Biometrics *31*, 423–447
June 1975

BEST LINEAR UNBIASED ESTIMATION AND PREDICTION UNDER A SELECTION MODEL

C. R. Henderson

Department of Animal Science, Cornell University, Ithaca, N. Y. 14850, U.S.A.

SUMMARY

Mixed linear models are assumed in most animal breeding applications. Convenient methods for computing BLUE of the estimable linear functions of the fixed elements of the model and for computing best linear unbiased predictions of the random elements of the model have been available. Most data available to animal breeders, however, do not meet the usual requirements of random sampling, the problem being that the data arise either from selection experiments or from breeders' herds which are undergoing selection. Consequently, the usual methods are likely to yield biased estimates and predictions. Methods for dealing with such data are presented in this paper.

1. INTRODUCTION

Data available to animal breeding research workers for estimation of genetic and environmental parameters and to practitioners for making selection decisions almost invariably are provided by herds in which selection has been practiced. As a consequence, the usual assumptions of random sampling invoked for estimation and prediction are seldom valid. For example, production records in dairy cows have been used to estimate genetic and environmental trends. It is obvious from knowledge of the industry that the mean of first lactation records on cows that also produce a second record is higher than the mean of first records on contemporaries that do not produce a second record. This should suggest caution in applying the usual linear model methods to these data. Henderson [1949] and Henderson *et al.* [1959] presented methods for obtaining unbiased estimators in such cases. Similarly, Lush and Shrode [1950] described the bias in estimation of age effects due to culling on previous production. A similar problem occurs in evaluation of sires used in artificial breeding programs. Progeny tests are made on large numbers of sires. Then, only those with the best progeny are continued in service to produce many more progeny with records. Further, there is a tendency among those retained for the number of subsequent progeny to be positively correlated with past progeny production.

Two different possibilities for the solution of problems of this sort suggest themselves: (1) invoking a joint distribution of certain random variables of the model and the design matrix and (2) writing a distribution conditional on selected random variables. The first of these methods has proved to be extremely difficult, but some useful results can be attained using the second, as described in this paper. These are based on deriving best linear unbiased estimators and predictors under a model conditional on selection of certain linear functions of random variables jointly distributed with the random variables of the usual linear model. For example, in the cow culling case the selection variable could be the difference between means of first records on cows continuing in production and those that are culled.

It is obvious that the subclass numbers are unequal in data which have arisen from selection situations, and in fact there are usually many empty subclasses. Professor Snedecor helped many research workers deal with the problems of unequal subclass numbers through his teaching and his book with its examples of strains of mice crossclassified with isolations of typhoid bacilli (Snedecor [1946]).

This paper presents some general results on best linear unbiased estimation and prediction from mixed linear models conditional on a selected vector variable. The methods are an extension of some of the author's mixed model techniques for the unconditional model. These latter methods are summarized in the next section.

2. BEST LINEAR UNBIASED PREDICTION

2.1 A mixed linear model

A mixed linear model is assumed in many genetic applications and can be represented as follows.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \tag{1}$$

where $\mathbf{y}$ is an $n \times 1$ observation vector, $\mathbf{X}$ is a known, $n \times p$ matrix, $\boldsymbol{\beta}$ is an unknown, fixed vector, and $\mathbf{Z}$ is a known, $n \times q$ matrix. $\mathbf{u}$ and $\mathbf{e}$ are nonobservable random vectors with null means and

$$\mathrm{Var}\begin{bmatrix}\mathbf{u}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{G} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}\end{bmatrix}\sigma^2$$

where σ^2 is a scalar, possibly unknown, and $\mathbf{G}$ and $\mathbf{R}$ are both nonsingular.

Now, given a sample vector, $\mathbf{y}$, we wish to do one or more of the following:

1. Estimate some estimable linear function of $\boldsymbol{\beta}$.
2. Test hypotheses regarding $\boldsymbol{\beta}$.
3. Estimate $\mathbf{G}$, $\mathbf{R}$, and σ^2.
4. Predict $\mathbf{u}$ or some linear functions of $\mathbf{u}$.
5. Predict linear functions of $\boldsymbol{\beta}$ and $\mathbf{u}$ jointly.

The animal breeding research worker is usually concerned with the first three, and the practitioner with the last two of these. The research worker is concerned also with developing better methods for the practitioner to accomplish the last two.

2.2 Best linear prediction

It is known that if $\boldsymbol{\beta}$, $\mathbf{G}$, and $\mathbf{R}$ are known, the best *linear* predictor of $\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}$ is

$$\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \tag{2}$$

where $\mathbf{V} = \mathbf{R} + \mathbf{Z}\mathbf{G}\mathbf{Z}'$. In the multivariate normal case, (2) is of course $E(\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u} \mid \mathbf{y})$ and consequently is the *best* predictor. Whether or not normality is implied, (2) is the usual selection index used in animal breeding.

2.3 Best linear unbiased prediction

Now if $\boldsymbol{\beta}$ is unknown, as is usually the case, the method of (2) cannot be used. A modification invoking unbiasedness can be employed, however. By unbiased we mean that $E(\text{predictor}) = E(\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}) = \mathbf{k}'\boldsymbol{\beta}$. Henderson [1963] showed that the best linear unbiased predictor (BLUP) of $\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}$ is

$$\mathbf{k}'\hat{\boldsymbol{\beta}} + \mathbf{m}'\mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \tag{3}$$

where $\hat{\boldsymbol{\beta}}$ is any solution to (4), the generalized least squares (GLS) equations,

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{y}. \tag{4}$$

The difficulty with this method is that $\mathbf{V}$ is often a matrix so large that its inversion is very costly. An alternative method was suggested by Henderson [1950]. The prediction of $\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u}$ is $\mathbf{k}'\hat{\boldsymbol{\beta}} + \mathbf{m}'\hat{\mathbf{u}}$ where $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ are any solution to (5);

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}. \tag{5}$$

Henderson *et al.* [1959] proved that $\hat{\boldsymbol{\beta}}$ of (5) is a solution to (4), the GLS equations. The proof involved showing that

$$\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1} = \mathbf{V}^{-1}. \tag{6}$$

Henderson [1963] proved that $\hat{\mathbf{u}}$ of (5) is equal to $\mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$ of (3). This was done by showing that

$$(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1} = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1} \tag{7}$$

and by noting that $\mathbf{X}\hat{\boldsymbol{\beta}}$ of (5) is the GLS estimator of $\mathbf{X}\boldsymbol{\beta}$.

The obvious advantage computationally of (5) over (3) is that neither $\mathbf{V}$ nor its inverse is required. Of course $\mathbf{R}$ has the same dimensions as $\mathbf{V}$, but it usually is an identity matrix. Further, $\mathbf{G}$ is often diagonal, and $\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1}$ is either diagonal or has a large diagonal submatrix.

2.4 Distributional properties of $\hat{\boldsymbol{\beta}}$ *and* $\hat{\mathbf{u}}$

Let some symmetric g-inverse of the coefficient matrix of (5) be

$$\begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{bmatrix}. \tag{8}$$

$\mathbf{C}_{22}$ is always unique, but $\mathbf{C}_{11}$ and $\mathbf{C}_{12}$ are unique only if $\mathbf{X}$ has full column rank. Now assuming that $\mathbf{K}'\boldsymbol{\beta}$ is estimable,

$$\begin{aligned}
&\mathrm{Var}\,(\mathbf{K}'\hat{\boldsymbol{\beta}}) = \mathbf{K}'\mathbf{C}_{11}\mathbf{K}\sigma^2, \\
&\mathrm{Cov}\,(\mathbf{K}'\hat{\boldsymbol{\beta}}, \hat{\mathbf{u}}') = \mathbf{0}, \\
&\mathrm{Cov}\,(\mathbf{K}'\hat{\boldsymbol{\beta}}, \mathbf{u}') = -\mathbf{K}'\mathbf{C}_{12}\sigma^2, \\
&\mathrm{Cov}\,(\mathbf{K}'\hat{\boldsymbol{\beta}}, \hat{\mathbf{u}}' - \mathbf{u}') = \mathbf{K}'\mathbf{C}_{12}\sigma^2, \\
&\mathrm{Var}\,(\hat{\mathbf{u}}) = (\mathbf{G} - \mathbf{C}_{22})\sigma^2, \\
&\mathrm{Cov}\,(\hat{\mathbf{u}}, \mathbf{u}') = \mathrm{Var}\,(\hat{\mathbf{u}}), \\
&\mathrm{Var}\,(\hat{\mathbf{u}} - \mathbf{u}) = \mathbf{C}_{22}\sigma^2.
\end{aligned}$$

Accordingly,

$$\mathrm{Var}\,(\mathbf{K}'\hat{\boldsymbol{\beta}} + \mathbf{M}'\hat{\mathbf{u}} - \mathbf{K}'\boldsymbol{\beta} - \mathbf{M}'\mathbf{u}) = [\mathbf{K}' \ \vdots \ \mathbf{M}'] \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{K} \\ \mathbf{M} \end{bmatrix} \sigma^2.$$

The proof of these results for the full rank case are presented in Appendix A.

3. EXAMPLES OF BLUP IN THE NO SELECTION MODEL

Suppose we have records on progeny of two sires in two herds with progeny numbers as follows.

Sire	Herd 1	Herd 2
1	5	10
2	5	0

We assume that the model for the kth progeny of the ith sire in the jth herd is

$$y_{ijk} = \mu + s_i + h_j + e_{ijk} \tag{9}$$

where $\mathbf{e}' = [e_{111}, \cdots, e_{215}]$, $E(\mathbf{e}) = 0$, and $\mathrm{Var}(\mathbf{e}) = \mathbf{I}\sigma^2$.

We wish to predict $s_1 - s_2$. The best linear unbiased predictor depends upon the distribution of $\mathbf{s}' = [s_1 \quad s_2]$ and $\mathbf{h}' = [h_1 \quad h_2]$. Various assumptions are described below.

3.1 Sires and herds fixed

The regular least squares equations are

$$\begin{bmatrix} 20 & 15 & 5 & 10 & 10 \\ 15 & 15 & 0 & 5 & 10 \\ 5 & 0 & 5 & 5 & 0 \\ 10 & 5 & 5 & 10 & 0 \\ 10 & 10 & 0 & 0 & 10 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{h}_1 \\ \hat{h}_2 \end{bmatrix} = \begin{bmatrix} 5 & 10 & 5 \\ 5 & 10 & 0 \\ 0 & 0 & 5 \\ 5 & 0 & 5 \\ 0 & 10 & 0 \end{bmatrix} \bar{\mathbf{y}} \tag{10}$$

where $\bar{\mathbf{y}}' = [\bar{y}_{11.} \quad \bar{y}_{12.} \quad \bar{y}_{21.}]$. For any solution to (10),

$$\hat{s}_1 - \hat{s}_2 = [1 \quad 0 \quad -1]\bar{\mathbf{y}} \tag{11}$$

with variance $= 0.4\sigma^2$.

