An alternative derivation method of mixed model equations from best linear unbiased prediction (BLUP) and restricted BLUP of breeding values not using maximum likelihood

Masahiro Satoh

Graduate School of Agricultural Science, Tohoku University, Sendai, Japan

Correspondence
Masahiro Satoh, Graduate School of Agricultural Science, Tohoku University, Sendai, Japan.
Email: masahiro.satoh.d5@tohoku.ac.jp

Funding information
JSPS KAKENHI, Grant/Award Number: JP15k07751

Abstract
Two mixed model equations (MME) for best linear unbiased prediction (BLUP) of breeding values and for restricted BLUP of breeding values were derived by maximum likelihood from the joint normal probability distribution of the observations and breeding values. As a result, MME is actually more general than maximum likelihood because we can prove that each set of solutions of MME are identical to BLUP and restricted BLUP of breeding values and then it does not depend on normality. In the present study, the author shows deriving directly each MME from BLUP and restricted BLUP equations for breeding values without assuming the joint normal distribution of the data and random effects. However, if we cannot assume the multivariate normal density distribution of the estimated aggregate breeding value and each breeding value for selected traits, the response to selection by restricted BLUP may deviate from the expected values.

KEYWORDS
BLUP procedure, maximum likelihood, mixed model equations, restricted BLUP procedure

1 | INTRODUCTION

Henderson (1950) showed estimating breeding values based on maximizing the log-joint normal density of the data and random effects and later he considered the more general problem of predicting linear functions of the unknown fixed and random effects, that is mixed model equations (MME) (Henderson, 1973). Henderson, Kempthorne, Searle, and von Krosigk (1959) proved that the solutions of fixed effects in MME are identical to the solutions of generalized least-squares. Furthermore, Henderson (1963) showed that the solutions of random effects in MME are identical to the solutions of best linear unbiased prediction (BLUP) of breeding values.

On the other hand, Quaas and Henderson (1976) extended restricted selection index and proposed restricted BLUP procedure. The restricted BLUP of breeding values are derived by imposing constraints directly within a multiple trait mixed model. The MME for estimating restricted BLUP of breeding values also can be shown by the method that can be obtained by maximizing for variation in fixed effects and breeding values the joint density function of the data and breeding values, described by Henderson (1975).

However, as a result, MME is actually more general than maximum likelihood because it can be proven that fixed effect solutions of MME are identical to solutions of generalized least-squares equations and random effect solutions of two MME are identical to BLUP and restricted BLUP of breeding values, respectively, and then it does not depend on normality.

The objective of this study is to derive directly each MME from BLUP and restricted BLUP equations algebraically without assuming the joint normal density function of the data and breeding values.
2 | METHODS

2.1 | Derivation of BLUP of breeding values

An additive genetic mixed model can be represented as follows:

\[ y = Xb + Zu + e, u \sim (0, G), e \sim (0, R), \text{cov}(u, e') = 0 \]  

(1)

where \( y \) is a vector of observations and \( b \) is an unknown vector of fixed effects, \( u \) and \( e \) are unknown vectors of random and residual effects, respectively, \( X \) and \( Z \) are unknown incidence matrices relating elements of \( b \) to \( y \) and \( u \) to \( y \), respectively, and \( \text{var}(y) = V = ZGZ' + R \). In model (1), distributions of \( y, u, e \) are not assumed to be multivariate normal.

Henderson (1973) considered the general problem with respect to model (1) is to predict a linear function of the unknown \( b \) and \( u \), that is \( w = k'b + a'u \), using a linear function of \( y \). The predictor \( \hat{w} \) can be written in any vector \( p \) and a scalar \( q \) as follows:

\[
\hat{w} = p'y + q.
\]

The BLUP of \( w \), that is \( \hat{w} \) satisfies (i) a linear function of \( y \), (ii) an unbiased predictor of \( w \), that is \( E(\hat{w}) = E(w) \), (iii) least squares error, that is \( E(\hat{w} - w)^2 \) is minimized. The BLUP of \( w \) is

\[
\hat{w} = k'b + a'GZV^{-1}(y - Xb).
\]

where \( b \) is the generalized least-squares solutions for \( b \), that is

\[
XV^{-1}Xb = XV^{-1}y.
\]

(2)

Then BLUP of \( u \), that is \( \hat{u} \) is

\[
\hat{u} = GV^{-1}(y - Xb).
\]

(3)

We refer to (3) as BLUP equations. If \( k'b \) is estimable, \( k'b \) is the best linear unbiased estimator of \( k'b \).

