Best linear unbiased prediction when error vector is correlated with other random vectors in the model

L.R. SCHAEFFER and C.R. HENDERSON*

*Department of Animal and Poultry Science,
University of Guelph,
Guelph, Ontario N1G 2W1 Canada
*Department of Animal Science,
Cornell University,
Ithaca, New York 14850 USA*

Summary

Non-zero covariances between random factors of a linear model with the residual or error vector can be handled with best linear unbiased prediction techniques. An equivalent model for describing y in which the covariances between random vectors with residual vectors are zero is the key to the solution. Computational difficulties depend on the structure of the covariance matrix. An example is used to illustrate the calculations.

Key-words : linear prediction, correlated vectors.

Résumé

Meilleure prédiction linéaire sans biais lorsque le vecteur d'erreurs est corrélé aux autres effets aléatoires du modèle

On peut traiter le cas de covariances non nulles entre, d'une part, les facteurs aléatoires d'un modèle linéaire et, d'autre part, le vecteur des résidus en utilisant les techniques du BLUP. La clé du problème réside dans l'écriture d'un modèle équivalent décrivant les données y de sorte que les covariances entre les vecteurs des effets aléatoires et des effets résiduels soient nulles. Les difficultés de calcul sont liées à la structure de la matrice de covariances entre ces deux types d'effets. Un exemple est donné qui illustre ces considérations.

Mots-clés : Prédiction linéaire, vecteurs corrélés.

I. Introduction

In mixed linear models, the covariances among the residual or error vector with other random factors in the model are assumed to be zero. This assumption is ordinarily applied to most practical applications in the biological sciences when the assumption is invalid. HENDERSON (1975) presented best linear unbiased prediction (BLUP) of random elements for a general linear model and also under a selection model. The objective of this paper is to extend HENDERSON's results to the case where the assumed covariances between residual effects and other random effects are not zero. The results suggest an equivalent model that could be used.

II. Model

Let the general linear model be

$$y = X\beta + Zu + e \tag{2.1}$$

where

$$E \begin{Bmatrix} y \\ u \\ e \end{Bmatrix} = \begin{Bmatrix} X\beta \\ 0 \\ 0 \end{Bmatrix},$$

$$V \begin{Bmatrix} y \\ u \\ e \end{Bmatrix} = \begin{pmatrix} V & CG & F \\ GC' & G & S' \\ F' & S & R \end{pmatrix}, \tag{2.2}$$

and

$$\begin{aligned} C &= Z + SG^{-1} \\ F &= ZS' + R \\ V &= ZGZ' + ZS' + SZ' + R \\ &= CGC' + R - SG^{-1}S'. \end{aligned} \tag{2.3}$$

Let

$$B = R - SG^{-1}S' \tag{2.4}$$

then

$$B^{-1} = R^{-1} - R^{-1}S(S'R^{-1}S - G)^{-1}S'R^{-1}$$

and

$$V^{-1} = B^{-1} - B^{-1}C(C'B^{-1}C + G^{-1})^{-1}C'B^{-1}.$$

Normally, S is assumed to be null, and R is taken to be $I\sigma^2$. Another paper could be written on the problem of estimation of S, R and G by either restricted maximum likelihood or minimum variance quadratic unbiased estimation, but in this paper S, R and G are known. We know that $\hat{\beta} = (X'V^{-1}X)^- X'V^{-1}y$ is BLUE and $\hat{u} = GC'V^{-1}(y - X\hat{\beta})$ is BLUP, but these are not easy computing algorithms.

III. An equivalent model

Models are defined to be equivalent if they yield the same variance-covariance matrix of y, and $E(y)$ is the same for both models.

