatrix}.
$$

The resulting $\hat{\mu}_{ij}$ are

$$
\begin{pmatrix}
1.243 & 1.533 & .752 & 6.462 \\
1.959 & 1.461 & .857 & .872 \\
2.945 & .839 & 3.349 & .409 \\
.528 & 2.908 & 2.635 & 1.541
\end{pmatrix}.
$$

Note that these $\hat{\mu}_{ij} \neq \overline{y}_{ij}$ but are not markedly different from them. The same $\hat{\mu}_{ij}$ can be obtained by using

$$
\begin{aligned}
Var(\boldsymbol{\gamma}_{ii}) &= -2.88\,\mathbf{I}, \\
Var(\boldsymbol{\gamma}_{ij}) &= 7.2\,\mathbf{I}, \\
Var(\mathbf{r}) &= 2.64\,\mathbf{I}.
\end{aligned}
$$

The solution to these mixed model equations is different from before, but the resulting $\hat{\mu}_{ij}$ are identical. Ordinarily one would not accept a negative "variance". The reason for this in our example was a bad choice of the parameters of $\mathbf{P}$. The OLS coefficient matrix for this solution is in (30.7) … (30.9). The right hand sides are (18, 22, 23, 22, 25, 16, 22, 22, 6, 3, 2, 7, 8, 3, 5, 6, 9, 4, 7, 3, 2, 6, 8, 6, 11, 11, 9, 9, 12, 11). μ^o and d_4^o are deleted giving a solution of 0 for them. The OLS equations for the preceding method are the same as these except the last 6 equations and unknowns are deleted. The solution is

$$
\begin{aligned}
\mu^o &= 0, \\
\mathbf{s}^o &= (1.460,\ 1.379,\ 2.838,\ 1.058)', \\
\mathbf{d}^o &= (-.844,\ .301,\ 1.375,\ 0)', \\
p^o &= .007, \\
\mathbf{r}^o &= (.253,\ -.239,\ 1.125,\ -.888,\ .220,\ -.471)'.
\end{aligned}
$$

$$
\hat{\boldsymbol{\gamma}} \text{ in tabular form} =
\begin{pmatrix}
.621 & -.481 & -1.844 & 3.877 \\
1.172 & -.226 & -1.009 & -.727 \\
1.191 & -1.412 & -.872 & -1.958 \\
-.810 & 1.328 & .673 & .477
\end{pmatrix}.
$$

This solution gives the same result for $\hat{\mu}_{ij}$ as before.

Upper left 15 × 15

$$
\begin{pmatrix}
11 & 0 & 0 & 0 & 5 & 2 & 3 & 5 & 5 & 2 & 3 & 1 & 0 & 0 & 0 \\
 & 19 & 0 & 0 & 4 & 2 & 6 & 2 & 0 & 0 & 0 & 0 & 4 & 2 & 6 \\
 & & 18 & 0 & 3 & 5 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & 13 & 4 & 2 & 3 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 16 & 0 & 0 & 5 & 5 & 0 & 0 & 0 & 4 & 0 & 0 \\
 & & & & & 11 & 0 & 2 & 0 & 2 & 0 & 0 & 0 & 2 & 0 \\
 & & & & & & 14 & 2 & 0 & 0 & 3 & 0 & 0 & 0 & 6 \\
 & & & & & & & 13 & 5 & 0 & 0 & 0 & 0 & 2 & 0 \\
 & & & & & & & & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 2 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 3 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & 1 & 0 & 0 & 0 \\
 & & & & & & & & & & & & 4 & 0 & 0 \\
 & & & & & & & & & & & & & 2 & 0 \\
 & & & & & & & & & & & & & & 6
\end{pmatrix}. \tag{7}
$$

Upper right 15 × 15 and (lower left 15 × 15)'

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 3 & 1 & 0 & 0 & 0 \\
7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 6 & 7 & 0 \\
0 & 3 & 5 & 2 & 8 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 5 & 0 & 8 \\
0 & 0 & 0 & 0 & 0 & 4 & 2 & 3 & 4 & 0 & 0 & 4 & 0 & 2 & 3 \\
0 & 3 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 & 3 & 4 & 0 & 0 & 0 \\
0 & 0 & 5 & 0 & 0 & 0 & 2 & 0 & 0 & 2 & 0 & 0 & 5 & 2 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 3 & 0 & 0 & 3 & 0 & 6 & 0 & 3 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0
\end{pmatrix}. \tag{8}
$$

Lower right 15×15

$$\begin{pmatrix}
7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 & 0 \\
 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\
 & & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 \\
 & & & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \\
 & & & & & 4 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\
 & & & & & & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\
 & & & & & & & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\
 & & & & & & & \text{dg} & (4 & 6 & 6 & 5 & 11 & 9 & 11)
\end{pmatrix} \tag{9}$$

The method just preceding is convenient for missing subclasses. In that case γ_{ij} associated with $n_{ij} = 0$ are included in the mixed model equations.

Chapter 31
Maternal Effects

C. R. Henderson

1984 - Guelph

Many traits are influenced by the environment contributed by the dam. This is particularly true for traits measured early in life and for species in which the dam nurses the young for several weeks or months. Examples are 3 month weights of pigs, 180 day weights of beef calves, and weaning weights of lambs. In fact, genetic merit for maternal ability can be an important trait for which to select. This chapter is concerned with some models for maternal effects and with BLUP of them.