2.2 | Derivation of restricted BLUP of breeding values

In a general mixed model (1), let the set of restrictions on \( u \) be \( C'u \). Kemphorne and Nordskog (1959) defined \( C \) for no-change constraints and Mallard (1972) and Itoh and Yamada (1987) extended it for predetermined proportional changes. The predictor is to be chosen such that \( E(p'y - a'u)^2 \) is minimum subject to the restriction that \( E(C'u p'y) = 0 \) (Quaas & Henderson, 1976). The restricted BLUP of \( u \), that is \( \hat{u}_R \) is

\[
\hat{u}_R = GZV^{-1}(y - X\hat{b}_R - ZGC\hat{t}),
\]

(4)

where \( \hat{b}_R \) and \( \hat{t} \) are any solutions to (5):

\[
\begin{bmatrix} XV^{-1}X & XV^{-1}ZG \\ CGZV^{-1}X & CGZV^{-1}ZGC \end{bmatrix} \begin{bmatrix} \hat{b}_R \\ \hat{t} \end{bmatrix} = \begin{bmatrix} XV^{-1}y \\ CGZV^{-1}y \end{bmatrix}.
\]

(5)

Note that these are generalized least-squares equations under a model \( E(y) = Xb + ZGCt \); var(\( y \)) = \( V \) (Quaas & Henderson, 1976).

2.3 | Derivation of MME from BLUP equations using maximum likelihood

Matrix \( V \) in (2) and (3) is often so large that its inversion is not computationally feasible. An alternative method without the need for computing \( V^{-1} \) was suggested by Henderson (1950). The prediction of \( w \) is \( k'b + a'\hat{u} \), where \( b \) and \( u \) are any solutions to (6):

\[
\begin{bmatrix} XR^{-1}X & XR^{-1}Z \\ ZR^{-1}X & ZR^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} XR^{-1}y \\ ZR^{-1}y \end{bmatrix}.
\]

(6)

When \( u \) and \( e \) are normally joint distributed, MME can be obtained by maximizing for variations in \( b \) and \( u \). The proofs that \( b \) and \( u \) in (6) are equal to \( \hat{b}_R \) in (2) and \( \hat{u}_R \) in (3), respectively were given by Henderson, Kemphorne, Seale, and von Krisig (1959) and Henderson (1963), respectively. Therefore, MME are satisfied not assuming joint normal distribution of \( y \) and \( u \).

2.4 | Derivation of MME from restricted BLUP equations using maximum likelihood

Quaas and Henderson (1976) described that the predictor is \( k'b_R + a'\hat{u}_R \), where \( b_R \) and \( u_R \) are solutions to (7), that is restricted MME:

\[
\begin{bmatrix} XR^{-1}X & XR^{-1}ZG \\ ZR^{-1}X & ZR^{-1}ZGC \end{bmatrix} \begin{bmatrix} b_R \\ u_R \end{bmatrix} = \begin{bmatrix} XR^{-1}y \\ ZR^{-1}y \end{bmatrix}.
\]

(7)

The proofs that \( \hat{b}_R \) and \( \hat{u}_R \) in (7) are equal to \( \hat{b}_R \) in (5) and \( \hat{u}_R \) in (4), respectively, can be shown by the methods described in Henderson (1975). Therefore, restricted MME are satisfied not assuming joint normal distribution of \( y \) and \( u \).