Let

$$y = X\beta + Cu + \varepsilon \tag{3.1}$$

where

$$E \begin{Bmatrix} y \\ u \\ \varepsilon \end{Bmatrix} = \begin{Bmatrix} X\beta \\ 0 \\ 0 \end{Bmatrix}$$

and

$$V \begin{Bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{\epsilon} \end{Bmatrix} = \begin{Bmatrix} \mathbf{V} & \mathbf{CG} & \mathbf{B} \\ \mathbf{GC'} & \mathbf{G} & \mathbf{0} \\ \mathbf{B} & \mathbf{0} & \mathbf{B} \end{Bmatrix}$$

and $\mathbf{V} = \mathbf{CGC'} + \mathbf{B}$ which is the same as (2.3). Thus, a model with covariances between $\mathbf{u}$ and $\mathbf{e}$ can be transformed to an equivalent model with zero covariances between $\mathbf{\epsilon}$ and $\mathbf{u}$. The computational problems depend on the structure of $\mathbf{S}$ which influences the form of $\mathbf{C}$ and $\mathbf{B}$.

The equivalent model allows one to directly use the usual mixed model equations of Henderson (1975), which are in this case

$$\begin{bmatrix} \mathbf{X'B^{-1}X} & \mathbf{X'B^{-1}C} \\ \mathbf{C'B^{-1}X} & \mathbf{C'B^{-1}C + G^{-1}} \end{bmatrix} \begin{bmatrix} \mathbf{\beta^\circ} \\ \mathbf{\hat{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'B^{-1}y} \\ \mathbf{C'B^{-1}y} \end{bmatrix} \qquad (3.2)$$

Let

$$\begin{bmatrix} \mathbf{P_{11}} & \mathbf{P_{12}} \\ \mathbf{P'_{12}} & \mathbf{P_{22}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'B^{-1}X} & \mathbf{X'B^{-1}C} \\ \mathbf{C'B^{-1}X} & \mathbf{C'B^{-1}C + G^{-1}} \end{bmatrix}^{-},$$

then

$$V(\mathbf{\hat{u}}) = \mathbf{G} - \mathbf{P_{22}}$$
$$\text{Cov}(\mathbf{u}, \mathbf{\hat{u}}) = V(\mathbf{\hat{u}})$$
$$\text{Cov}(\mathbf{\hat{u}}, \mathbf{\beta^\circ}) = 0$$
$$V(\mathbf{\hat{u}} - \mathbf{u}) = \mathbf{P_{22}}$$

and

$V(\mathbf{K'\beta^\circ}) = \mathbf{K'P_{11}K}$ for $\mathbf{K'\beta}$ being estimable.

Alternative equations to (3.2) that do not require the inverse of $\mathbf{B}$ are

$$\begin{Bmatrix} \mathbf{X'R^{-1}X} & \mathbf{X'R^{-1}C} & \mathbf{X'R^{-1}S} \\ \mathbf{C'R^{-1}X} & \mathbf{C'R^{-1}C + G^{-1}} & \mathbf{C'R^{-1}S} \\ \mathbf{S'R^{-1}X} & \mathbf{S'R^{-1}C} & \mathbf{S'R^{-1}S - G} \end{Bmatrix} \begin{Bmatrix} \mathbf{\beta^\circ} \\ \mathbf{\hat{u}} \\ \mathbf{\theta} \end{Bmatrix} = \begin{Bmatrix} \mathbf{X'R^{-1}y} \\ \mathbf{C'R^{-1}y} \\ \mathbf{S'R^{-1}y} \end{Bmatrix}. \qquad (3.3)$$

The disadvantage of these equations is that $(\mathbf{S'R^{-1}S - G})$ is negative definite, and consequently Gauss-Seidel iteration would not be guaranteed to converge. The advantage of (3.3) is that the inverse of $\mathbf{B}$ is not needed and $\mathbf{R}$ may be diagonal. The order of equations (3.3) would be almost twice as large as that of equations (3.2) since the number of elements in $\mathbf{\hat{u}}$ would be the same as in $\mathbf{\theta}$.

IV. An easier derivation

Work by HENDERSON (1950, 1959, 1963) has shown that under normality assumptions maximizing the joint density of $\mathbf{y}$ and $\mathbf{u}$, $f(\mathbf{y}, \mathbf{u})$ gives BLUE of $\mathbf{K'\beta}$ and BLUP of $\mathbf{u}$ under any distribution. This derivation follows the MAP procedure of MELSA & COHN (1978).

