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

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

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 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

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

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.

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

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,

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

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:

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

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.

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{G} \\ \mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{G} + \mathbf{G} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\boldsymbol{\alpha}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{G}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} \tag{51}$$

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

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

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 $\mathbf{u}$,

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

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 $\mathbf{G}$. 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 $\mathbf{G}$ 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 **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)

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

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

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

The symmetric counterpart of these equations is

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

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

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

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

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

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

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

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

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.

$$\begin{pmatrix} \mathbf{K}'\boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \left(\begin{matrix} .33333 & .33333 & .33333 & 0 \\ .04504 & .04504 & -.09009 & .35135 \\ .03604 & .03604 & -.07207 & .08108 \\ -.07207 & -.07207 & .14414 & -.16216 \end{matrix}\right.$$

$$\left.\begin{matrix} 0 & 0 & 0 \\ .21622 & .21622 & .21622 \\ -.02703 & -.02703 & -.02703 \\ .05405 & .05405 & .05405 \end{matrix}\right)\mathbf{y} \quad (72)$$

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

$$= \text{ [matrix (72)] } \mathbf{R} \text{ [matrix (72)]}'$$

$$= \begin{pmatrix} 1.6667 & 0 & 0 & 0 \\ & 1.37935 & .10348 & -.20696 \\ & & .08278 & -.16557 \\ & & & .33114 \end{pmatrix}. \quad (73)$$

For $\mathbf{u}$ in

$$\begin{pmatrix} \mathbf{K}'\boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{K}' \\ \mathbf{I}_2 \end{pmatrix} \text{ [matrix (72)] } \mathbf{Z}$$

$$= \begin{pmatrix} .66667 & .33333 \\ .44144 & .55856 \\ .15315 & -.15315 \\ -.30631 & .30631 \end{pmatrix}. \quad (74)$$

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

$$= \text{ [matrix (74)] } \mathbf{G} \text{ [matrix (74)]}'$$

$$= \begin{pmatrix} 1.6667 & 1.66662 & 0 & 0 \\ & 1.81885 & -.10348 & .20696 \\ & & .07037 & -.14074 \\ & & & .28143 \end{pmatrix}. \quad (75)$$

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

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

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

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

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

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

Then

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

Let an equivalent model be

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Tu} + \boldsymbol{\varepsilon}, \tag{80}$$

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

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

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 $\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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).