3 | RESULTS

For deriving two MME, that is (6) and (7) without assuming any distribution, at first we use Matrix inversion lemma (Sherman-Morrison-Woodbury identity) and MME identity, that is (8) and (9), respectively:

\[
V^{-1} = (R + ZGZ')^{-1} = R^{-1} - R^{-1}Z(G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1},
\]

(8)

\[
(G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1} = GZV^{-1}.
\]

(9)

The proofs of (8) and (9) are shown in the Appendix. Alternative derivation of (8), leading to a generalized expression, can be found in Henderson and Searle (1981) and Tylavsky and Sohie (1986).

3.1 | Derivation of MME from BLUP equations without joint normal distribution

Substituting (8) for (2)

\[
XR^{-1}Xb + XR^{-1}Z(G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}b = XR^{-1}y + XR^{-1}Z(G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}y.
\]

(10)
Substituting (9) for (3)

\[ \hat{u} = (G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}y - (G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}x_b. \]

Then

\[ (G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}x_b = (G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}y - \hat{u}. \] (11)

Therefore

\[ ZR^{-1}x_b + (G^{-1} + ZR^{-1}Z)\hat{u} = ZR^{-1}y. \] (12)

Substituting (11) for (10) gives

\[ XR^{-1}x_b + XR^{-1}\hat{u} = XR^{-1}y. \] (13)

Therefore, (12) and (13) are exactly the same as (6). This derivation of MME does not need to assume any distribution of \( y \) and \( u \).

### 3.2 Derivation of restricted MME from restricted BLUP equations without joint normal distribution

Substituting (9) for (4)

\[ \hat{u}_R = (G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}(y - Xb_R - ZGc). \]

Then

\[ (G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}Xb_R + \hat{u}_R + (G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}ZGc = (G^{-1} + ZR^{-1}Z)^{-1}ZR^{-1}y. \] (14)

That is,

\[ ZR^{-1}Xb_R + (G^{-1} + ZR^{-1}Z)\hat{u}_R + ZR^{-1}ZGc = ZR^{-1}y. \] (15)

Substituting (8) for the first and second equations in (5) and using (14)

\[ XR^{-1}x_b + XR^{-1}\hat{u} + XR^{-1}ZGc = XR^{-1}y \] (16)

and

\[ CGZR^{-1}x_b + CGZR^{-1}\hat{u} + CGZR^{-1}ZGc = CGZR^{-1}y. \] (17)

Therefore, (15), (16), and (17) are exactly the same as (7). This derivation of restricted MME does not need to assume any distribution of \( y \) and \( u \).

### 4 DISCUSSION

Henderson (1963) showed that the BLUP of \( kb + a'v \) is \( k\hat{b} + a'GV^{-1}(y - \hat{X}b) \), where \( \hat{b} \) is any solution to \( X'V^{-1}Xb = X'V^{-1}y \). The difficulty with this method is that \( V \) is often a matrix so large that its inversion is very costly. Before that, Henderson (1950) derived MME by maximizing for variation in \( b \) and \( u \) using the joint normal density function of \( y \) and \( u \). Furthermore, Henderson, Kempthorne, Searle, and von Krogh (1959) proved the solutions of MME (6) are identical to the solutions of least-squares (2) and BLUP equations (3). On the other hand, Quaas and Henderson (1976) proposed restricted BLUP procedure under \( E(Cu'p'y) = 0 \) and the MME for estimating restricted BLUP of breeding values also can be derived by the methods described in Henderson (1975).

Each MME from BLUP and restricted BLUP equations need to assume at first the joint normal density function of \( y \) and \( u \). Because finally we can prove that the solutions of each MME are identical to BLUP and restricted BLUP equations, respectively, both MME do not need to be the joint normal density function of \( y \) and \( u \). On the other hand, the viewpoint of this study is to derive directly MME from BLUP and restricted BLUP equations algebraically without assuming joint normal density distribution of \( y \) and \( u \).