3.2 Sires fixed and herds random

$E(\mathbf{h}) = \mathbf{0}$ and $\mathrm{Var}(\mathbf{h}) = 0.5\mathbf{I}\sigma^2$. Then the BLUP equations according to (5) are the same as (10) except that a diagonal matrix with diagonal coefficients

$$[0 \quad 0 \quad 0 \quad 2 \quad 2]$$

is added to the coefficient matrix of (10). Now for any solution,

$$\hat{s}_1 - \hat{s}_2 = [11 \quad 2 \quad -13]\bar{\mathbf{y}}/13 \tag{12}$$

with variance $= 24\sigma^2/65$, which is less than $0.4\sigma^2$. Therefore, if in fact $\mathbf{h}$ were random but computations were done as in (10), the prediction error variance would have been unnecessarily large.

3.3 Sires random and herds fixed

$E(\mathbf{s}) = \mathbf{0}$ and $\mathrm{Var}(\mathbf{s}) = 0.1\mathbf{I}\sigma^2$. Now the BLUP equations are like those of (10) except that a diagonal matrix with diagonal elements

$$[0 \quad 10 \quad 10 \quad 0 \quad 0]$$

is added to the coefficient matrix. The resulting solution is

$$\hat{s}_1 - \hat{s}_2 = [1 \quad 0 \quad -1]\bar{\mathbf{y}}/3 \tag{13}$$

with prediction error variance $= 2\sigma^2/15$. We illustrate calculation of this variance from a g-inverse of the coefficient matrix. A g-inverse is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & -3 & -5 \\ 0 & 1 & 5 & -3 & -1 \\ 0 & -3 & -3 & 9 & 3 \\ 0 & -5 & -1 & 3 & 11 \end{bmatrix} \frac{1}{60}$$

and $\mathrm{Var}(\hat{s}_1 - \hat{s}_2 - s_1 + s_2) = \sigma_2[5 - (2)(1) + 5]/60 = 2\sigma^2/15$.

3.4 Sires and herds random

Now the diagonal matrix with diagonals

$$[0 \quad 10 \quad 10 \quad 2 \quad 2]$$

is added to the coefficient matrix of (10). The solution to $\hat{s}_1 - \hat{s}_2$ is

$$[11 \quad 2 \quad -13]\bar{\mathbf{y}}/37 \tag{14}$$

with prediction error variance $= 24\sigma^2/185$.

4. EXAMPLES OF EFFECT OF SELECTION

The example of section 3 is used to illustrate the consequences of selection. Note that sire 2 has no progeny in herd 2. Suppose we know that the progeny in herd 1 were produced earlier than those in herd 2 and we suspect that the reason sire 2 was not used in herd 2 was that his progeny were inferior to those of sire 1 in herd **1.** If this is true, is the prediction of $s_1 - s_2$ biased as a consequence? In the absence of selection,

$$\begin{aligned} E(\bar{y}_{11.} - \bar{y}_{21.}) &= E(s_1 - s_2) \\ &= s_1 - s_2 \text{ in the fixed sire model} \\ &= 0 \text{ in the random sire model.} \end{aligned}$$

Suppose that due to selection this expectation is $s_1 - s_2 + \alpha$ in the fixed sire case or is α in the random sire case. Now assuming a multivariate normal distribution we examine the bias in prediction of $s_1 - s_2$ under the various models and suggest some unbiased predictors.

4.1 Fixed sires and herds

Conditional means, given that $E(\bar{y}_{11.} - \bar{y}_{21.}) = s_1 - s_2 + \alpha$, are

$$E\begin{bmatrix} \bar{y}_{11.} \\ \bar{y}_{12.} \\ \bar{y}_{21.} \\ s_1 - s_2 \end{bmatrix} = \begin{bmatrix} \mu + s_1 + h_1 \\ \mu + s_1 + h_2 \\ \mu + s_2 + h_1 \\ s_1 - s_2 \end{bmatrix} + \begin{bmatrix} 0.2 \\ 0 \\ -0.2 \\ 0 \end{bmatrix} t$$

where $t = [\text{Var}(\bar{y}_{11.} - \bar{y}_{21.})/\sigma^2]^{-1}\alpha$. Consequently, $\hat{s}_1 - \hat{s}_2$ of (11) has expectation $s_1 - s_2 + .4t$, a biased predictor. Does an unbiased linear predictor exist? Let a predictor be $\mathbf{b}'\bar{\mathbf{y}}$. For unbiasedness, the following conditions are sufficient:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0.2 & 0 & -0.2 \end{bmatrix} \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ -01 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

No solution exists to $\mathbf{b}$ so no unbiased predictor exists.

4.2 Sires fixed and herds random

Now the conditional means given $E(\bar{y}_{11.} - \bar{y}_{21.}) = s_1 - s_2 + \alpha$ are

$$E\begin{bmatrix} \bar{y}_{11.} \\ \bar{y}_{12.} \\ \bar{y}_{21.} \\ s_1 - s_2 \end{bmatrix} = \begin{bmatrix} \mu + s_1 \\ \mu + s_1 \\ \mu + s_2 \\ s_1 - s_2 \end{bmatrix} + \begin{bmatrix} 0.2 \\ 0 \\ -0.2 \\ 0 \end{bmatrix} t.$$

Then the expectation of $\hat{s}_1 - \hat{s}_2$ from (12) is $s_1 - s_2 + 24t/65$, a biased predictor. If $\mathbf{b}'\bar{\mathbf{y}}$ is to be an unbiased predictor, a sufficient condition is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0.2 & 0 & -0.2 \end{bmatrix} \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}.$$

A unique solution exists. It is $\mathbf{b}' = [-1 \quad 2 \quad -1]$. Thus, a single unbiased predictor exists.

4.3 Sires random and herds fixed

The conditional means given $E(\bar{y}_{11.} - \bar{y}_{21.}) = \alpha$ are

$$E\begin{bmatrix} \bar{y}_{11.} \\ \bar{y}_{12.} \\ \bar{y}_{21.} \\ s_1 - s_2 \end{bmatrix} = \begin{bmatrix} \mu + h_1 \\ \mu + h_2 \\ \mu + h_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.3 \\ 0.1 \\ -0.3 \\ 0.2 \end{bmatrix} t.$$

From the solution in (13), $E(\hat{s}_1 - \hat{s}_2) = 0.2t$, an unbiased predictor. Note that if $s_1 - s_2$ had been predicted as though s were fixed as in (11), the expectation of the predictor would be $0.6t$. That is, the predictor of the difference is too large. It is easy to show that (13) is the only unbiased predictor.

4.4 Sires and herds random

The conditional means given $E(\bar{y}_{11.} - \bar{y}_{21.}) = \alpha$ are

$$E\begin{bmatrix} \bar{y}_{11.} \\ \bar{y}_{12.} \\ \bar{y}_{21.} \\ s_1 - s_2 \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \mu \\ 0 \end{bmatrix} + \begin{bmatrix} 0.3 \\ 0.1 \\ -0.3 \\ 0.2 \end{bmatrix} t.$$

From (14), $E(\hat{s}_1 - \hat{s}_2) = .2t$, an unbiased predictor. This is not the only one, the complete set being $\mathbf{b}'\mathbf{y}$ where $\mathbf{b}' = [1 + 2a \,\vdots\, -1 - 3a \,\vdots\, a]$. The solution of (14) has minimum variance as we shall see in the next section.

4.5 Selection on **s**

Suppose that instead of production in herd 1 the choice of sire to be used in herd 2 was based on some prior knowledge of the sires regarded as random in the unconditional distribution.

$$E(s_1 - s_2) = \alpha \neq 0.$$

Then the conditional means are

$$E\begin{bmatrix} \bar{y}_{11.} \\ \bar{y}_{12.} \\ \bar{y}_{21.} \\ s_1 - s_2 \end{bmatrix} = \begin{bmatrix} 0.1 \\ 0.1 \\ -0.1 \\ 0.2 \end{bmatrix} t$$

where $t = (2\sigma_s^2/\sigma^2)^{-1}\alpha$. Assuming herds random, the BLUP solution for the unconditional distribution of (14) gives $E(\hat{s}_1 - \hat{s}_2) = 0.0703t$, which is biased.

If instead we use the solution in (12) in which sires are regarded as fixed, $E(\hat{s}_1 - \hat{s}_2) = 0.2t$, an unbiased predictor. A general unbiased predictor is $[a \,\vdots\, 1 - a \,\vdots\, -1]\bar{\mathbf{y}}$.

5. BLUP IN A SELECTION MODEL

In this section, we derive the general solution to best linear unbiased prediction and estimation in a selection model. The examples of section 4 can all be described as a particular case of a model described in section 5.1.

5.1 A mixed model with selection

The unconditional model for **y** is the same as (1) but with the additional condition that the form of distribution is multivariate normal. We also invoke an additional vector variable **w**, jointly normally distributed with **y**, **u**, and **e**. The following additional parameters exist.

$$E(\mathbf{w}) = \mathbf{d}. \tag{15}$$

$$\mathrm{Cov}\left(\begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{e} \end{bmatrix}, \mathbf{w}'\right) = \begin{bmatrix} \mathbf{B} \\ \mathbf{B}_u \\ \mathbf{B}_e \end{bmatrix} \sigma^2 \tag{16}$$

where $\mathbf{B} = \mathbf{Z}\mathbf{B}_u + \mathbf{B}_e$.

$$\mathrm{Var}(\mathbf{w}) = \mathbf{H}\sigma^2. \tag{17}$$

Now suppose that some sort of selection on $\mathbf{w}$ has occurred such that in the selection model

$$\begin{aligned} E(\mathbf{w}) &= \mathbf{s} \neq \mathbf{d}, \\ \mathrm{Var}\ (\mathbf{w}) &= \mathbf{H}_s\sigma^2 \neq \mathbf{H}\sigma^2. \end{aligned} \tag{18}$$

For example, in the model of section 4.4,

$$\mathbf{w} = \bar{y}_{11.} - \bar{y}_{21.}\ ,$$

$$E(\mathbf{w}) = \mathbf{s} \neq 0,$$

$$\mathbf{H} = \mathrm{Var}\ (\bar{y}_{11.} - \bar{y}_{21.})/\sigma^2 = 0.6,$$

$$\mathbf{B}_u = \frac{1}{\sigma^2}\,\mathrm{Cov}\left[\begin{pmatrix} s_1 \\ s_2 \\ h_1 \\ h_2 \end{pmatrix}, (\bar{y}_{11.} - \bar{y}_{21.})\right] = \begin{bmatrix} 0.1 \\ -0.1 \\ 0 \\ 0 \end{bmatrix},$$

$$\mathbf{B}_e = \frac{1}{\sigma^2}\,\mathrm{Cov}\left[\begin{pmatrix} \bar{e}_{11.} \\ \bar{e}_{12.} \\ \bar{e}_{21.} \end{pmatrix}, (\bar{y}_{11.} - \bar{y}_{21.})\right] = \begin{bmatrix} 0.2 \\ 0 \\ -0.2 \end{bmatrix},$$

$$\mathbf{B} = \frac{1}{\sigma^2}\,\mathrm{Cov}\left[\begin{pmatrix} \bar{y}_{11.} \\ \bar{y}_{12.} \\ \bar{y}_{21.} \end{pmatrix}, (\bar{y}_{11.} - \bar{y}_{21.})\right] = \begin{bmatrix} 0.3 \\ 0.1 \\ -0.3 \end{bmatrix}.$$

Given these parameters of (1), (15), (16), (17), and (18), we can write the remaining parameters of the conditional (selection) model. For this purpose, we use a result due to Pearson [1903]. $[\mathbf{v}_1' \vdots \mathbf{v}_2']$ is normally distributed with mean $= [\boldsymbol{\mu}_1' \vdots \boldsymbol{\mu}_2']$ and variance