Note that,

$$f(\mathbf{y}, \mathbf{u}) = g(\mathbf{y}|\mathbf{u}) h(\mathbf{u}),$$
$$E(\mathbf{y}|\mathbf{u}) = \mathbf{X\beta} + (\mathbf{Z} + \mathbf{SG^{-1}})\mathbf{u},$$

and

$$V(y \mid u) = B = R - SG^{-1}S'.$$

Then except for a constant,

$$\log[g(y \mid u)\, h(u)] = (y - X\beta - Cu)'B^{-1}(y - X\beta - Cu) + u'G^{-1}u.$$

Differentiating these equations with respect to β and u and equating to 0 gives (3.2). A similar result can be obtained using a Bayesian approach.

V. An example

Suppose that in dairy cattle there is a positive covariance between the genetic value of a bull and the residual effects associated with each daughter production record. Take ten observations on daughters of three sires in two herds, where

$$
y = \begin{pmatrix} 150 \\ 123 \\ 120 \\ 95 \\ 146 \\ 160 \\ 153 \\ 130 \\ 86 \\ 92 \end{pmatrix}
\quad
X\beta = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}
\begin{Bmatrix} h_1 \\ h_2 \end{Bmatrix}
\quad
Zu = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}
\begin{Bmatrix} s_1 \\ s_2 \\ s_3 \end{Bmatrix}
$$

Sires 1 and 2 are related so that

$$V(u) = G = \begin{Bmatrix} 1 & .5 & 0 \\ .5 & 1 & 0 \\ 0 & 0 & 1 \end{Bmatrix} \sigma_s^2,$$

$$V(e) = R = I\sigma_e^2,$$

and

$$k = \sigma_e^2 / \sigma_s^2 = 9. \text{ Let } \sigma_e^2 = 1 \text{ and } \sigma_s^2 = 1/9.$$

Assuming $S = 0$, the usual mixed model equations would be

$$
\begin{bmatrix} 4 & 0 & 1 & 2 & 1 \\ 0 & 6 & 3 & 1 & 2 \\ 1 & 3 & 16 & -6 & 0 \\ 2 & 1 & -6 & 15 & 0 \\ 1 & 2 & 0 & 0 & 12 \end{bmatrix}
\begin{Bmatrix} h_1^0 \\ h_2^0 \\ s_1^0 \\ s_2^0 \\ s_3^0 \end{Bmatrix}
= \begin{Bmatrix} 488 \\ 767 \\ 609 \\ 373 \\ 273 \end{Bmatrix}
$$

with solution

$$
\begin{Bmatrix} h_1^0 \\ h_2^0 \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \end{Bmatrix}
= \begin{Bmatrix} 120.0095 \\ 125.7033 \\ 8.4406 \\ 3.8614 \\ -8.2013 \end{Bmatrix}
$$

Now suppose that $S = .3Z\sigma_s^2$, then

$$C = Z + SG^{-1}$$

$$= \begin{pmatrix}
1.4 & -.2 & 0 \\
-.2 & 1.4 & 0 \\
-.2 & 1.4 & 0 \\
0 & 0 & 1.3 \\
1.4 & -.2 & 0 \\
1.4 & -.2 & 0 \\
1.4 & -.2 & 0 \\
-.2 & 1.4 & 0 \\
0 & 0 & 1.3 \\
0 & 0 & 1.3
\end{pmatrix},$$

and

$$R - SG^{-1}S'$$

$$= 1/9 \begin{pmatrix}
8.88 & .06 & .06 & 0 & -.12 & -.12 & -.12 & .06 & 0 & 0 \\
.06 & 8.88 & -.12 & 0 & .06 & .06 & .06 & -.12 & 0 & 0 \\
.06 & -.12 & 8.88 & 0 & .06 & .06 & .06 & -.12 & 0 & 0 \\
0 & 0 & 0 & 8.91 & 0 & 0 & 0 & 0 & -.09 & -.09 \\
-.12 & .06 & .06 & 0 & 8.88 & -.12 & -.12 & .06 & 0 & 0 \\
-.12 & .06 & .06 & 0 & -.12 & 8.88 & -.12 & .06 & 0 & 0 \\
-.12 & .06 & .06 & 0 & -.12 & -.12 & 8.88 & .06 & 0 & 0 \\
.06 & -.12 & -.12 & 0 & .06 & .06 & .06 & 8.88 & 0 & 0 \\
0 & 0 & 0 & -.09 & 0 & 0 & 0 & 0 & 8.91 & -.09 \\
0 & 0 & 0 & -.09 & 0 & 0 & 0 & 0 & -.09 & 8.91
\end{pmatrix}.$$