1 Model For Maternal Effects

Maternal effects can be estimated only through the progeny performance of a female or the progeny performance of a related female when direct and maternal effects are uncorrelated. If they are correlated, maternal effects can be evaluated whenever direct can be. Because the maternal ability is actually a phenotypic manifestation, it can be regarded as the sum of a genetic effect and an environmental effect. The genetic effect can be partitioned at least conceptually into additive, dominance, additive $\times$ additive, etc. components. The environmental part can be partitioned, as is often done for lactation yield in dairy cows, into temporary and permanent environmental effects. Some workers have suggested that the permanent effects can be attributed in part to the maternal contribution of the dam of the dam whose maternal effects are under consideration.

Obviously if one is to evaluate individuals for maternal abilities, estimates of the underlying variances and covariances are needed. This is a difficult problem in part due to much confounding between maternal and direct genetic effects. BLUP solutions are probably quite sensitive to errors in estimates of the parameters used in the prediction equations. We will illustrate these principles with some examples.

2 Pedigrees Used In Example

Individual No.	Sex	Sire	Dam	Record
1	Male	Unknown	Unknown	6
2	Female	Unknown	Unknown	9
3	Female	1	2	4
4	Female	1	2	7
5	Male	Unknown	Unknown	8
6	Male	Unknown	Unknown	3
7	Male	6	3	6
8	Male	5	4	8

This gives an **A** matrix as follows:

$$\begin{pmatrix} 1 & 0 & .5 & .5 & 0 & 0 & .25 & .25 \\ & 1 & .5 & .5 & 0 & 0 & .25 & .25 \\ & & 1 & .5 & 0 & 0 & .5 & .25 \\ & & & 1 & 0 & 0 & .25 & .5 \\ & & & & 1 & 0 & 0 & .5 \\ & & & & & 1 & .5 & 0 \\ & & & & & & 1 & .125 \\ & & & & & & & 1 \end{pmatrix}.$$

The corresponding dominance relationship matrix is a matrix with 1's in the diagonals, and the only non-zero off-diagonal element is that for $d_{34} = .25$.

For our first example we assume a model with both additive direct and additive maternal effects. We assume that $\sigma_e^2 = 1$, σ_a^2 (direct) $= .5$, σ_a^2 (maternal) $= .4$, covariance direct with maternal $= .2$. We assume $\mathbf{X}\boldsymbol{\beta} = \mathbf{1}\mu$. In all of our examples we have assumed that the permanent environmental contribution to maternal effects is negligible. If one does not wish to make this assumption, a vector of such effects can be included. Its variance is $\mathbf{I}\sigma_p^2$, and is assumed to be uncorrelated with any other variables. Then permanent environmental effects can be predicted only for those animals with recorded progeny. Then the incidence matrix excluding **p** is

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (1)$$

Cols. 2-9 represent **a** and cols 10-17 represent **m**. This gives the following OLS equations.

$$
\left(\begin{array}{ccccccccccccccccc}
8 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\
 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & 1 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & & 0 & 0 & 0 \\
 & & & & & & & & & & & & & & & 0 & 0 \\
 & & & & & & & & & & & & & & & & 0
\end{array}\right)
\left(\begin{array}{c} \hat{\mu} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{m}} \end{array}\right)
=
\left(\begin{array}{c} 51 \\ 6 \\ 9 \\ 4 \\ 7 \\ 8 \\ 3 \\ 6 \\ 8 \\ 0 \\ 11 \\ 6 \\ 8 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right). \qquad (2)
$$

$$
\mathbf{G} = \left(\begin{array}{cc} .5\mathbf{A} & .2\mathbf{A} \\ .2\mathbf{A} & .4\mathbf{A} \end{array}\right).
$$

Adding the inverse of **G** to the lower 16 × 16 submatrix of (31.2) gives the mixed model equations, the solution to which is

$$
\begin{aligned}
\hat{\mu} &= 6.386, \\
\hat{\mathbf{a}} &= (-.241,\ .541,\ -.269,\ .400,\ .658,\ -1.072,\ -.585,\ .709)', \\
\hat{\mathbf{m}} &= (.074,\ -.136,\ -.144,\ .184,\ .263,\ -.429,\ -.252,\ .296)'.
\end{aligned}
$$

In contrast, if covariance $(\mathbf{a}, \mathbf{m}') = \mathbf{0}$, the maternal predictions of 5 and 6 are 0. With $\sigma_a^2 = .5$, $\sigma_m^2 = .4$, $\sigma_{am}^2 = 0$ the solution is

$$
\begin{aligned}
\hat{\mu} &= 6.409, \\
\hat{\mathbf{a}} &= (-.280,\ .720,\ -.214,\ .440,\ .659,\ -1.099,\ -.602,\ .742)', \\
\hat{\mathbf{m}} &= (.198,\ -.344,\ -.029,\ .081,\ 0,\ 0,\ -.014,\ .040)'.
\end{aligned}
$$

Note now that 5 and 6 cannot be evaluated for **m** since they are males and have no female relatives with progeny.