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

Then, if $\mathbf{v}_2$ is selected such that $E(\mathbf{v}_2) = \boldsymbol{\mu}_2 + \mathbf{k}$ and $\mathrm{Var}(\mathbf{v}_2) = \mathbf{C}_s$, the parameters of the conditional distribution are

$$E(\mathbf{v}_1) = \boldsymbol{\mu}_1 + \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{k}$$

$$\mathrm{Var}\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_0\mathbf{C}_{12}' & \vdots & \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_s \\ & \vdots & \\ \mathbf{C}_s\mathbf{C}_{22}^{-1}\mathbf{C}_{12}' & \vdots & \mathbf{C}_s \end{bmatrix} \tag{19}$$

where $\mathbf{C}_0 = \mathbf{C}_{22}^{-1}(\mathbf{C}_{22} - \mathbf{C}_s)\mathbf{C}_{22}^{-1}$. Note that if $\mathbf{v}_2$ is fixed and consequently $\mathbf{C}_s = \mathbf{0}$, we have the well-known result $\mathrm{Var}(\mathbf{v}_1) = \mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_{12}'$. Applying the result in (19) to our mixed model, the parameters of the conditional distribution are

$$E\begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} \mathbf{X}\boldsymbol{\beta} + \mathbf{B}\mathbf{t} \\ \mathbf{B}_u\mathbf{t} \\ \mathbf{s} \end{bmatrix} \tag{20}$$

where $\mathbf{t} = \mathbf{H}^{-1}(\mathbf{s} - \mathbf{d})$,

$$\mathrm{Var}\begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} \mathbf{V} - \mathbf{B}\mathbf{H}_0\mathbf{B}' & \mathbf{Z}\mathbf{G} - \mathbf{B}\mathbf{H}_0\mathbf{B}_u' & \mathbf{B}\mathbf{H}^{-1}\mathbf{H}_s \\ \mathbf{G}\mathbf{Z}' - \mathbf{B}_u\mathbf{H}_0\mathbf{B}' & \mathbf{G} - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u' & \mathbf{B}_u\mathbf{H}^{-1}\mathbf{H}_s \\ \mathbf{H}_s\mathbf{H}^{-1}\mathbf{B}' & \mathbf{H}_s\mathbf{H}^{-1}\mathbf{B}_u' & \mathbf{H}_s \end{bmatrix} \tag{21}$$

where $\mathbf{H}_0 = \mathbf{H}^{-1}(\mathbf{H} - \mathbf{H}_s)\mathbf{H}^{-1}$.

5.2 Best linear unbiased prediction

We wish to predict

$$\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u} + \mathbf{f}'(\mathbf{w} - \mathbf{d})$$

by $\mathbf{b}'\mathbf{y}$ such that the predictor is unbiased and has minimum variance of prediction error under the conditional model.

$$E[\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{u} + \mathbf{f}'(\mathbf{w} - \mathbf{d})] = \mathbf{k}'\boldsymbol{\beta} + \mathbf{m}'\mathbf{B}_u\mathbf{t} + \mathbf{f}'\mathbf{H}\mathbf{t}$$

$$E(\mathbf{b}'\mathbf{y}) = \mathbf{b}'\mathbf{X}\boldsymbol{\beta} + \mathbf{b}'\mathbf{B}\mathbf{t}$$

For unbiasedness, it is required that $\mathbf{X}'\mathbf{b} = \mathbf{k}$ and $\mathbf{B}'\mathbf{b} = \mathbf{B}_u'\mathbf{m} + \mathbf{H}\mathbf{f}$. The variance of the error of prediction is

$$\begin{aligned} \mathrm{Var}(\mathbf{b}'\mathbf{y} - \mathbf{m}'\mathbf{u} - \mathbf{f}'\mathbf{w}) = \{&\mathbf{b}'(\mathbf{V} - \mathbf{B}\mathbf{H}_0\mathbf{B}')\mathbf{b} - 2\mathbf{b}'(\mathbf{Z}\mathbf{G} - \mathbf{B}\mathbf{H}_0\mathbf{B}_u')\mathbf{m} \\ &- 2\mathbf{b}'\mathbf{B}\mathbf{H}^{-1}\mathbf{H}_s\mathbf{f} + \mathrm{Var}\,(\mathbf{m}'\mathbf{u} + \mathbf{f}'\mathbf{w})\}\sigma^2. \end{aligned}$$

This is minimized subject to unbiasedness by $\mathbf{b}$, the solution if it exists to (22),

$$\begin{bmatrix} \mathbf{V} - \mathbf{B}\mathbf{H}_0\mathbf{B}' & \mathbf{X} & \mathbf{B} \\ \mathbf{X}' & \mathbf{0} & \mathbf{0} \\ \mathbf{B}' & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \boldsymbol{\theta} \\ \boldsymbol{\phi} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}\mathbf{G}\mathbf{m} - \mathbf{B}\mathbf{H}_0\mathbf{B}_u'\mathbf{m} + \mathbf{B}\mathbf{H}^{-1}\mathbf{H}_s\mathbf{f} \\ \mathbf{k} \\ \mathbf{B}_u'\mathbf{m} + \mathbf{H}\mathbf{f} \end{bmatrix}. \tag{22}$$

From the third equation of (22), $\mathbf{B}'\mathbf{b} = \mathbf{B}_u'\mathbf{m} + \mathbf{H}\mathbf{f}$. Substituting this in the first equation of (22), we get

$$\begin{bmatrix} \mathbf{V} & \mathbf{X} & \mathbf{B} \\ \mathbf{X}' & \mathbf{0} & \mathbf{0} \\ \mathbf{B}' & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \boldsymbol{\theta} \\ \boldsymbol{\phi} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}\mathbf{G}\mathbf{m} + \mathbf{B}\mathbf{f} \\ \mathbf{k} \\ \mathbf{B}_u'\mathbf{m} + \mathbf{H}\mathbf{f} \end{bmatrix}. \tag{23}$$

This is a fortunate result because the predictor does not require knowledge of $\mathbf{s}$ and $\mathbf{H}_s$. Provided a solution exists, it is

$$\mathbf{b}'\mathbf{y} = \mathbf{k}'\hat{\boldsymbol{\beta}} + (\mathbf{m}'\mathbf{B}_u + \mathbf{f}'\mathbf{H})\hat{\mathbf{t}} + (\mathbf{m}'\mathbf{G}\mathbf{Z}' + \mathbf{f}'\mathbf{B}')\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{B}\hat{\mathbf{t}}) \tag{24}$$

where $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{t}}$ are any solution to (25),

$$\begin{bmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{V}^{-1}\mathbf{B} \\ \mathbf{B}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{B}'\mathbf{V}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ \mathbf{B}'\mathbf{V}^{-1}\mathbf{y} \end{bmatrix}. \tag{25}$$

From the second equation of (25),

$$\mathbf{B}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\beta} - \mathbf{B}\hat{\mathbf{t}}) = 0.$$

Substituting this in (24), we obtain

$$\mathbf{b}'\mathbf{y} = \mathbf{k}'\hat{\beta} + (\mathbf{m}'\mathbf{B}_u + \mathbf{f}'\mathbf{H})\hat{\mathbf{t}} + (\mathbf{m}'\mathbf{G}\mathbf{Z}'\mathbf{V}^{-1})(\mathbf{y} - \mathbf{X}\hat{\beta} - \mathbf{B}\hat{\mathbf{t}}). \tag{26}$$

Now from the matrix identities in (6) and (7),

$$\mathbf{b}'\mathbf{y} = \mathbf{k}'\hat{\beta} + (\mathbf{m}'\mathbf{B}_u + \mathbf{f}'\mathbf{H})\hat{\mathbf{t}} + \mathbf{m}'\hat{\mathbf{v}} \tag{27}$$

where $\hat{\beta}$, $\hat{\mathbf{t}}$, and $\hat{\mathbf{v}}$ are any solution to (28),

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{B} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{B} \\ \mathbf{B}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{B}'\mathbf{R}^{-1}\mathbf{Z} & \vdots & \mathbf{B}'\mathbf{R}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{v}} \\ \hat{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{B}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \cdot \tag{28}$$

The advantages of the method of (27) and (28) over (24) and (25) are the same as described for the advantages of (5) over (3) and (4).

A further modification of (28) that yields a direct solution to $\hat{\mathbf{u}}$, the best linear unbiased predictor of $\mathbf{u}$, can be accomplished as follows. Premultiply each side of equation (28) by

$$\mathbf{P} = \begin{bmatrix} \mathbf{I} & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & -\mathbf{B}_u' & \mathbf{I} \end{bmatrix} \cdot \tag{29}$$

This gives

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{B} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{B} \\ \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{Z} - \mathbf{B}_u\mathbf{G}^{-1} & \vdots & \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{v}} \\ \hat{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \cdot \tag{30}$$

Now insert $\mathbf{P}'(\mathbf{P}')^{-1} = \mathbf{I}$ between the coefficient matrix and $[\hat{\beta}' \quad \hat{\mathbf{v}}' \quad \hat{\mathbf{t}}']'$ of (30) and simplify. This gives

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{B}_e \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{B}_e - \mathbf{G}^{-1}\mathbf{B}_u \\ \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{Z} - \mathbf{B}_u'\mathbf{G}^{-1} & \vdots & \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{B}_e + \mathbf{B}_u'\mathbf{G}^{-1}\mathbf{B}_u \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{v}} + \mathbf{B}_u\hat{\mathbf{t}} \\ \hat{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \cdot \tag{31}$$

Now define $\hat{\mathbf{v}} + \mathbf{B}_u\hat{\mathbf{t}}$ as $\hat{\mathbf{u}}$ and substitute this in (27) and (31). This is a logical definition since $E(\hat{\mathbf{u}}) = \mathbf{B}_u\mathbf{t}$. Now (27) can be written as

$$\begin{aligned} \mathbf{b}'\mathbf{y} &= \mathbf{k}'\hat{\beta} + \mathbf{m}'(\hat{\mathbf{v}} + \mathbf{B}_u\hat{\mathbf{t}}) + \mathbf{f}'\mathbf{H}\hat{\mathbf{t}} \\ &= \mathbf{k}'\hat{\beta} + \mathbf{m}'\hat{\mathbf{u}} + \mathbf{f}'\mathbf{H}\hat{\mathbf{t}} \end{aligned} \tag{32}$$

where $\hat{\beta}$, $\hat{\mathbf{u}}$, and $\hat{\mathbf{t}}$ are solutions to (31) with $\hat{\mathbf{u}} = \hat{\mathbf{v}} + B_u\hat{\mathbf{t}}$.