Then

$$\begin{bmatrix} X'B^{-1}X & X'B^{-1}C \\ C'B^{-1}X & C'B^{-1}C + G^{-1} \end{bmatrix} \quad \begin{bmatrix} \beta^o \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'B^{-1}y \\ C'B^{-1}y \end{bmatrix}$$

are

$$\begin{Bmatrix}
4.0516 & .0401 & .9850 & 2.6879 & 1.3402 \\
.0401 & 6.1394 & 4.2029 & 0.7384 & 2.6804 \\
.9850 & 4.2029 & 20.4610 & -8.2355 & 0 \\
2.6879 & 0.7384 & -8.2355 & 18.3470 & 0 \\
1.3402 & 2.6804 & 0 & 0 & 14.2268
\end{Bmatrix} \begin{Bmatrix}
h_1^o \\ h_2^o \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3
\end{Bmatrix} = \begin{Bmatrix}
498.3216 \\ 791.2177 \\ 810.7590 \\ 398.9563 \\ 365.8763
\end{Bmatrix}$$

with solutions

$$\begin{Bmatrix}
h_1^o \\ h_2^o \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3
\end{Bmatrix} = \begin{Bmatrix}
120.3063 \\ 125.2487 \\ 9.4405 \\ 3.3165 \\ -9.2134
\end{Bmatrix}$$

In practice, B and B^{-1} may be difficult to construct depending on the definition of S. In such cases, the alternative equations (3.2) may be used, especially if S is simply a multiple of Z. For the example data, let $\sigma_e^2 = 1$ and $\sigma_s^2 = 1/9$, then equations (4.1) would be

$$
\begin{pmatrix}
4 & 0 & 1 & 2.6 & 1.3 & .0333 & .0667 & .0333 \\
0 & 6 & 4 & 0.8 & 2.6 & .1 & .0333 & .0667 \\
1 & 4 & 19.96 & -7.96 & .0 & .1867 & -.0200 & .0 \\
2.6 & 0.8 & -7.96 & 18.04 & .0 & -.0267 & .1400 & .0 \\
1.3 & 2.6 & 0 & 0 & 14.07 & .0 & .0 & .1300 \\
.0333 & 0.1 & .1867 & -.0267 & .0 & -.1067 & -.0556 & .0 \\
.0667 & .0333 & -.0200 & .1400 & .0 & -.0556 & -.1078 & .0 \\
.0333 & .0667 & .0 & .0 & .1300 & .0 & .0 & -.1078
\end{pmatrix}
$$

$$
\begin{Bmatrix}
h_1^o \\
h_2^o \\
\hat{s}_1 \\
\hat{s}_2 \\
\hat{s}_3 \\
\theta_1 \\
\theta_2 \\
\theta_3
\end{Bmatrix}
=
\begin{Bmatrix}
488 \\
767 \\
778 \\
400.4 \\
354.9 \\
20.3 \\
12.4333 \\
9.1
\end{Bmatrix}
$$

with solutions

$$
\begin{Bmatrix}
h_1^o \\
h_2^o \\
\hat{s}_1 \\
\hat{s}_2 \\
\hat{s}_3 \\
\theta_1 \\
\theta_2 \\
\theta_3
\end{Bmatrix}
=
\begin{Bmatrix}
120.3063 \\
125.2487 \\
9.4405 \\
3.3165 \\
-9.2134 \\
-27.0471 \\
14.2901 \\
19.1355
\end{Bmatrix}
$$

References

HENDERSON C.R., 1950. Estimation of genetic parameters. *Biometrics*, **6**, 186.

HENDERSON C.R., 1963. Selection index and expected genetic advance *In: Statistical Genetics and Plant Breeding*, NAS NRC pp. 141-163.

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

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

MELSA J.L., COHN D.L., 1978. *Decision and Estimation Theory*. New York, McGraw-Hill, Inc.