3 Additive And Dominance Maternal And Direct Effects

If we assume that additive and dominance affect both direct and maternal merit, the incidence matrix of (31.1) is augmented on the right by the last 16 columns of (31.1) giving an 8×33 matrix. Assume the same additive direct and maternal parameters as before and let the dominance parameters be .3 for direct variance, .2 for maternal, and .1 for their covariance. Then

$$\mathbf{G} = \begin{pmatrix} .5\mathbf{A} & .2\mathbf{A} & \mathbf{0} & \mathbf{0} \\ .2\mathbf{A} & .4\mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & .3\mathbf{D} & .1\mathbf{D} \\ \mathbf{0} & \mathbf{0} & .1\mathbf{D} & .2\mathbf{D} \end{pmatrix}.$$

The solution is

$$\begin{aligned} \hat{\mu} &= 6.405, \\ \mathbf{a} \text{ direct} &= (-.210,\ .478,\ -.217,\ .350,\ .545,\ -.904,\ -.503,\ .588)', \\ \mathbf{a} \text{ maternal} &= (.043,\ -.083,\ -.123,\ .156,\ .218,\ -.362,\ -.220,\ .243)', \\ \mathbf{d} \text{ direct} &= (-.045,\ .392,\ -.419,\ .049,\ .242,\ -.577,\ .069,\ .169)', \\ \mathbf{d} \text{ maternal} &= (-.015,\ -.078,\ -.078,\ .119,\ .081,\ -.192,\ .023,\ .056)'. \end{aligned}$$

Quadratics to compute to estimate variances and covariances by MIVQUE would be

$$\begin{aligned} &\hat{\mathbf{a}}(\text{direct})'\mathbf{A}^{-1}\hat{\mathbf{a}}(\text{direct}), \\ &\hat{\mathbf{a}}(\text{direct})'\mathbf{A}^{-1}\hat{\mathbf{a}}(\text{maternal}), \\ &\hat{\mathbf{a}}(\text{maternal})'\mathbf{A}^{-1}\hat{\mathbf{a}}(\text{maternal}), \\ &\hat{\mathbf{d}}(\text{direct})'\mathbf{D}^{-1}\hat{\mathbf{d}}(\text{direct}), \\ &\hat{\mathbf{d}}(\text{direct})'\mathbf{D}^{-1}\hat{\mathbf{d}}(\text{maternal}), \\ &\hat{\mathbf{d}}(\text{maternal})'\mathbf{D}^{-1}\hat{\mathbf{d}}(\text{maternal}), \\ &\hat{\mathbf{e}}'\hat{\mathbf{e}}. \end{aligned}$$

Of course the data of our example would be quite inadequate to estimate these variances and covariances.

Chapter 32
Three Way Mixed Model

C. R. Henderson

1984 - Guelph

Some of the principles of preceding chapters are illustrated in this chapter using an unbalanced 3 way mixed model. The method used here is one of several alternatives that appeals to me at this time. However, I would make no claims that it is "best".

1 The Example

Suppose we have a 3 way classification with factors denoted by A, B, C. The levels of A are random and those of B and C are fixed. Accordingly a traditional mixed model would contain factors and interactions as follows, a, b, c, ab, ac, bc, abc with b, c, and bc fixed, and the others random. The subclass numbers are as follows.

	BC subclasses								
A	11	12	13	21	22	23	31	32	33
1	5	2	3	6	0	3	2	5	0
2	1	2	4	0	5	2	3	6	0
3	0	4	8	2	3	5	7	0	0

The associated ABC subclass means are

$$\begin{pmatrix} 3 & 5 & 2 & 4 & - & 8 & 9 & 2 & - \\ 5 & 6 & 7 & - & 8 & 5 & 2 & 6 & - \\ - & 9 & 8 & 4 & 3 & 7 & 5 & - & - \end{pmatrix}.$$

Because there are no observations on bc_{33}, estimates and tests of b, c, and $b \times c$ that mimic the filled subclass case cannot be accomplished using unbiased estimators. Accordingly, we might use some prior on squares and products of bc_{jk} and obtain biased estimators. Let us assume the following prior values, $\sigma_e^2/\sigma_a^2 = 2$, $\sigma_e^2/\sigma_{ab}^2 = 3$, $\sigma_e^2/\sigma_{ac}^2 = 4$, $\sigma_e^2/$pseudo $\sigma_{bc}^2 = 6$, $\sigma_e^2/\sigma_{abc}^2 = 5$.

2 Estimation And Prediction

The OLS equations that include missing observations have 63 unknowns as follows

a	1 – 3	**ac**	19 – 27
b	4 – 6	**bc**	28 – 36
c	7 – 9	**abc**	37 – 63
ab	10 – 18		

$\overline{\mathbf{W}}$ is a 20 × 63 matrix with 1's in the following columns of the 20 rows. The other elements are 0.