5.3 Estimability

In the no selection case, the solution to $\hat{\mathbf{u}}$ is always unique, but β is estimable only if $\mathbf{X}$ has full column rank. In the selection case, however, $\mathbf{u}$ is not necessarily predictable. The requirement for this is that $\mathbf{B}_u\mathbf{t}$ be estimable. Of course, even if this is not true, there may be certain linear functions of interest, say $\mathbf{m}'\mathbf{u}$, that are predictable. This will be true provided $\mathbf{m}'\mathbf{B}_u\mathbf{t}$ is estimable. One way to check estimability is to determine if $[\mathbf{k}' \vdots \mathbf{m}' \vdots \mathbf{f}'\mathbf{H}]'E[\hat{\boldsymbol{\beta}}' \quad \hat{\mathbf{u}}' \quad \hat{\mathbf{t}}']' = \mathbf{k}'\boldsymbol{\beta} + (\mathbf{m}'\mathbf{B}_u + \mathbf{f}'\mathbf{H})\mathbf{t}$ where

$$E\begin{bmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{u}}\\ \hat{\mathbf{t}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\\ \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{X}\end{bmatrix}\boldsymbol{\beta} + \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{B}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{B}\\ \mathbf{B}_e'\mathbf{R}^{-1}\mathbf{B}\end{bmatrix}\hat{\mathbf{t}}$$

premultiplied by the g-inverse of the coefficient matrix of (31) used in the solution to $[\hat{\boldsymbol{\beta}}' \vdots \hat{\mathbf{u}}' \vdots \hat{\mathbf{t}}']$.

5.4 Distributional properties of estimators and predictors

Let some symmetric g-inverse of the coefficient matrix in (31) be

$$\begin{bmatrix}\mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13}\\ \mathbf{C}_{12}' & \mathbf{C}_{22} & \mathbf{C}_{23}\\ \mathbf{C}_{13}' & \mathbf{C}_{23}' & \mathbf{C}_{33}\end{bmatrix}. \tag{33}$$

Then, $\mathrm{Var}(\mathbf{K}'\hat{\boldsymbol{\beta}} + \mathbf{M}'\hat{\mathbf{u}} + \mathbf{F}'\hat{\mathbf{t}})/\sigma^2$ is

$$[\mathbf{K}' \vdots \mathbf{M}' \vdots \mathbf{F}'] \cdot \begin{bmatrix}\mathbf{C}_{11} & \vdots & \mathbf{C}_{13}\mathbf{B}_u' & \vdots & \mathbf{C}_{13}\\ \mathbf{B}_u\mathbf{C}_{13}' & \vdots & \mathbf{G} - \mathbf{C}_{22} + \mathbf{C}_{23}\mathbf{B}_u' + \mathbf{B}_u\mathbf{C}_{23}' - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u' & \vdots & \mathbf{B}_u\mathbf{C}_{33} - \mathbf{B}_u\mathbf{H}_0\\ \mathbf{C}_{13}' & \vdots & \mathbf{C}_{33}\mathbf{B}_u' - \mathbf{H}_0\mathbf{B}_u' & \vdots & \mathbf{C}_{33} - \mathbf{H}_0\end{bmatrix}\begin{bmatrix}\mathbf{K}\\ \mathbf{M}\\ \mathbf{F}\end{bmatrix}. \tag{34}$$

$\mathrm{Cov}\,([\mathbf{K}'\hat{\boldsymbol{\beta}} + \mathbf{M}'\hat{\mathbf{u}} + \mathbf{F}'\hat{\mathbf{t}}], [\mathbf{u}' \vdots \mathbf{w}'])/\sigma^2$

$$= [\mathbf{K}' \vdots \mathbf{M}' \vdots \mathbf{F}']\begin{bmatrix}-\mathbf{C}_{12} + \mathbf{C}_{13}\mathbf{B}_u' & \vdots & \mathbf{0}\\ \mathbf{G} - \mathbf{C}_{22} + \mathbf{C}_{23}\mathbf{B}_u' - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u' & \vdots & \mathbf{B}_u\mathbf{H}^{-1}\mathbf{H}_s\\ -\mathbf{C}_{23}' + \mathbf{C}_{33}\mathbf{B}_u' - \mathbf{H}_0\mathbf{B}_u' & \vdots & \mathbf{H}^{-1}\mathbf{H}_s\end{bmatrix}. \tag{35}$$

Note that in contrast to the no selection case $\mathrm{Var}(\hat{\mathbf{u}}) \neq \mathrm{Cov}(\hat{\mathbf{u}}, \mathbf{u}')$. From (21),

$$\frac{1}{\sigma^2}\mathrm{Var}\begin{bmatrix}\mathbf{u}\\ \mathbf{w}\end{bmatrix} = \begin{bmatrix}\mathbf{G} - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u' & \vdots & \mathbf{B}_u\mathbf{H}^{-1}\mathbf{H}_s\\ \mathbf{H}_s\mathbf{H}^{-1}\mathbf{B}_u' & \vdots & \mathbf{H}_s\end{bmatrix}. \tag{36}$$

Combining the results of (34), (35), and (36), we obtain the results of (37) for variances of prediction errors where $\mathbf{H}\hat{\mathbf{t}}$ is the predictor of $\mathbf{w} - \mathbf{d}$ in the full rank case. The variance of error of prediction of $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u} + \mathbf{F}'(\mathbf{w} - \mathbf{d})$ is

$$[\mathbf{K}' \vdots \mathbf{M}' \vdots \mathbf{F}']\begin{bmatrix}\mathbf{C}_{11} & \vdots & \mathbf{C}_{12} & \vdots & \mathbf{C}_{13}\mathbf{H}\\ \mathbf{C}_{12}' & \vdots & \mathbf{C}_{22} & \vdots & \mathbf{C}_{23}\mathbf{H}\\ \mathbf{H}\mathbf{C}_{13}' & \vdots & \mathbf{H}\mathbf{C}_{23}' & \vdots & \mathbf{H}\mathbf{C}_{33}\mathbf{H} - \mathbf{H}\end{bmatrix}\begin{bmatrix}\mathbf{K}\\ \mathbf{M}\\ \mathbf{F}\end{bmatrix}\sigma^2. \tag{37}$$

From (31) and (37), we now have the fortunate result that best linear unbiased predictors and the prediction error variances require no knowledge of the intensity and not even the direction of selection. In contrast, it can be seen that the correlation between $\mathbf{u}_i$ and $\hat{\mathbf{u}}_i$ cannot be computed unless $\mathbf{H}_s$ is known. The results in (34) and (35) are proved for the full rank case in Appendix B.

6. SOME APPLICATIONS

6.1 *Selection on* y

The most common type of selection involves selection on $\mathbf{y}$. For this case, let $\mathbf{w} = \mathbf{L}'\mathbf{y}$ where $\mathbf{L}'$ has full row rank. Then

$$\begin{aligned}
\mathbf{B}_u &= \text{Cov}(\mathbf{u}, \mathbf{y}'\mathbf{L})/\sigma^2 = \mathbf{GZ}'\mathbf{L},\\
\mathbf{B}_e &= \text{Cov}(\mathbf{e}, \mathbf{y}'\mathbf{L})/\sigma^2 = \mathbf{RL},\\
\mathbf{B} &= \text{Cov}(\mathbf{y}, \mathbf{y}'\mathbf{L})/\sigma^2 = \mathbf{VL},\\
\mathbf{H} &= \text{Var}(\mathbf{L}'\mathbf{y})/\sigma^2 = \mathbf{L}'\mathbf{VL}.
\end{aligned}$$

Substituting these in (31), the BLUP equations for the selection model are

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{L} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{L}'\mathbf{X} & \mathbf{0} & \mathbf{L}'\mathbf{VL} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{L}'\mathbf{y} \end{bmatrix}. \tag{38}$$

It should be noted in particular that if $\mathbf{L}'\mathbf{X} = \mathbf{0}$, $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ are any solution to the BLUP equations in the no selection model. [See Eq. (5)].

We illustrate $\mathbf{L}'\mathbf{y}$ type of selection with a sire example. Suppose we have the following progeny from 4 sires in 2 herds.

		Herd	
Group	Sire	1	2
1	1	5	20
	2	3	0
2	3	4	0
	4	6	0

The model assumed is

$$y_{ijkl} = g_i + s_{ij} + h_k + e_{ijkl}$$

where $\mathbf{g}$ is fixed; $\mathbf{s}$, $\mathbf{h}$, and $\mathbf{e}$ are normally distributed with null means, and

$$\text{Var}\begin{bmatrix} \mathbf{s} \\ \mathbf{h} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} 0.1\mathbf{I} & 0 & 0 \\ 0 & 0.5\mathbf{I} & 0 \\ 0 & 0 & \mathbf{I} \end{bmatrix} \sigma^2.$$

Now we assume that sire 1 was selected to have progeny in herd 2 because his progeny were better in herd 1 than were those of the other sires. In terms of the model,

$$E(g_1 + s_1 + h_1 + \bar{e}_{111.}) > E(g_1 + s_2 + h_1 + \bar{e}_{121.})$$

or $E[(s_1 + \bar{e}_{111.}) - (s_2 + \bar{e}_{121.})] > 0$,

$$E(g_1 + s_1 + h_1 + \bar{e}_{111.}) > E(g_2 + s_3 + h_1 + \bar{e}_{231.})$$

or $E[(s_1 + \bar{e}_{111.}) - (s_3 + \bar{e}_{231.})] > (g_2 - g_1)$.

Also $E[(s_1 + \bar{e}_{111.}) - (s_4 + \bar{e}_{241.})] > (g_2 - g_1)$.

Rewrite the model as

$$\bar{y}_{ijk.} = g_i + s_{ij} + h_k + \bar{e}_{ijk.}$$

where **s** and **h** have the same distributional properties as before and

$$\mathrm{Var}\,(\bar{\mathbf{e}}) = \begin{bmatrix} 5 & & & & & \\ & 20 & & & 0 & \\ & & 3 & & & \\ & & & 4 & & \\ & 0 & & & & \\ & & & & & 6 \end{bmatrix}^{-1} \sigma^2.$$

Now we can describe the selection as

$$\mathbf{L}'\bar{\mathbf{y}} = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \bar{y}_{111.} \\ \vdots \\ \bar{y}_{241.} \end{bmatrix}.$$

$$\mathbf{V} = \begin{bmatrix} 0.8 & 0.1 & 0.5 & 0.5 & 0.5 \\ 0.1 & 0.65 & 0 & 0 & 0 \\ 0.5 & 0 & 0.933\ldots & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 & 0.85 & 0.5 \\ 0.5 & 0 & 0.5 & 0.5 & 0.766\ldots \end{bmatrix}.$$

$$\mathbf{L}'\mathbf{V}\mathbf{L} = \begin{bmatrix} 0.733\ldots & 0.3 & 0.3 \\ 0.3 & 0.65 & 0.3 \\ 0.3 & 0.3 & 0.566\ldots \end{bmatrix}.$$

$$\mathbf{L}'\mathbf{X} = \mathbf{L}' \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & -1 \\ 1 & -1 \end{bmatrix}.$$

Then the BLUP equations according to (38) are

$$\begin{bmatrix} 28 & 0 & 25 & 3 & 0 & 0 & 8 & 20 & 0 & 1 & 1 \\ 0 & 10 & 0 & 0 & 4 & 6 & 10 & 0 & 0 & -1 & -1 \\ 25 & 0 & 35 & 0 & 0 & 0 & 5 & 20 & 0 & 0 & 0 \\ 3 & 0 & 0 & 13 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 14 & 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 & 16 & 6 & 0 & 0 & 0 & 0 \\ 8 & 10 & 5 & 3 & 4 & 6 & 20 & 0 & 0 & 0 & 0 \\ 20 & 0 & 20 & 0 & 0 & 0 & 0 & 22 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.733\ldots & 0.3 & 0.3 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0.65 & 0.3 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0.3 & 0.566\ldots \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ h_1 \\ h_2 \\ t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} 5 & 20 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 6 \\ 5 & 20 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 6 \\ 5 & 0 & 3 & 4 & 6 \\ 0 & 20 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} \bar{\mathbf{y}}. \tag{39}$$