The main purpose of the restricted BLUP procedure is to maximize the genetic changes in one or more traits equal to zero or some proportional constraint on genetic gains, or to achieve desired genetic gains for all traits. For achieving this purpose, we have to assume the multivariate normal density distribution of the estimated aggregate breeding value and each breeding value for selected traits. Therefore, even though the MME from restricted BLUP equations does not need to assume the joint normal density function of \( y \) and \( u \), the condition is not sufficient for selection with constraint.

### ACKNOWLEDGMENTS

This study was supported by JSPS KAKENHI Grant Number JP15k07751.

### REFERENCES

Henderson, C. R. (1950). Estimation of genetic parameters. Annals of Mathematical Statistics, 21, 309–310.

Henderson, C. R. (1963). Selection index and expected genetic advance. In: W. D. Hanson & H. F. Robinson (Eds), Statistical genetics and plant breeding (pp. 141–163), Publication 982, Washington DC: National Academy of Sciences, National Research Council.

Henderson, C. R. (1973). Sire evaluation and genetic trends. In: Proceedings of the animal breeding and genetics symposium in honor of J. L. Lush (pp. 10–41). Blackburgh, Champaign, IL: American Society for Animal Science.

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. Biometrics, 31, 423–447. https://doi.org/10.2307/2529430

Henderson, C. R., Kempthorne, O., Searle, S. R., & von Krogh, C. M. (1959). The estimation of environmental and genetic trends from records subject to culling. Biometrics, 15, 192–218. https://doi.org/10.2307/2527669

Henderson, H. V., & Searle, S. R. (1981). On deriving the inverse of a sum of matrices. Society for Industrial Applied Mathematics Review, 23, 53–60.

Itoh, Y., & Yamada, Y. (1987). Comparison of selection indices achieving predetermined proportional gains. Genetik Selektion Evolution, 19, 69–82.

Kempthorne, O., & Nordskog, A. W. (1959). Restricted selection indices. Biometrics, 15, 10–19. https://doi.org/10.2307/2527598

Mallard, J. (1972). La theorie et la calcul des index de selection avec restrictions: Synthese critique. Biometrics, 28, 713–735. https://doi.org/10.2307/2528758

Quaas, R. L., & Henderson, C. R. (1976). Restricted best linear unbiased prediction of breeding values (pp. 1–14). Ithaca, NY: Mimeo, Cornell Univ.
APPENDIX

Matrix inversion lemma

Let $A_{n \times n}$ and $C_{m \times m}$ be of full rank and $B$ and $D$ have orders $n \times m$ and $m \times n$, respectively. Then:

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B\left(C^{-1} + DA^{-1}B\right)^{-1}DA^{-1}.$$ 

Proof:

First consider these useful identities,

$$B + BCD^{-1}B = BC\left(C^{-1} + DA^{-1}B\right) = (A + BCD)A^{-1}B.$$ 

Therefore:

$$(A + BCD)^{-1}BC = A^{-1}B\left(C^{-1} + DA^{-1}B\right)^{-1}.$$ 

Now,

$$A^{-1} = (A + BCD)^{-1}(A + BCD)^{-1} = A^{-1} - A^{-1}B\left(C^{-1} + DA^{-1}B\right)^{-1}DA^{-1}.$$ 

Hence, using Matrix inversion lemma,

$$\begin{align*}
(A + BCD)^{-1} &= A^{-1} - A^{-1}B\left(C^{-1} + DA^{-1}B\right)^{-1}DA^{-1} \\
&= (A + BCD)^{-1}\begin{pmatrix}
I & B(C^{-1} + DA^{-1}B)\end{pmatrix} \\
&= (A + BCD)^{-1}\begin{pmatrix}
A & B(C^{-1} + DA^{-1}B)\end{pmatrix}\begin{pmatrix}
A^{-1} & -A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}
\end{pmatrix} \\
&= (A + BCD)^{-1}\begin{pmatrix}
A^{-1} & -A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}
\end{pmatrix} \\
&= (A + BCD)^{-1}\begin{pmatrix}
A^{-1} & -A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}
\end{pmatrix} \\
&= (A + BCD)^{-1}\begin{pmatrix}
A^{-1} & -A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}
\end{pmatrix}.
\end{align*}$$