Levels of			
a	b	c	Cols. with 1
1	1	1	1,4,7,10,19,28,37
1	1	2	1,4,8,10,20,29,38
1	1	3	1,4,9,10,21,30,39
1	2	1	1,5,7,11,19,31,40
1	2	3	1,5,9,11,21,33,42
1	3	1	1,6,7,12,19,34,43
1	3	2	1,6,8,12,20,35,44
2	1	1	2,4,7,13,22,28,46
2	1	2	2,4,8,13,23,29,47
2	1	3	2,4,9,13,24,30,48
2	2	2	2,5,8,14,23,32,50
2	2	3	2,5,9,14,24,33,51
2	3	1	2,6,7,15,22,34,52
2	3	2	2,6,8,15,23,35,53
3	1	2	3,4,8,16,26,29,56
3	1	3	3,4,9,16,27,30,57
3	2	1	3,5,7,17,25,31,58
3	2	2	3,5,8,17,26,32,59
3	2	3	3,5,9,17,27,33,60
3	3	1	3,6,7,18,25,34,61

Let $\mathbf{N}$ be a 20 × 20 diagonal matrix with filled subclass numbers in the diagonal, that is $\mathbf{N}$ = diag(5,2,…,5,7). Then the OLS coefficient matris is $\overline{\mathbf{W}}'\mathbf{N}\overline{\mathbf{W}}$, and the right hand side vector is $\mathbf{W}'\mathbf{N}\bar{\mathbf{y}}$, where $\bar{\mathbf{y}} = (3\ \ 5\ \dots\ 7\ \ 5)'$. The right hand side vector is (107, 137, 187, 176, 150, 105, 111, 153, 167, 31, 48, 28, 45, 50, 42, 100, 52, 35, 57, 20, 30, 11, 88, 38, 43, 45, 99, 20, 58, 98, 32, 49, 69, 59, 46, 0, 15, 10, 6, 24, 0, 24, 18, 10, 0, 5, 12, 28, 0, 40, 10, 6, 36, 0, 0, 36, 64, 8, 9, 35, 35, 0, 0).

Now we add the following diagonal matrix to the coefficient matrix, (2, 2, 2, 0, 0, 0, 0, 0, 0, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 6, 6, 6, 6, 6, 6, 6, 6, 6, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5). A resulting mixed model solution is

$$\begin{aligned}
\mathbf{a} &= (-.54801,\ .10555,\ .44246)'. \\
\mathbf{b} &= (6.45206,\ 6.13224,\ 5.77884)'. \\
\mathbf{c} &= (-1.64229,\ -.67094,\ 0)'. \\
\mathbf{ab} &= (-1.21520,\ .46354,\ .38632,\ .14669,\ .29571,\ -.37204, \\
&\quad 1.06850,\ -.75924,\ -.01428)'. \\
\mathbf{ac} &= (.60807,\ -.70385,\ -.17822,\ -.63358,\ .85039,\ -.16403, \\
&\quad .02552,\ -.14653,\ .34225)'. \\
\mathbf{bc} &= (-.12431,\ .47539,\ -.35108,\ -.30013,\ -.05095,\ .35108, \\
&\quad .42444,\ -.42444,\ 0)'. \\
\mathbf{abc} &= (-.26516,\ .34587,\ -.80984,\ -.38914,\ 0,\ .66726,\ 1.14075, \\
&\quad -.90896,\ 0,\ .11598,\ -.38832,\ .36036,\ 0,\ .66901,\ -.49158, \\
&\quad -.62285,\ .39963,\ 0,\ 0,\ .61292,\ .02819,\ .02898,\ -.73014, \\
&\quad .24561,\ -.00857,\ 0,\ 0)'.
\end{aligned}$$

From these results the biased prediction of subclass means are in (32.1).

	B_1			B_2			B_3		
A	C_1	C_2	C_3	C_1	C_2	C_3	C_1	C_2	C_3
1	3.265	4.135	3.350	4.324	4.622	6.888	6.148	2.909	5.439
2	4.420	6.971	6.550	3.957	7.331	6.229	3.038	5.667	5.348
3	6.222	8.234	7.982	3.928	4.217	6.754	5.006	4.965	6.549

(1)

Note that these are different from the $\overline{y}_{ijk}$ for filled subclasses, the latter being BLUE. Also subclass means are predicted for those cases with no observations.

3 Tests Of Hypotheses

Suppose we wish to test the following hypotheses regarding **b**, **c**, and **bc**. Let

$$\mu_{jk} = \mu + b_j + c_k + bc_{jk}.$$

We test $\overline{\mu}_{j.}$ are equal, $\overline{\mu}_{.k}$ are equal, and that all $\overline{\mu}_{jk}$ - $\overline{\mu}_{j.}$ - $\overline{\mu}_{.k}$ + $\overline{\mu}_{..}$ are equal. Of course these functions are not estimable if any jk subclass is missing as is true in our example. Consequently we must resort to biased estimation and accompanying approximate tests based on estimated MSE rather than sampling variances. We assume that our priors are the correct values and proceed for the first test.

$$\mathbf{K}'\boldsymbol{\beta} = \left(\begin{array}{cccccccccccccc} 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 \end{array}\right.$$
$$\left.\begin{array}{ccccccccccccc} 0 & -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \end{array}\right) \boldsymbol{\beta},$$

where $\boldsymbol{\beta}$ is the vector of μ_{ijk} ordered k in j in i. From (32.1) the estimate of these functions is (6.05893, 3.18058). To find the mean squared errors of this function we first compute the mean squared errors of the $\overline{\mu}_{ijk}$. This is $\overline{\mathbf{W}}\mathbf{C}\overline{\mathbf{W}}' \equiv \mathbf{P}$, where $\overline{\mathbf{W}}$ is the matrix relating ijk subclass means to the 63 elements of our model. $\mathbf{C}$ is a g-inverse of the mixed model coefficient matrix. Then the mean squared error of $\mathbf{K}'\boldsymbol{\beta}$ is

$$\mathbf{K}'\mathbf{P}\mathbf{K} = \begin{pmatrix} 17.49718 & 13.13739 \\ & 16.92104 \end{pmatrix}.$$