The sires are evaluated by $\hat{g}_i + \hat{s}_{ij}$. The solution is

$$\begin{bmatrix} \hat{g}_1 + \hat{s}_1 \\ \hat{g}_1 + \hat{s}_2 \\ \hat{g}_2 + \hat{s}_3 \\ \hat{g}_2 + \hat{s}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -0.3846 & 1.1538 & 0.2308 & 0 & 0 \\ 0.6653 & -0.9148 & 0.2495 & 0.5946 & 0.4054 \\ 0.5821 & -0.8004 & 0.2183 & 0.2703 & 0.7297 \end{bmatrix} \bar{\mathbf{y}}. \tag{40}$$

To check that this solution is unbiased, note that

$$E(\bar{\mathbf{y}}) = \begin{bmatrix} g_1 \\ g_1 \\ g_1 \\ g_2 \\ g_2 \end{bmatrix} + \begin{bmatrix} 0.3 & 0.3 & 0.3 \\ 0.1 & 0.1 & 0.1 \\ -0.433\ldots & 0 & 0 \\ 0 & -0.35 & 0 \\ 0 & 0 & -0.266\ldots \end{bmatrix} \mathbf{t}.$$

Using this result and the solution of (40),

$$E \begin{bmatrix} \hat{g}_1 + \hat{s}_1 \\ \hat{g}_1 + \hat{s}_2 \\ \hat{g}_2 + \hat{s}_3 \\ \hat{g}_2 + \hat{s}_4 \end{bmatrix} = \begin{bmatrix} g_1 \\ g_1 \\ g_2 \\ g_2 \end{bmatrix} + \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ -0.1 & 0 & 0 \\ 0 & -0.1 & 0 \\ 0 & 0 & -0.1 \end{bmatrix} \mathbf{t},$$

which is unbiased because this is the expectation of $\mathbf{g} + \mathbf{s}$. Observe that if the model had been

$$y_{ijk} = \mu + s_i + h_j + e_{ijk} ,$$

$\mathbf{X}' = [1 \cdots 1]$ and $\mathbf{L}'\mathbf{X} = \mathbf{0}$. Consequently, regular BLUP would be unbiased.

6.2 Selection on $\mathbf{u}$

In some cases selection may have occurred on $\mathbf{u}$. In this case, let $\mathbf{w} = \mathbf{L}'\mathbf{u}$. Then

$$\begin{aligned} \mathbf{B}_u &= \text{Cov}\,(\mathbf{u}, \mathbf{u}'\mathbf{L})/\sigma^2 = \mathbf{GL}, \\ \mathbf{B}_e &= \text{Cov}\,(\mathbf{e}, \mathbf{u}'\mathbf{L})/\sigma^2 = \mathbf{0}, \\ \mathbf{B} &= \text{Cov}\,(\mathbf{y}, \mathbf{u}'\mathbf{L})/\sigma^2 = \mathbf{ZGL}, \\ \mathbf{H} &= \text{Var}\,(\mathbf{L}'\mathbf{u})/\sigma^2 = \mathbf{L}'\mathbf{GL}. \end{aligned}$$

Substituting in (31), we get as the BLUP equations

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \vdots & \mathbf{0} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \vdots & -\mathbf{L} \\ \mathbf{0} & \vdots & -\mathbf{L}' & \vdots & \mathbf{L}'\mathbf{GL} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{0} \end{bmatrix}. \tag{41}$$

As an example of this, suppose we have progeny on 6 sires numbering 5, 8, 3, 2, 4, and 2. The model is

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

where μ is fixed, s_i and e_{ij} are normally distributed with means $= 0$,

$$\text{Var (s)} = \frac{1}{9}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}\sigma^2,$$

$\text{Var}(\mathbf{e}) = \mathbf{I}\sigma^2$, and $\text{Cov}(\mathbf{s}, \mathbf{e}') = \mathbf{0}$. Var(s) might imply an additive genetic model with all sires unrelated except 2 and 3 which are full sibs. Now suppose we have prior information on the sires that makes us believe that $s_1 + s_2 > s_3 + s_4$ and $s_1 + s_2 > s_5 + s_6$. Then,

$$\mathbf{L}'\mathbf{u} = \begin{bmatrix} 1 & 1 & -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & -1 \end{bmatrix}\mathbf{u}.$$

The equations of (41) become

$$\begin{bmatrix} 24 & 5 & 8 & 3 & 2 & 4 & 2 & 0 & 0 \\ 5 & 14 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 8 & 0 & 20 & -6 & 0 & 0 & 0 & -1 & -1 \\ 3 & 0 & -6 & 15 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 11 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 0 & 13 & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 & 0 & 11 & 0 & 1 \\ 0 & -1 & -1 & 1 & 1 & 0 & 0 & 0.33\ldots & 0.166\ldots \\ 0 & -1 & -1 & 0 & 0 & 1 & 1 & 0.166\ldots & 0.44\ldots \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{s}_4 \\ \hat{s}_5 \\ \hat{s}_6 \\ \hat{t}_1 \\ \hat{t}_2 \end{bmatrix} = \begin{bmatrix} 5 & 8 & 3 & 2 & 4 & 2 \\ 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\bar{\mathbf{y}}$$

where $\bar{\mathbf{y}}' = [\bar{y}_{1.} \cdots \bar{y}_{6.}]$. The solution is approximately

$$\begin{bmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{s}_4 \\ \hat{s}_5 \\ \hat{s}_6 \end{bmatrix} = \begin{bmatrix} 0.134 & 0.180 & 0.148 & 0.150 & 0.244 & 0.144 \\ 0.594 & 0.165 & -0.195 & -0.226 & -0.213 & -0.125 \\ 0.082 & 0.519 & -0.043 & -0.138 & -0.264 & -0.156 \\ -0.213 & 0.101 & 0.388 & 0.143 & -0.263 & -0.156 \\ -0.323 & -0.129 & 0.292 & 0.466 & -0.192 & -0.114 \\ -0.095 & -0.221 & -0.162 & -0.125 & 0.493 & 0.110 \\ -0.088 & -0.228 & -0.165 & -0.120 & 0.264 & 0.338 \end{bmatrix} \bar{\mathbf{y}}.$$

$$E(\bar{\mathbf{y}}) = \begin{bmatrix} \mu \\ \mu \\ \mu \\ \mu \\ \mu \\ \mu \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0.5 & 1 \\ -0.5 & 0.5 \\ -1 & 0 \\ 0 & -1 \\ 0 & -1 \end{bmatrix} \mathbf{t}/9.$$

Then,

$$E \begin{bmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{s}_4 \\ \hat{s}_5 \\ \hat{s}_6 \end{bmatrix} = \begin{bmatrix} \mu \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0.5 & 1 \\ -0.5 & 0.5 \\ -1 & 0 \\ 0 & -1 \\ 0 & -1 \end{bmatrix} \mathbf{t}/9,$$

which is the expectation of

$$\begin{bmatrix} \mu \\ \mathbf{s} \end{bmatrix}.$$

In some applications, equations (41) may be difficult to write or to solve. An unbiased but not minimum prediction error variance method is available. Suppose selection has the form $\mathbf{L}_1'\mathbf{u}_1$ where $\mathbf{u}' = [\mathbf{u}_1' \vdots \mathbf{u}_2']$ and

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

Then it is easy to prove that a solution to equations of the form of (5) with $\mathbf{u}_1$ regarded as fixed yield these results: 1) $\hat{\mathbf{u}}_2$ is a unique, unbiased predictor of $\mathbf{u}_2$ and 2) $\mathbf{k}'\hat{\boldsymbol{\beta}} + \mathbf{m}_1'\hat{\mathbf{u}}_1$ is a unique, unbiased predictor of $\mathbf{k}'\boldsymbol{\beta} + \mathbf{m}_1'\mathbf{u}_1$ provided this is an estimable function under a fixed $\mathbf{u}_1$ model.

6.3 Selection on **e**

In this case, let $\mathbf{w} = \mathbf{L}'\mathbf{e}$. Then,

$$\mathbf{B}_u = \text{Cov}(\mathbf{u}, \mathbf{e}'\mathbf{L})/\sigma^2 = \mathbf{0},$$
$$\mathbf{B}_e = \text{Cov}(\mathbf{e}, \mathbf{e}'\mathbf{L})/\sigma^2 = \mathbf{RL},$$
$$\mathbf{B} = \text{Cov}(\mathbf{y}, \mathbf{e}'\mathbf{L})/\sigma^2 = \mathbf{RL},$$
$$\mathbf{H} = \text{Var}(\mathbf{L}'\mathbf{e})/\sigma^2 = \mathbf{L}'\mathbf{RL}.$$

Substituting these in (31), we get

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \vdots & \mathbf{X}'\mathbf{L} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \vdots & \mathbf{Z}'\mathbf{L} \\ \mathbf{L}'\mathbf{X} & \vdots & \mathbf{L}'\mathbf{Z} & \vdots & \mathbf{L}'\mathbf{RL} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{L}'\mathbf{y} \end{bmatrix}. \tag{42}$$

As an example, suppose that 4 sires have progeny numbering 5, 10, 5, and 4. The model is

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

where **s** and **e** are normally distributed with means zero and

$$\text{Var}\begin{bmatrix} \mathbf{s} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} .1\mathbf{I} & 0 \\ 0 & \mathbf{I} \end{bmatrix}\sigma^2.$$

We suspect that the $\bar{e}_i$ associated with sires 1 and 2 are better than those for 3 and 4 because of differential selection of mates. Then,

$$\mathbf{L}'\bar{\mathbf{e}} = [1 \quad 1 \quad -1 \quad -1]\bar{\mathbf{e}}$$

where $\bar{\mathbf{e}}' = [\bar{e}_{1.} \cdots \bar{e}_{4.}]$. Then the equations according to (42) are

$$\begin{bmatrix} 24 & 5 & 10 & 5 & 4 & 0 \\ 5 & 15 & 0 & 0 & 0 & 1 \\ 10 & 0 & 20 & 0 & 0 & 1 \\ 5 & 0 & 0 & 15 & 0 & -1 \\ 4 & 0 & 0 & 0 & 14 & -1 \\ 0 & 1 & 1 & -1 & -1 & 0.75 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{s}_4 \\ \hat{t} \end{bmatrix} = \begin{bmatrix} 5 & 10 & 5 & 4 \\ 5 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 4 \\ 1 & 1 & -1 & -1 \end{bmatrix} \bar{\mathbf{y}}$$

where $\bar{\mathbf{y}}' = [\bar{y}_{1.} \cdots \bar{y}_{4.}]$. The solution is approximately

$$\begin{bmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{s}_4 \end{bmatrix} = \begin{bmatrix} 0.251 & 0.361 & 0.211 & 0.177 \\ 0.153 & -0.198 & 0.013 & 0.032 \\ -0.198 & 0.261 & -0.043 & -0.020 \\ 0.013 & -0.043 & 0.180 & -0.150 \\ 0.032 & -0.020 & -0.150 & 0.138 \end{bmatrix} \bar{\mathbf{y}},$$

$$E(\bar{\mathbf{y}}) = \begin{bmatrix} \mu \\ \mu \\ \mu \\ \mu \end{bmatrix} + \begin{bmatrix} 0.2 \\ 0.1 \\ -0.2 \\ -0.25 \end{bmatrix} \mathbf{t}.$$

Then,

$$E \begin{bmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{s}_4 \end{bmatrix} = \begin{bmatrix} \mu \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \mathbf{t},$$

as it should be for unbiasedness.

