

ANIMAL BREEDING NOTES**CHAPTER 10****BEST LINEAR UNBIASED PREDICTION****Derivation of the Best Linear Unbiased Predictor (BLUP)**

Let

$y = [y_1 \ y_2 \ \dots \ y_n]$ be an observable random vector, and

$g = [g_1 \ g_2 \ \dots \ g_p]$ be an unobservable random vector,

where y and g are jointly distributed.

Assume that:

- (1) The joint distribution of y and g as well as the means of y and g are unknown, and
- (2) All variances and covariances are known.

Let

$$\begin{bmatrix} y \\ g \end{bmatrix} \sim \left\{ \begin{bmatrix} X\beta \\ 0 \end{bmatrix}, \begin{bmatrix} V & C \\ C' & G \end{bmatrix} \right\},$$

where

X = known incidence matrix relating records to elements of β , and

β = vector of unknown constants (fixed effects).

The $E[g]$ was set to zero to retain the property of maximization of the probability of correct pairwise ranking, shown for the BLP and for the BP under normality.

We want to predict

$$w = K'\beta + L'g,$$

[10-2]

where K is a matrix of estimable contrasts and L is also a matrix of contrasts, **using**

$$\hat{w} = a + By,$$

where **vector a and matrix B are chosen so that they minimize the mean square error of prediction (MSEP)**, i.e., they minimize

$$E[(a + By - K'\beta - L'g)' A (a + By - K'\beta - L'g)],$$

where A is any s.p.d. matrix, **subject to the restriction** $E[\hat{w}] = E[w]$, i.e.,

$$E[a + By] = E[K'\beta + L'g]$$

$$a + BX\beta = K'\beta + 0$$

$$\Rightarrow a = 0, \text{ and}$$

$$BX = K'.$$

The resulting best linear unbiased predictor (BLUP) of w is:

$$\hat{w} = K'\beta^* + L'C'V^{-1}(y - X\beta^*),$$

where

$$\begin{aligned}\beta^* &= (X'V^{-1}X)^{-1}X'V^{-1}y \\ &= \text{GLS of } \beta,\end{aligned}$$

$K'\beta^*$ = Best Linear Unbiased Estimator (BLUE) of the set of estimable functions $K'\beta$ in the model $y = X\beta + \varepsilon$, $y \sim (X\beta, V)$, and

$$L'C'V^{-1}(y - X\beta^*) = L'\hat{g}, \text{ the BLUP of } L'g.$$

Thus, the BLUP of w is:

$$\hat{w} = K'\beta^* + L'\hat{g}.$$

Proof:

First, a brief explanation of **Lagrange multipliers**, a procedure used to impose restrictions when

[10-3]

maximizing or minimizing a linear function. This procedure will be used in the derivation of BLUP. Lagrange multipliers can be applied to a scalar, vector, or matrix. For example:

[1] Scalar Lagrange multiplier

Restriction: $ax = b \rightarrow \lambda(ax - b)$

[2] Vector of Lagrange multipliers

$$\text{Restrictions: } \begin{bmatrix} a_1 x_1 \\ a_2 x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \rightarrow \lambda_1(a_1 x_1 - b_1) + \lambda_2(a_2 x_2 - b_2)$$

[3] Matrix of Lagrange multipliers

$$\begin{aligned} \text{Restrictions: } & \begin{bmatrix} a_{11}x_1 & a_{12}x_2 \\ a_{21}x_1 & a_{22}x_2 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ \rightarrow & \text{tr} \left\{ \begin{bmatrix} \lambda_{11} & \lambda_{21} \\ \lambda_{12} & \lambda_{22} \end{bmatrix} \begin{bmatrix} a_{11}x_1 - b_{11} & a_{12}x_2 - b_{12} \\ a_{21}x_1 - b_{21} & a_{22}x_2 - b_{22} \end{bmatrix} \right\} \\ = & \text{tr} \{ T' (AX - B) \} \\ = & \lambda_{11}(a_{11}x_1 - b_{11}) + \lambda_{21}(a_{21}x_1 - b_{21}) + \lambda_{12}(a_{12}x_2 - b_{12}) + \lambda_{22}(a_{22}x_2 - b_{22}) \end{aligned}$$

A matrix of Lagrange multipliers will be used in the derivation of BLUP below.

The joint distribution of y and w is:

$$\begin{bmatrix} y \\ w \end{bmatrix} \sim \left\{ \begin{bmatrix} X\beta \\ K'\beta \end{bmatrix}, \begin{bmatrix} V & CL \\ L'C & L'GL \end{bmatrix} \right\}$$

Minimize

$$L = E[(a + By - K'\beta - L'g)'A(a + By - K'\beta - L'g)] + \text{tr}(2M'A(BX - K')),$$

where $2M'A = 2T'$ = matrix of Lagrange multipliers and $M' = T'A^{-1}$. The matrix A^{-1} exists because A is s.p.d.