Then

$$\boldsymbol{\beta}^{o\prime}\mathbf{K}'(\mathbf{K}'\mathbf{P}\mathbf{K})^{-1}\mathbf{K}'\boldsymbol{\beta}^o = 2.364,$$

and this is distributed approximately as χ^2 with 2 d.f. under the null hypothesis.

To test C we use

$$\mathbf{K}'\boldsymbol{\beta} = \left(\begin{matrix} 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \end{matrix}\right.$$
$$\left.\begin{matrix} -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \end{matrix}\right) \boldsymbol{\beta}.$$

$$\mathbf{K}'\boldsymbol{\beta}^o = (-14.78060,\ -6.03849)',$$

with

$$\text{MSE} = \begin{pmatrix} 17.25559 & 10.00658 \\ & 14.13424 \end{pmatrix}.$$

This gives the test criterion = 13.431, distributed approximately as χ^2 with 2 d.f. under the null hypothesis.

For $B \times C$ interaction we use

$$\mathbf{K}'\boldsymbol{\beta} = \left(\begin{matrix} 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 1 \end{matrix}\right.$$
$$\left.\begin{matrix} 0 & -1 & 0 & 1 & 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\ -1 & -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 \end{matrix}\right) \boldsymbol{\beta}.$$

This gives

$$\mathbf{K}'\boldsymbol{\beta}^o = (-.83026,\ 5.25381,\ -4.51772,\ .09417)',$$

with

$$\text{MSE} = \begin{pmatrix} 6.37074 & 4.31788 & 4.56453 & 3.64685 \\ & 6.09614 & 3.77847 & 4.70751 \\ & & 6.32592 & 4.23108 \\ & & & 6.31457 \end{pmatrix}.$$

The test criterion is 21.044 distributed approximately as χ^2 with 4 d.f. under the null hypothesis.

Note that in these examples of hypothesis testing the priors used were quite arbitrary. The tests are of little value unless one has good prior estimates. This of course is true for any unbalanced mixed model design.

4 REML Estimation By EM Method

We next illustrate one round of estimation of variances by the EM algorithm. We treat σ^2_{bc} as a variance. The first round of estimation is obtained from the mixed model solution of Section 32.2. For σ^2_e we compute

$$[\mathbf{y}'\mathbf{y} - \text{soln. vector (r.h.s. vector)}]/[\text{n} - \text{rank }(\mathbf{X})].$$

$$\mathbf{y}'\mathbf{y} = 2802.$$

$$\text{Red} = 2674.47.$$

$$\hat{\sigma}^2_e = (2802 - 2674.47)/(78 - 5) = 1.747.$$

$$\hat{\sigma}^2_a = \left(\hat{\mathbf{a}}'\hat{\mathbf{a}} + tr\ 1.747 \begin{pmatrix} .28568 & .10673 & .10759 \\ & .28645 & .10683 \\ & & .28558 \end{pmatrix}\right) /3 = .669$$

$$\hat{\sigma}^2_{ab} = \left(\hat{\mathbf{a}}\hat{\mathbf{b}}'\hat{\mathbf{a}}\hat{\mathbf{b}} + tr\ 1.747 \begin{pmatrix} .2346 & & \ldots \\ & \ddots & \\ & & .26826 \end{pmatrix}\right) /9 = .847.$$

$$\hat{\sigma}^2_{ac} = \left(\hat{\mathbf{a}}\hat{\mathbf{c}}'\hat{\mathbf{a}}\hat{\mathbf{c}} + tr\ 1.747 \begin{pmatrix} .19027 & & \ldots \\ \vdots & \ddots & \\ & & .18846 \end{pmatrix}\right) /9 = .580.$$

$$\hat{\sigma}^2_{bc} = \left(\hat{\mathbf{b}}\hat{\mathbf{c}}'\hat{\mathbf{b}}\hat{\mathbf{c}} + tr\ 1.747 \begin{pmatrix} .14138 & & \ldots \\ \vdots & \ddots & \\ & & .16607 \end{pmatrix}\right) /9 = .357.$$

$$\hat{\sigma}^2_{abc} = \left(\hat{\mathbf{a}}\hat{\mathbf{b}}\hat{\mathbf{c}}'\hat{\mathbf{a}}\hat{\mathbf{b}}\hat{\mathbf{c}} + tr\ 1.747 \begin{pmatrix} .16505 & & \ldots \\ \vdots & \ddots & \\ & & .20000 \end{pmatrix}\right) /127 = .534.$$

The solution for four rounds follows.