6.4 Sequential selection

The foregoing methods apply to sequential selection. To illustrate, suppose we wish to select for genetic merit of trait 1. We observe traits 1 and 2 on several animals. A selection index for trait 1 is constructed and a certain fraction of animals is selected. Later trait 3 is observed in the survivors of the initial selection. What is the best predictor of trait 3 and what are its distributional properties?

Suppose the model for traits 1, 2, and 3 is

$$\mathbf{y} = \mathbf{Ig} + \mathbf{Ie}$$

where $\mathbf{y}$, $\mathbf{g}$, and $\mathbf{e}$ are column vectors of order 3; $E(\mathbf{g}) = E(\mathbf{e}) = \mathbf{0}$;

$$\text{Var}\ (\mathbf{e}) = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \sigma^2;$$

$$\text{Var}\ (\mathbf{g}) = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 6 & 3 \\ -1 & 3 & 7 \end{bmatrix} \sigma^2.$$

Note that $E(\mathbf{y}) = \mathbf{0}$, a usual assumption in selection index applications but not a very realistic one.

Now, given a sample $[\bar{y}_1 \quad \bar{y}_2]$, the best predictor of $[g_1 \quad g_2]'$ is

$$\begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \end{bmatrix} = (\mathbf{R}_2^{-1} + \mathbf{G}_2^{-1})^{-1}\mathbf{R}_2^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{79} \begin{bmatrix} 48 & 11 \\ 14 & 46 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

where $\mathbf{R}_2$ and $\mathbf{G}_2$ refer to 2×2 submatrices of $\mathbf{R}$ and $\mathbf{G}$. Now we select on the basis of

$$\frac{1}{79} [48 \quad 11] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

such that the expectation of this function is not $\mathbf{0}$. That is, $\mathbf{L}' = [48 \quad 11]/79$. Clearly $\mathbf{L}'\mathbf{X}$ does not exist since $\mathbf{X}$ does not exist. Consequently, BLUP of $\mathbf{g}$, given $[y_1 \quad y_2 \quad y_3]$, is the same solution as in the unconditional model. That is,

$$\begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \\ \hat{g}_3 \end{bmatrix} = (\mathbf{R}^{-1} + \mathbf{G}^{-1})^{-1}\mathbf{R}^{-1}\mathbf{y}$$

$$= \begin{bmatrix} 1.2983 & -0.0305 & 0.4203 \\ -0.0305 & 2.1508 & 1.3661 \\ 0.4203 & 1.3661 & 1.9559 \end{bmatrix} \begin{bmatrix} 0.6154 & 0.3846 & -0.4615 \\ 0.3846 & 0.6154 & -0.5385 \\ -0.4615 & -0.5385 & 0.8462 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

The variance of the prediction error of g_1 is $1.2983\sigma^2$ from (37). Let us examine the correlation between $\hat{g}_1$ and g_1. For this, we need $\mathbf{B}_u$ and $\mathbf{H}_0$.

$$\mathbf{B}_u = \mathrm{Cov}\,(g_1, \mathbf{L}'\mathbf{y})/\sigma^2 = [5 \quad 2 \quad -1] \begin{bmatrix} 48 \\ 11 \\ 0 \end{bmatrix} \frac{1}{79} = 3.3165;$$

$$\mathbf{H} = \mathrm{Var}\,(\mathbf{L}'\mathbf{y})/\sigma^2 = 262/79 = 3.3165.$$

Let $\mathbf{H}_s = k\mathbf{H}$ where k is in the range 0 to 1. Then $\mathbf{H}_0 = \mathbf{H}^{-1}(\mathbf{H} - \mathbf{H}_s)\mathbf{H}^{-1} = 0.3015(1 - k)$. From (34),

$$\begin{aligned} \mathrm{Var}\,(\hat{g}_1)/\sigma &= 5 - 1.2983 - (3.3165)^2(0.3015)(1 - k) \\ &= 0.3853 + 3.3165k \end{aligned}$$

This also is $\mathrm{Cov}(\hat{g}_1, g_1)/\sigma^2$.

$$\begin{aligned} \mathrm{Var}\,(g_1)/\sigma^2 &= 5 - 3.3165(1 - k) \\ &= 1.6835 + 3.3165k \end{aligned}$$

Therefore,

$$r_{\hat{g}_1, g_1} = \sqrt{\frac{0.3853 + 3.3165k}{1.6835 + 3.3165k}}.$$

As would be expected, the more intense selection is in stage 1; that is, the smaller k is, the smaller the correlation is.

It should be noted that the results of this paper do not necessarily apply to more than two stage selection because the selected variates are not normally distributed and consequently the results of (21) do not apply.

7. BIAS DUE TO INCORRECT VARIANCES

It should be recognized that the equations of (31) give unbiased estimates and predictions under the assumption that the correct $\mathbf{R}$, $\mathbf{G}$, $\mathbf{B}_e$, and $\mathbf{B}_u$ are used. Statements of the bias due to using some other values can be written but are hard to interpret. In the case of selection on $\mathbf{y}$ and with $\mathbf{L}'\mathbf{X} = \mathbf{0}$, some fairly simple results can be stated.

Suppose that $\mathbf{R} = \mathbf{I}$ as is often the case, and we use an estimate of $\mathbf{G}$, say $\hat{\mathbf{G}}$. Then the BLUP equations for estimation of $\boldsymbol{\beta}$ and prediction of $\mathbf{u}$ are

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \vdots & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \vdots & \mathbf{Z}'\mathbf{Z} + \hat{\mathbf{G}}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}. \tag{43}$$

Now under the distribution conditional on $\mathbf{L}'\mathbf{y}$,

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} + \mathbf{ZGZ}')\mathbf{L}\hat{\mathbf{t}},$$

$$E(\mathbf{u}) = \mathbf{GZ}'\mathbf{L}\hat{\mathbf{t}}.$$

Assume that $\mathbf{X}$ is full rank. Then, an inverse of the coefficient matrix of (43) is

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

$$E(\hat{\boldsymbol{\beta}}) = (\mathbf{C}_{11}\mathbf{X}' + \mathbf{C}_{12}\mathbf{Z}')[\mathbf{X}\boldsymbol{\beta} + (\mathbf{I} + \mathbf{ZGZ}')\mathbf{Lt}].$$

Making use of $\mathbf{L}'\mathbf{X} = \mathbf{0}$, this simplifies to

$$\boldsymbol{\beta} + \mathbf{C}_{12}(\mathbf{G}^{-1} - \hat{\mathbf{G}}^{-1})\mathbf{GZ}'\mathbf{Lt}$$

which, of course, reduces to $\boldsymbol{\beta}$ if $\mathbf{G} = \hat{\mathbf{G}}$ or if no selection has been practiced. Note further that if regular least squares were used, $\hat{\mathbf{G}}^{-1} = \mathbf{0}$ and the bias is $\mathbf{C}_{12}\mathbf{Z}'\mathbf{Lt}$, assuming in this last case that $\boldsymbol{\beta}$ and $\mathbf{u}$ are estimable under a fixed $\mathbf{u}$ model. Similarly,

$$E(\hat{\mathbf{u}}) = \mathbf{GZ}'\mathbf{Lt} + \mathbf{C}_{22}(\mathbf{G}^{-1} - \hat{\mathbf{G}}^{-1})\mathbf{GZ}'\mathbf{Lt}$$

which reduces to the correct expectation, $\mathbf{GZ}'\mathbf{Lt}$, if $\mathbf{G} = \hat{\mathbf{G}}$ or if no selection had been practiced.

MEILLEUR ESTIMATEUR LINEAIRE SANS BIAIS (BLUE) ET PREDICTEUR SOUS UN MODELE DE SELECTION

RESUME

Les modèles linéaires mixtes sont choisis dans la plupart des applications sur lignées animales. Sont disponibles des méthodes de calcul:
—des BLUE des fonctions linéaires estimables des éléments fixés du modèle,
—des prédicteurs linéaires sans biais des éléments aléatoires du modèle.
La plupart des données relatives aux lignées animales ne satisfont cependant pas aux conditions usuelles de l'échantillonnage aléatoire. Le problème est que les données proviennent soit d'une expérimentation de sélection, soit de troupeaux qui subissent une sélection.

Par conséquent, les méthodes usuelles donnent vraisemblablement des estimateurs et des prédicteurs biaisés. On présente, dans ce papier, des méthodes pour traiter de telles données.

REFERENCES

Henderson, C. R. [1949]. Estimates of changes in herd environment. *J. Dairy Sci. 32*, 706.

Henderson, C. R. [1950]. Estimation of genetic parameters. *Ann. Math. Stat. 21*, 309.

Henderson, C. R. [1963]. Selection index and expected genetic advance. In: Statistical Genetics and Plant Breeding. Hanson, W. D. and Robinson, H. F. (Eds.), pp. 141–63. National Academy of Sciences—National Research Council, Washington. Publication 982.

Henderson, C. R., Kempthorne, O., Searle, S. R., and von Krosigk, C. M. [1959]. The estimation of environmental and genetic trends from records subject to culling. *Biometrics 15*, 192.

Lush, J. L. and Shrode, R. R. [1950]. Changes in milk production with age and milking frequency. *J. Dairy Sci. 33*, 338.

Pearson, K. [1903]. Mathematical contributions to the theory of evolution.—XI. On the influence of natural selection on the variability and correlation of organs. *Philosophical Transactions of the Royal Society of London, A, 200,* 1.
Snedecor, G. W. [1946]. *Statistical Methods.* Collegiate Press, Ames, Iowa. 4th ed.

APPENDIX A

Derivation of Variance-Covariance Matrix of Estimators, Predictors, and Prediction Errors in the No Selection Model.

Let a regular inverse of the coefficient matrix of (5) be

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

Then

$$\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{11} & \vdots & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \vdots & \mathbf{C}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1' \\ \mathbf{Q}_2' \end{bmatrix} \bar{\mathbf{y}} \text{ , say.}$$

The following identity is useful,

$$\begin{bmatrix} \mathbf{Q}_1' \\ \mathbf{Q}_2' \end{bmatrix} [\mathbf{X} \vdots \mathbf{Z}] = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{bmatrix} \left(\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \vdots & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} \end{bmatrix} \right)$$

$$= \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{C}_{12}\mathbf{G}^{-1} \\ \mathbf{0} & \mathbf{C}_{22}\mathbf{G}^{-1} \end{bmatrix}.$$

That is,

$$\mathbf{Q}_1'\mathbf{X} = \mathbf{I},$$

$$\mathbf{Q}_1'\mathbf{Z} = -\mathbf{C}_{12}\mathbf{G}^{-1},$$

$$\mathbf{Q}_2'\mathbf{X} = \mathbf{0},$$

$$\mathbf{Q}_2'\mathbf{Z} = \mathbf{I} - \mathbf{C}_{22}\mathbf{G}^{-1}.$$

Using these results,

$$\begin{aligned} \text{Var}(\hat{\boldsymbol{\beta}})/\sigma^2 &= \mathbf{Q}_1'(\mathbf{R} + \mathbf{ZGZ}')\mathbf{Q}_1 \\ &= \mathbf{Q}_1'\mathbf{XC}_{11} + \mathbf{Q}_1'\mathbf{ZC}_{12}' + \mathbf{Q}_1'\mathbf{ZGZ}'\mathbf{Q}_1 \\ &= \mathbf{IC}_{11} - \mathbf{C}_{12}\mathbf{G}^{-1}\mathbf{C}_{12}' + \mathbf{C}_{12}'\mathbf{G}^{-1}\mathbf{GG}^{-1}\mathbf{C}_{12} \\ &= \mathbf{C}_{11}\,. \end{aligned}$$

$$\begin{aligned} \text{Cov}(\hat{\boldsymbol{\beta}}, \hat{\mathbf{u}}')/\sigma^2 &= \mathbf{Q}_1'(\mathbf{R} + \mathbf{ZGZ}')\mathbf{Q}_2 \\ &= \mathbf{Q}_1'\mathbf{XC}_{12} + \mathbf{Q}_1'\mathbf{ZC}_{22} + \mathbf{Q}_1'\mathbf{ZGZ}'\mathbf{Q}_2 \\ &= \mathbf{IC}_{12} - \mathbf{C}_{12}\mathbf{G}^{-1}\mathbf{C}_{22} - \mathbf{C}_{12}\mathbf{G}^{-1}\mathbf{G}(\mathbf{I} - \mathbf{G}^{-1}\mathbf{C}_{22}) \\ &= \mathbf{0}. \end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}\,(\hat{\boldsymbol\beta}, \mathbf{u}')/\sigma^2 &= \mathbf{Q}_1'\mathbf{Z}\mathbf{G} \\
&= -\mathbf{C}_{12}\mathbf{G}^{-1}\mathbf{G} \\
&= -\mathbf{C}_{12}\,.
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}\,(\hat{\boldsymbol\beta}, \hat{\mathbf{u}}' - \mathbf{u}')/\sigma^2 &= \mathbf{0} - (-\mathbf{C}_{12}) \\
&= \mathbf{C}_{12}\,.
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}\,(\hat{\mathbf{u}})/\sigma^2 &= \mathbf{Q}_2'(\mathrm{R} + \mathbf{ZGZ}')\mathbf{Q}_2 \\
&= \mathbf{Q}_2'\mathbf{X}\mathbf{C}_{12} + \mathbf{Q}_2'\mathbf{Z}\mathbf{C}_{22} + \mathbf{Q}_2'\mathbf{ZGZ}'\mathbf{Q}_2 \\
&= \mathbf{0} + (\mathbf{I} - \mathbf{C}_{22}\mathbf{G}^{-1})\mathbf{C}_{22} + (\mathbf{I} - \mathbf{C}_{22}\mathbf{G}^{-1})\mathbf{G}(\mathbf{I} - \mathbf{G}^{-1}\mathbf{C}_{22}) \\
&= \mathbf{G} - \mathbf{C}_{22}\,.
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}\,(\hat{\mathbf{u}}, \mathbf{u})/\sigma^2 &= \mathbf{Q}_2'\mathbf{Z}\mathbf{G} \\
&= (\mathbf{I} - \mathbf{C}_{22}\mathbf{G}^{-1})\mathbf{G} \\
&= \mathbf{G} - \mathbf{C}_{22}\,.
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}\,(\hat{\mathbf{u}} - \mathbf{u})/\sigma^2 &= \mathbf{G} - \mathbf{C}_{22} - 2(\mathbf{G} - \mathbf{C}_{22}) + \mathbf{G} \\
&= \mathbf{C}_{22}\,.
\end{aligned}$$

APPENDIX B

Derivation of Variance-Covariance Matrix of Estimators, Predictors, and Prediction Error in Selection Model.

Let the regular inverse of the coefficient matrix of (25) be

$$\begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{13} \\ \mathbf{T}_{13}' & \mathbf{T}_{33} \end{bmatrix}. \tag{B1}$$

Then,

$$\begin{bmatrix} \hat{\boldsymbol\beta} \\ \hat{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{13} \\ \mathbf{T}_{13}' & \mathbf{T}_{33} \end{bmatrix} \begin{bmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ \mathbf{B}'\mathbf{V}^{-1}\mathbf{y} \end{bmatrix} - \begin{bmatrix} \mathbf{Q}_1' \\ \mathbf{Q}_3' \end{bmatrix} y, \quad \text{say.}$$

The following identity is useful,

$$\begin{bmatrix} \mathbf{Q}_1' \\ \mathbf{Q}_3' \end{bmatrix} [\mathbf{X} \,\vdots\, \mathbf{B}] = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{13} \\ \mathbf{T}_{13}' & \mathbf{T}_{33} \end{bmatrix} \begin{bmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{V}^{-1}\mathbf{B} \\ \mathbf{B}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{B}'\mathbf{V}^{-1}\mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}.$$

Thus, $\mathbf{Q}_1'\mathbf{X} = \mathbf{I}$, $\mathbf{Q}_1'\mathbf{B} = \mathbf{0}$, $\mathbf{Q}_3'\mathbf{X} = \mathbf{0}$, and $\mathbf{Q}_3'\mathbf{B} = \mathbf{I}$.

$$\begin{aligned}
\mathrm{Var}\,(\hat{\boldsymbol\beta})/\sigma^2 &= \mathbf{Q}_1'(\mathbf{V} - \mathbf{B}\mathbf{H}_0\mathbf{B}')\mathbf{Q}_1 \\
&= \mathbf{Q}_1'\mathbf{X}\mathbf{T}_{11} + \mathbf{Q}_1'\mathbf{B}\mathbf{T}_{13}' - \mathbf{Q}_1'\mathbf{B}\mathbf{H}_0\mathbf{B}'\mathbf{Q}_1 \\
&= \mathbf{I}\mathbf{T}_{11} + \mathbf{0} - \mathbf{0} \\
&= \mathbf{T}_{11}\,.
\end{aligned}$$

$$
\begin{aligned}
\mathrm{Cov}\,(\hat{\boldsymbol\beta}, \hat{\mathbf{t}})/\sigma^2 &= \mathbf{Q}_1'(\mathbf{V} - \mathbf{B}\mathbf{H}_0\mathbf{B}')\mathbf{Q}_3 \\
&= \mathbf{Q}_1'\mathbf{X}\mathbf{T}_{13} + \mathbf{Q}_1'\mathbf{B}\mathbf{T}_{33} - \mathbf{Q}_1'\mathbf{B}\mathbf{H}_0\mathbf{B}'\mathbf{Q}_3 \\
&= \mathbf{I}\mathbf{T}_{13} + 0 - 0 \\
&= \mathbf{T}_{13}\,. \\
\mathrm{Cov}\,(\hat{\boldsymbol\beta}, \mathbf{w}')/\sigma^2 &= \mathbf{Q}_1'\mathbf{B}\mathbf{H}^{-1}\mathbf{H}_s \\
&= \mathbf{0}. \\
\mathrm{Var}\,(\hat{\mathbf{t}})/\sigma^2 &= \mathbf{Q}_3'(\mathbf{V} - \mathbf{B}\mathbf{H}_0\mathbf{B}')\mathbf{Q}_3 \\
&= \mathbf{Q}_3'\mathbf{X}\mathbf{T}_{13} + \mathbf{Q}_3'\mathbf{B}\mathbf{T}_{33} - \mathbf{Q}_3'\mathbf{B}\mathbf{H}_0\mathbf{B}'\mathbf{Q}_3 \\
&= \mathbf{0} + \mathbf{T}_{33} - \mathbf{H}_0 \\
&= \mathbf{T}_{33} - \mathbf{H}_0\,. \\
\mathrm{Cov}\,(\hat{\mathbf{t}}, \mathbf{w}')/\sigma^2 &= \mathbf{Q}_3'\mathbf{B}\mathbf{H}^{-1}\mathbf{H}_s \\
&= \mathbf{H}^{-1}\mathbf{H}_s\,.
\end{aligned}
$$

Let the regular inverse of the coefficient matrix of (28) be

$$
\begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} & \mathbf{T}_{13} \\ \mathbf{T}_{12}' & \mathbf{T}_{22} & \mathbf{T}_{23} \\ \mathbf{T}_{13}' & \mathbf{T}_{23}' & \mathbf{T}_{33} \end{bmatrix}. \tag{B2}
$$

From partitioned matrix inversion methods and the identity of (6), it is clear that $\mathbf{T}_{11}$, $\mathbf{T}_{13}$, and $\mathbf{T}_{33}$ of (B2) are identical to submatrices of (B1) with the same identification. From (B2) and (28),

$$
\begin{bmatrix} \hat{\boldsymbol\beta} \\ \hat{\mathbf{v}} \\ \hat{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} & \mathbf{T}_{13} \\ \mathbf{T}_{12}' & \mathbf{T}_{22} & \mathbf{T}_{23} \\ \mathbf{T}_{13}' & \mathbf{T}_{23}' & \mathbf{T}_{33} \end{bmatrix} \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1} \\ \mathbf{Z}'\mathbf{R}^{-1} \\ \mathbf{B}'\mathbf{R}^{-1} \end{bmatrix} \mathbf{y} = \begin{bmatrix} \mathbf{Q}_1' \\ \mathbf{Q}_2' \\ \mathbf{Q}_3' \end{bmatrix} \mathbf{y}, \quad \text{say.}
$$

$$
\begin{bmatrix} \mathbf{Q}_1' \\ \mathbf{Q}_2' \\ \mathbf{Q}_3' \end{bmatrix} [\mathbf{X} \,\vdots\, \mathbf{Z} \,\vdots\, \mathbf{B}] = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} & \mathbf{T}_{13} \\ \mathbf{T}_{12}' & \mathbf{T}_{22} & \mathbf{T}_{23} \\ \mathbf{T}_{13}' & \mathbf{T}_{23}' & \mathbf{T}_{33} \end{bmatrix} \left(\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{B} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{B} \\ \mathbf{B}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{B}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{B}'\mathbf{R}^{-1}\mathbf{B} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \right)
$$

$$
= \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{T}_{12}\mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_{22}\mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_{23}'\mathbf{G}^{-1} & \mathbf{0} \end{bmatrix}
$$

That is, $\mathbf{Q}_1'\mathbf{X} = \mathbf{I}$, $\mathbf{Q}_1'\mathbf{Z} = -\mathbf{T}_{12}\mathbf{G}^{-1}$, $\mathbf{Q}_1'\mathbf{B} = \mathbf{0}$, $\mathbf{Q}_2'\mathbf{X} = \mathbf{0}$, $\mathbf{Q}_2'\mathbf{Z} = \mathbf{I} - \mathbf{T}_{22}\mathbf{G}^{-1}$, $\mathbf{Q}_2'\mathbf{B} = \mathbf{0}$, $\mathbf{Q}_3'\mathbf{X} = \mathbf{0}$, $\mathbf{Q}_3'\mathbf{Z} = -\mathbf{T}_{23}'\mathbf{G}^{-1}$, and $\mathbf{Q}_3'\mathbf{B} = \mathbf{I}$. These identities are used extensively.