	1	2	3	4
σ_e^2	1.747	1.470	1.185	.915
σ_a^2	.169	.468	.330	.231
σ_{ab}^2	.847	.999	1.090	1.102
σ_{ac}^2	.580	.632	.638	.587
σ_{bc}^2	.357	.370	.362	.327
σ_{abc}^2	.534	.743	1.062	1.506

It appears that $\hat{\sigma}_e^2$ and $\hat{\sigma}_{abc}^2$ may be highly confounded, and convergence will be slow. Note that $\hat{\sigma}_e^2 + \hat{\sigma}_{abc}^2$ does not change much.

Chapter 33
Selection When Variances are Unequal

C. R. Henderson

1984 - Guelph

The mixed model equations for BLUP are well adapted to deal with variances that differ from one subpopulation to another. These unequal variances can apply to either **e** or to **u** or a subvector of **u**. For example, cows are to be selected from several herds, but the variances differ from one herd to another. Some possibilities are the following.

1. σ^2_a , additive genetic variance, is the same in all herds but the within herd σ^2_e differ.

2. σ^2_e is constant from one herd to another but intra-herd σ^2_a differ.

3. Both σ^2_a and σ^2_e differ from herd to herd, but σ^2_a/σ^2_e is constant. That is, intra-herd h^2 is the same in all herds, but the phenotypic variance is different.

4. Both σ^2_a and σ^2_e differ among herds and so does σ^2_a/σ^2_e.

1 Sire Evaluation With Unequal Variances

As an example, AI sires are sometimes evaluated across herds using

$$
\begin{aligned}
y_{ijk} &= s_i + h_j + e_{ijk}. \\
Var(\mathbf{s}) &= \mathbf{A}\sigma^2_s, \\
Var(\mathbf{e}) &= \mathbf{I}\sigma^2_e, \\
Cov(\mathbf{a}, \mathbf{e}') &= \mathbf{0}.
\end{aligned}
$$

h is fixed. Suppose, however, that we assume, probably correctly, that within herd σ^2_e varies from herd to herd, probably related to the level of production. Suppose also that σ^2_s is influenced by the herd. That is, in the population of sires σ^2_s is different when sires are used in herd 1 as compared to σ^2_s when these same sires are used in herd 2. Suppose further that σ^2_s/σ^2_e is the same for every herd. This may be a somewhat unrealistic assumption, but it may be an adequate approximation. We can treat this as a multiple trait problem, trait 1 being progeny values in herd 1, trait 2 being progeny values in herd 2, etc. For purposes of illustration let us assume that all additive genetic correlations between pairs of traits are 1. In that case if the true rankings of sires for herd 1 were known, then these would be the true rankings in herd 2.

Let us order the progeny data by sire within herd. Then

$$\mathbf{R} = \begin{pmatrix} \mathbf{I}v_1 & 0 & \dots & 0 \\ 0 & \mathbf{I}v_2 & \dots & \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \mathbf{I}v_t \end{pmatrix},$$

where there are t herds.

$$\mathbf{G} = \begin{pmatrix} \mathbf{A}w_{11} & \mathbf{A}w_{12} & \dots & \mathbf{A}w_{1t} \\ \mathbf{A}w_{12} & \mathbf{A}w_{22} & \dots & \mathbf{A}w_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}w_{1t} & \mathbf{A}w_{2t} & & \mathbf{A}w_{tt} \end{pmatrix},$$

where v_i/w_{ii} is the same for all $i = 1, \dots, t$. Further $w_{ij} = (w_{ii}w_{jj})^{.5}$. This is, of course, an oversimplified model since it does not take into account season and year of freshening. It would apply to a situation in which all data are from one year and season.

We illustrate this model with a small set of data.

	n_{ij}			y_{ij}		
Sires	1	2	3	1	2	3
1	5	8	0	6	12	-
2	3	4	7	5	8	9
3	0	5	9	-	10	12

$$\mathbf{A} = \begin{pmatrix} 1 & .5 & .5 \\ & 1 & .25 \\ & & 1 \end{pmatrix}.$$

σ_e^2 for the 3 herds is 48, 108, 192, respectively. $Var(\mathbf{s})$ for the 3 herds is

$$\begin{pmatrix} 4\mathbf{A} & 6\mathbf{A} & 8\mathbf{A} \\ & 9\mathbf{A} & 12\mathbf{A} \\ & & 16\mathbf{A} \end{pmatrix}.$$

Note that $6 = [4(9)]^{.5}$, $8 = [4(16)]^{.5}$, and $12 = [9(16)]^{.5}$. Accordingly $\mathbf{G}$ is singular and we need to use the method described in Chapter 5 for singular $\mathbf{G}$. Now the GLS coefficient matrix for fixed $\mathbf{s}$ is in (33.1) . . . (33.3). This corresponds to ordering $(s_{11}, s_{21}, s_{31}, s_{12}, s_{22}, s_{32}, s_{13}, s_{23}, s_3$ The first subscript on s refers to sire number and the second to herd number. The right hand side vector is (.1250, .1042, 0, .1111, .0741, .0926, 0, .0469, .0625, .2292, .2778, .1094)'.

The upper diagonal element of (33.1) to (33.3) is 5/48, 5 being the number of progeny of sire 1 in herd 1, and 48 being σ_e^2 for herd 1. The lower diagonal is 16/192. The first element of the right hand side is 6/48, and the last is 21/192.