$$
\begin{aligned}
\mathrm{Cov}\,(\hat{\boldsymbol\beta}, \hat{\mathbf{v}}')/\sigma^2 &= \mathbf{Q}_1'(\mathbf{R} + \mathbf{Z}\mathbf{G}\mathbf{Z}' - \mathbf{B}\mathbf{H}_0\mathbf{B}')\mathbf{Q}_2 \\
&= \mathbf{Q}_1'\mathbf{X}\mathbf{T}_{12} + \mathbf{Q}_1'\mathbf{Z}\mathbf{T}_{22} + \mathbf{Q}_1'\mathbf{B}\mathbf{T}_{23}' + \mathbf{Q}_1'\mathbf{Z}\mathbf{G}\mathbf{Z}'\mathbf{Q}_2 + \mathbf{Q}_1'\mathbf{B}\mathbf{H}_0\mathbf{B}'\mathbf{Q}_2 \\
&= \mathbf{I}\mathbf{T}_{12} - \mathbf{T}_{12}\mathbf{G}^{-1}\mathbf{T}_{22} + \mathbf{0}(\mathbf{T}_{23}') - \mathbf{T}_{12}\mathbf{G}^{-1}\mathbf{G}(\mathbf{I} - \mathbf{G}^{-1}\mathbf{T}_{22}) - \mathbf{0}(\mathbf{H}_0\mathbf{B}')\mathbf{Q}_2 \\
&= \mathbf{0}.
\end{aligned}
$$

$$\begin{aligned}\text{Cov } \hat{\boldsymbol{\beta}}, \mathbf{u}')/\sigma^2 &= \mathbf{Q}_1(\mathbf{ZG} - \mathbf{BH}_0\mathbf{B}_u')\\ &= -\mathbf{T}_{12}\mathbf{G}^{-1}\mathbf{G} - \mathbf{0}(\mathbf{H}_0\mathbf{B}_u')\\ &= -\mathbf{T}_{12}\ .\end{aligned}$$

$$\begin{aligned}\text{Var }(\hat{\mathbf{v}})/\sigma^2 &= \mathbf{Q}_2'(\mathbf{R} + \mathbf{ZGZ}' - \mathbf{BH}_0\mathbf{B}')\mathbf{Q}_2\\ &= \mathbf{Q}_2'\mathbf{XT}_{12} + \mathbf{Q}_2'\mathbf{ZT}_{22} + \mathbf{Q}_2'\mathbf{BT}_{23}' + \mathbf{Q}_2'\mathbf{ZGZ}'\mathbf{Q}_2 - \mathbf{Q}_2'\mathbf{BH}_0\mathbf{B}'\mathbf{Q}_2\\ &= (\mathbf{0})(\mathbf{T}_{12}) + (\mathbf{I} - \mathbf{T}_{22}\mathbf{G}^{-1})\mathbf{T}_{22} + (\mathbf{0})(\mathbf{T}_{23}')\\ &\quad + (\mathbf{I} - \mathbf{T}_{22}\mathbf{G}^{-1})\mathbf{G}(\mathbf{I} - \mathbf{G}^{-1}\mathbf{T}_{22}) + (\mathbf{0})(\mathbf{H}_0)(\mathbf{0})\\ &= \mathbf{G} - \mathbf{T}_{22}\ .\end{aligned}$$

$$\begin{aligned}\text{Cov }(\hat{\mathbf{v}}, \hat{\mathbf{t}}')/\sigma^2 &= \mathbf{Q}_2'(\mathbf{R} + \mathbf{ZGZ}' - \mathbf{BH}_0\mathbf{B}')\mathbf{Q}_3\\ &= \mathbf{Q}_2'\mathbf{XT}_{13} + \mathbf{Q}_2'\mathbf{ZT}_{23} + \mathbf{Q}_2'\mathbf{BT}_{33} + \mathbf{Q}_2'\mathbf{ZGZ}'\mathbf{Q}_3 - \mathbf{Q}_2'\mathbf{BH}_0\mathbf{B}'\mathbf{Q}_3\\ &= \mathbf{0}(\mathbf{T}_{13}) + (\mathbf{I} - \mathbf{T}_{22}\mathbf{G}^{-1})\mathbf{T}_{23} + \mathbf{0}(\mathbf{T}_{33})\\ &\quad + (\mathbf{I} - \mathbf{T}_{22}\mathbf{G}^{-1})\mathbf{G}(-\mathbf{G}^{-1}\mathbf{T}_{23}) - (\mathbf{0})(\mathbf{H}_0\mathbf{B}')\mathbf{Q}_3\\ &= \mathbf{0}.\end{aligned}$$

$$\begin{aligned}\text{Cov }(\hat{\mathbf{v}}, \mathbf{u}')/\sigma^2 &= \mathbf{Q}_2'(\mathbf{ZG} - \mathbf{BH}_0\mathbf{B}_u')\\ &= (\mathbf{I} - \mathbf{T}_{22}\mathbf{G}^{-1})\mathbf{G} - (\mathbf{0})(\mathbf{H}_0\mathbf{B}_u')\\ &= \mathbf{G} - \mathbf{T}_{22}\end{aligned}$$

$$\begin{aligned}\text{Cov }(\hat{\mathbf{v}}, \mathbf{w}')/\sigma^2 &= \mathbf{Q}_2'\mathbf{BH}^{-1}\mathbf{H}_s\\ &= \mathbf{0}.\end{aligned}$$

$$\begin{aligned}\text{Cov }(\hat{\mathbf{t}}, \mathbf{u}')/\sigma^2 &= \mathbf{Q}_3'(\mathbf{ZG} - \mathbf{BH}_0\mathbf{B}_u')\\ &= -\mathbf{T}_{23}' - \mathbf{H}_0\mathbf{B}_u'.\end{aligned}$$

Now utilizing $\hat{\mathbf{u}} = \hat{\mathbf{v}} + \mathbf{B}_u\hat{\mathbf{t}}$,

$$\begin{aligned}\text{Cov }(\hat{\boldsymbol{\beta}}, \hat{\mathbf{u}}') &= \text{Cov }(\hat{\boldsymbol{\beta}}, \hat{\mathbf{v}}') + \text{Cov }(\hat{\boldsymbol{\beta}}, \hat{\mathbf{t}}'\mathbf{B}_u')\\ &= \mathbf{T}_{13}\mathbf{B}_u'\sigma^2.\end{aligned}$$

$$\begin{aligned}\text{Var }(\hat{\mathbf{u}}) &= \text{Var }(\hat{\mathbf{v}} + \mathbf{B}_u\hat{\mathbf{t}})\\ &= (\mathbf{G} - \mathbf{T}_{22} + \mathbf{B}_u\mathbf{T}_{33}\mathbf{B}_u' - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u')\sigma^2.\end{aligned}$$

$$\begin{aligned}\text{Cov }(\hat{\mathbf{u}}, \hat{\mathbf{t}}') &= \text{Cov }(\hat{\mathbf{v}}, \hat{\mathbf{t}}') + \mathbf{B}_u \text{ Var }(\hat{\mathbf{t}})\\ &= (\mathbf{B}_u\mathbf{T}_{33} - \mathbf{B}_u\mathbf{H}_0)\sigma^2.\end{aligned}$$

$$\begin{aligned}\text{Cov }(\hat{\mathbf{u}}, \mathbf{u}') &= \text{Cov }(\hat{\mathbf{v}}, \mathbf{u}') + \text{Cov }(\mathbf{B}_u\hat{\mathbf{t}}, \mathbf{u}')\\ &= (\mathbf{G} - \mathbf{T}_{22} - \mathbf{B}_u\mathbf{T}_{23}' - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u')\sigma^2.\end{aligned}$$

$$\begin{aligned}\text{Cov }(\hat{\mathbf{u}}', \mathbf{w}') &= \text{Cov }(\hat{\mathbf{v}}, \mathbf{w}') + \text{Cov }(\mathbf{B}_u\hat{\mathbf{t}}, \mathbf{w}')\\ &= \mathbf{B}_u\mathbf{H}^{-1}\mathbf{H}_s\sigma^2.\end{aligned}$$

Now we look at the elements in the inverse of the coefficient matrix in (31),

$$\begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{13}' & \mathbf{C}_{23}' & \mathbf{C}_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & -\mathbf{B}_u \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} & \mathbf{T}_{13} \\ \mathbf{T}_{12}' & \mathbf{T}_{22} & \mathbf{T}_{23} \\ \mathbf{T}_{13}' & \mathbf{T}_{23}' & \mathbf{T}_{33} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{B}_u' & \mathbf{I} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \mathbf{T}_{11} & \vdots & \mathbf{T}_{12} + \mathbf{T}_{13}\mathbf{B}_u' & \vdots & \mathbf{T}_{13} \\ \mathbf{T}_{12}' + \mathbf{B}_u\mathbf{T}_{13}' & \vdots & \mathbf{T}_{22} + \mathbf{B}_u\mathbf{T}_{23}' + \mathbf{T}_{23}\mathbf{B}_u' + \mathbf{B}_u\mathbf{T}_{33}\mathbf{B}_u' & \vdots & \mathbf{T}_{23} + \mathbf{B}_u\mathbf{T}_{33} \\ \mathbf{T}_{13}' & \vdots & \mathbf{T}_{23}' + \mathbf{T}_{33}\mathbf{B}_u' & \vdots & \mathbf{T}_{33} \end{bmatrix}.$$

Using these relationships between the elements of $\mathbf{C}$ and of $\mathbf{T}$ applied to the preceding variances and covariances of this appendix, we obtain

$$\frac{1}{\sigma^2}\operatorname{Cov}\left[\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{t}} \end{bmatrix}, \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{t}} \\ \mathbf{u} \\ \mathbf{w} \end{bmatrix}'\right] = \begin{bmatrix} \mathbf{C}_{11} & \vdots & \mathbf{C}_{13}\mathbf{B}_u' & \vdots \\ \mathbf{B}_u\mathbf{C}_{13}' & \vdots & \mathbf{G} - \mathbf{C}_{22} + \mathbf{C}_{23}\mathbf{B}_u' + \mathbf{B}_u\mathbf{C}_{23}' - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u' & \vdots \\ \mathbf{C}_{13}' & \vdots & \mathbf{C}_{33}\mathbf{B}_u' - \mathbf{H}_0\mathbf{B}_u' & \vdots \end{bmatrix}$$

$$\begin{matrix} \mathbf{C}_{13} & \vdots & -\mathbf{C}_{12} + \mathbf{C}_{13}\mathbf{B}_u' & \vdots & \mathbf{0} \\ \mathbf{B}_u\mathbf{C}_{33} - \mathbf{B}_u\mathbf{H}_0 & \vdots & \mathbf{G} - \mathbf{C}_{22} + \mathbf{C}_{23}\mathbf{B}_u' - \mathbf{B}_u\mathbf{H}_0\mathbf{B}_u' & \vdots & \mathbf{B}_u\mathbf{H}^{-1}\mathbf{H}_s \\ \mathbf{C}_{33} - \mathbf{H}_0 & \vdots & -\mathbf{C}_{23}' + \mathbf{C}_{33}\mathbf{B}_u' - \mathbf{H}_0\mathbf{B}_u' & \vdots & \mathbf{H}^{-1}\mathbf{H}_s \end{matrix} \Bigg].$$

Received November 1974

Key Words: Unbiased predictors; Selection bias; Prediction error variance; Sequential selection; Selection index.