Upper left 6 × 6

$$\text{diag}(.10417,\ .06250,\ 0,\ .07407,\ .03704,\ .04630). \tag{1}$$

Upper right 6 × 6 and (lower left 6 × 6)'

$$\begin{pmatrix} 0 & 0 & 0 & .10417 & 0 & 0 \\ 0 & 0 & 0 & .06250 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .07407 & 0 \\ 0 & 0 & 0 & 0 & .03704 & 0 \\ 0 & 0 & 0 & 0 & .04630 & 0 \end{pmatrix}. \tag{2}$$

Lower right 6 × 6

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ & .03646 & 0 & 0 & 0 & .03646 \\ & & .04687 & 0 & 0 & .04687 \\ & & & .16667 & 0 & 0 \\ & & & & .15741 & 0 \\ & & & & & .08333 \end{pmatrix}. \tag{3}$$

Now we multiply these equations by

$$\begin{pmatrix} 4\mathbf{A} & 6\mathbf{A} & 8\mathbf{A} & 0 \\ & 9\mathbf{A} & 12\mathbf{A} & 0 \\ & & 16\mathbf{A} & 0 \\ & & & \mathbf{I}_3 \end{pmatrix},$$

and add 1 to each of the first 9 diagonal elements. Solving these equations the solution is (-.0720, .0249, .0111, -.1080, .0373, .0166, -.1439, .0498, .0222, 1.4106, 1.8018, 1.2782)′. Note that $\hat{s}_{i1}/\hat{s}_{i2} = 2/3$, $\hat{s}_{i1}/\hat{s}_{i3} = 1/2$, $\hat{s}_{i2}/\hat{s}_{i3} = 3/4$. These are in the proportion (2:3:4) which is ($4^{.5}$:$9^{.5}$:$16^{.5}$). Because of this relationship we can reduce the mixed model equations to a set involving s_{i1} and h_j by premultiplying the equations by

$$\begin{pmatrix} 1. & 0 & 0 & 1.5 & 0 & 0 & 2. & 0 & 0 & 0 & 0 & 0 \\ 0 & 1. & 0 & 0 & 1.5 & 0 & 0 & 2. & 0 & 0 & 0 & 0 \\ 0 & 0 & 1. & 0 & 0 & 1.5 & 0 & 0 & 2. & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1. & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1. & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1. \end{pmatrix}. \tag{4}$$

Then the resulting coefficient matrix is post-multiplied by the transpose of matrix (33.4).

This gives equation (33.5).

$$\begin{pmatrix} 15.104 & 4.229 & 4.229 & 3.927 & 5.035 & 2.417 \\ 3.927 & 15.708 & 2.115 & 3.323 & 3.726 & 2.794 \\ 3.927 & 2.115 & 15.708 & 1.964 & 4.028 & 3.247 \\ 0.104 & 0.062 & 0.0 & 0.167 & 0.0 & 0.0 \\ 0.111 & 0.056 & 0.069 & 0.0 & 0.157 & 0.0 \\ 0.0 & 0.073 & 0.094 & 0.0 & 0.0 & 0.083 \end{pmatrix} \begin{pmatrix} \hat{s}_{11} \\ \hat{s}_{21} \\ \hat{s}_{31} \\ \hat{h}_1 \\ \hat{h}_2 \\ \hat{h}_3 \end{pmatrix} = \begin{pmatrix} 16.766 \\ 15.104 \\ 14.122 \\ .229 \\ .278 \\ .109 \end{pmatrix}. \tag{5}$$

The solution is (-.0720, .0249, .0111, 1.4016, 1.8018, 1.2782)′. These are the same as before.

How would one report sire predictions in a problem like this? Probably the logical thing to do is to report them for a herd with average σ_e^2 . Then it should be pointed out that sires are expected to differ more than this in herds with large σ_e^2 and to differ less in herds with small σ_e^2. A simpler method is to set up equations at once involving only s_{i1} or any other chosen s_{ij} (j fixed). We illustrate with s_{i1}. The $\overline{\mathbf{W}}$ matrix for our example with subclass means ordered sires in herds is

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1.5 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1.5 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1.5 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{pmatrix}.$$

This corresponds to $\hat{s}_{i2} = 1.5\ \hat{s}_{i1}$, and $\hat{s}_{i3} = 2\ \hat{s}_{i1}$. Now compute the diagonal matrix

$$\text{diag}(5, 3, 8, 4, 5, 7, 9)\ [\text{dg}(48, 48, 108, 108, 108, 192, 192)]^{-1}\ \equiv\ \mathbf{D}.$$

Then the GLS coefficient matrix is $\overline{\mathbf{W}}'\mathbf{D}\overline{\mathbf{W}}$ and the right hand side is $\overline{\mathbf{W}}'\mathbf{D}\overline{\mathbf{y}}$, where $\overline{\mathbf{y}}$ is the subclass mean vector. This gives

$$\begin{pmatrix} .2708 & 0 & 0 & .1042 & .1111 & 0 \\ & .2917 & 0 & .0625 & .0556 & .0729 \\ & & .2917 & 0 & .0694 & .0937 \\ & & & .1667 & 0 & 0 \\ & & & & .1574 & 0 \\ & & & & & .0833 \end{pmatrix} \begin{pmatrix} \hat{s}_{11} \\ \hat{s}_{21} \\ \hat{s}_{31} \\ \hat{h}_1 \\ \hat{h}_2 \\ \hat{h}_3 \end{pmatrix} = \begin{pmatrix} .2917 \\ .3090 \\ .2639 \\ .2292 \\ .2778 \\ .1094 \end{pmatrix}. \tag{6}$$

Then add $(4\mathbf{A})^{-1}$ to the upper 3 x 3 submatrix of (33.6) to obtain mixed model equations. Remember $4\mathbf{A}$ is the variance of the sires in herd 1. The solution to these equation is as before, (-.0720, .0249, .0111, 1.4106, 1.8018, 1.2782)′.

2 Cow Evaluation With Unequal Variances

Next we illustrate inter-herd joint cow and sire when herd variances are unequal. We assume a simple model

$$y_{ij} = h_i + a_j + e_{ij}.$$

h is fixed, **a** is additive genetic merit with

$$Var(\mathbf{a}) = \begin{pmatrix} \mathbf{A}g_{11} & \mathbf{A}g_{12} & \dots & \mathbf{A}g_{1t} \\ \mathbf{A}g_{12} & \mathbf{A}g_{22} & \dots & \mathbf{A}g_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}g_{1t} & \mathbf{A}g_{2t} & & \mathbf{A}g_{tt} \end{pmatrix}.$$

A is the numerator relationship for all animals. There are t herds, and we treat production as a different trait in each herd. We assume genetic correlations of 1. Therefore $g_{ij} = (g_{ii}g_{jj})^{.5}$.

$$Var(\mathbf{e}) = \begin{pmatrix} \mathbf{I}v_1 & & & \mathbf{0} \\ & \mathbf{I}v_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{I}v_t \end{pmatrix}.$$

First we assume σ_a^2/σ_e^2 is the same for all herds. Therefore g_{ii}/v_i is the same for all herds.

As an example suppose that we have 2 herds with cows 2, 3 making records in herd 1 and cows 4, 5 making records in herd 2. These animals are out of unrelated dams, and the sire of 2 and 4 is 1. The records are 3, 2, 5, 6.

$$\mathbf{A} = \begin{pmatrix} 1 & .5 & 0 & .5 & 0 \\ & 1 & 0 & .25 & 0 \\ & & 1 & 0 & 0 \\ & & & 1 & 0 \\ & & & & 1 \end{pmatrix}.$$

Ordering the data by cow number and the unknowns by h_1, h_2, **a** in herd 1, **a** in herd 2 the incidence matrix is

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Suppose that

$$\mathbf{G} = \begin{pmatrix} 4\mathbf{A} & 8\mathbf{A} \\ 8\mathbf{A} & 16\mathbf{A} \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 12\mathbf{I} & \mathbf{0} \\ \mathbf{0} & 48\mathbf{I} \end{pmatrix}.$$

Then $\sigma_e^2/\sigma_a^2 = 3$ in each herd, implying that $h^2 = .25$. Note that **G** is singular so the method for singular **G** is used. With these parameters the mixed model solution is

$$\begin{aligned}\hat{\mathbf{h}} &= (2.508,\ 5.468). \\ \hat{\mathbf{a}} \text{ in herd 1} &= (.030,\ .110,\ -.127,\ -.035,\ .066). \\ \hat{\mathbf{a}} \text{ in herd 2} &= (.061,\ .221,\ -.254,\ -.069,\ .133).\end{aligned}$$

Note that $\hat{a}_i$ in herd 2 = 2 $\hat{a}_i$ in herd 1 corresponding to $(16/4)^{.5} = 2$.

A simpler method is to use an incidence matrix as follows.

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

This corresponds to unknowns h_1, h_2, **a** in herd 1. Now $\mathbf{G} = 4\mathbf{A}$ and **R** is the same as before. The resulting solution is the same as before for $\hat{\mathbf{h}}$ and $\hat{\mathbf{a}}$ in herd 1. Then $\hat{\mathbf{a}}$ in herd 2 is 2 times $\hat{\mathbf{a}}$ in herd 1.

Now suppose that **G** is the same as before but σ_e^2 = 12,24 respectively. Then h^2 is higher in herd 2 than in herd 1. This leads again to the $\hat{a}$ in herd 2 being twice $\hat{a}$ in herd 1, but the $\hat{a}$ for cows making records in herd 2 are relatively more variable, and if we were selecting a single cow, say for planned mating, the chance that she would come from herd 2 is increased. The actual solution in this example is

$$\begin{aligned}\hat{\mathbf{h}} &= (2.513,\ 5.468). \\ \hat{\mathbf{a}} \text{ in herd 1} &= (.011,\ .102,\ -.128,\ -.074,\ .106). \\ \hat{\mathbf{a}} \text{ in herd 2} &= \text{twice those in herd 1.}\end{aligned}$$

The only reason we can compare cows in different herds is the use of sires across herds.

A problem with the methods of this chapter is that the individual intra-herd variances must be estimated with limited data. It would seem, therefore, that it might be advisable to take as the estimate for an individual herd, the estimate coming from that herd regressed toward the mean of variances of all herds, the amount of regression depending upon the number of observations. This would imply, perhaps properly, that intra- herd variances are a sample of some population of variances. I have not derived a method comparable to BLUP for this case.