[10-4]

$$\begin{aligned}
 L &= E[a'Aa + a'ABy - a'AK'\beta - a'AL'g \\
 &\quad + y'B'Aa + y'B'ABy - y'B'AK'\beta - y'B'AL'g \\
 &\quad - \beta'KAa - \beta'KABy + \beta'KAK'\beta + \beta'KAL'g \\
 &\quad - g'LAa - g'LABy + g'LAK'\beta + g'LAL'g] \\
 &\quad + \text{tr}(2M'A(BX - K')) \\
 L &= a'Aa + a'ABX\beta - a'AK'\beta - 0 \\
 &\quad + \beta'X'B'Aa + \text{tr}(B'ABV) + \beta'X'B'ABX\beta - \beta'X'B'AK'\beta - \text{tr}(B'AL'C') - 0 \\
 &\quad - \beta'KAa - \beta'KABX\beta + \beta'KAK'\beta + 0 \\
 &\quad - 0 - \text{tr}(LABC) + 0 + 0 + \text{tr}(LAL'G) + 0 \\
 &\quad + \text{tr}(2M'ABX) - \text{tr}(2M'AK')
 \end{aligned}$$

$$\frac{\partial L}{\partial a} = 2Aa + 2ABX\beta - 2AK'\beta = 0$$

$$a + BX\beta = K'\beta$$

$$\Rightarrow a = K'\beta - BX\beta$$

$$\frac{\partial L}{\partial M'} = 2ABX - 2AK' = 0$$

$$\Rightarrow BX = K'$$

$$\Rightarrow a = BX\beta - BX\beta$$

$$a = 0$$

$$\frac{\partial L}{\partial B} = 2Aa\beta'X' + 2ABV + 2ABX\beta\beta'X' - 2AK'\beta\beta'X' - 2AL'C' + 2AMX' = 0$$

However, $a = 0$ and multiplication by $\frac{1}{2} A^{-1}$ gives

[10-5]

$$BV + BX\beta\beta'X - K'\beta\beta'X - L'C' + MX' = 0$$

In addition, because $K' = BX$,

$$BV + BX\beta\beta'X - BX\beta\beta'X - L'C' + MX' = 0$$

Thus, because of the constraint $K' = BX$, the BLUP of $L'g$ will be invariant to β .

Now we solve for B in the following set of equations:

$$BV + MX' = L'C' \quad (1)$$

$$BX = K' \quad (2)$$

From (1),

$$B = L'C'V^{-1} - MX'V^{-1} \quad (3)$$

Substituting the expression of B in (3) for B in (2) yields,

$$L'C'V^{-1}X - MX'V^{-1}X = K'$$

$$\Rightarrow M = -K'(X'V^{-1}X)^{-} + L'C'V^{-1}X(X'V^{-1}X)^{-}$$

Using this expression for M in (3) gives:

$$B = L'C'V^{-1} + K'(X'V^{-1}X)^{-}X'V^{-1} - L'C'V^{-1}X(X'V^{-1}X)^{-}X'V^{-1}$$

$$\Rightarrow \hat{w} = a + By$$

$$\hat{w} = K'(X'V^{-1}X)^{-}X'V^{-1}y + L'C'V^{-1}(y - (X'V^{-1}X)^{-}X'V^{-1}y)$$

$$\hat{w} = K'\beta^{\circ} + L'\hat{g}, \text{ the BLUP of } w.$$

where

$$\beta^{\circ} = (X'V^{-1}X)^{-}X'V^{-1}y, \text{ the GLS of } \beta.$$

From the model $y = X\beta + \varepsilon$, $y \sim (X\beta, V)$,

$$K'\beta^{\circ} = \text{BLUE of } K'\beta, \text{ and}$$

$$\hat{g} = C'V^{-1}(y - X\beta^{\circ}), \text{ the BLUP of } g.$$

[10-6]

The expectations of $K'\beta^*$ and \hat{g} are:

$$\begin{aligned} E[K'\beta^*] &= K'\beta \\ E[\hat{g}] &= C'V^{-1}(X\beta - X\beta) \\ &= 0 \end{aligned}$$

Translation invariance of Best Linear Unbiased Predictor

Definition: a predictor $\hat{w} = a + B\gamma$ is translation invariant if

$$a + B\gamma = a + B(\gamma + Xt) \text{ for any vector } t.$$

Theorem: The BLUP of $L'g$, i.e., $L'\hat{g}$, is translation invariant to the value of β .

Proof:

Minimize:

$$\begin{aligned} L &= E[(a + B(\gamma + Xt) - K'\beta - L'g)'A(a + B(\gamma + Xt) - K'\beta - L'g)] \\ &\quad + \text{tr}(2M'A(BX - K')) \end{aligned}$$

But,

$$\begin{aligned} E[\gamma + Xt] &= X\beta + Xt \\ &= X(\beta + t) \\ &= X\beta^*, \text{ for } \beta^* = (\beta + t) \end{aligned}$$

Thus,

$$\begin{aligned} L &= a'Aa + a'ABX\beta^* - a'AK'\beta^* - 0 \\ &\quad + \beta^{*'}X'B'Aa + \text{tr}(B'ABV) + \beta^{*'}X'B'ABX\beta^* - \beta^{*'}X'B'AK'\beta^* - \text{tr}(B'AL'C' + 0) \\ &\quad + \beta^{*'}KAa - \beta^{*'}KABX\beta^* + \beta^{*'}KAK'\beta^* + 0 \\ &\quad - 0 - \text{tr}(LABC + 0) + 0 + \text{tr}(LAL'G + 0) \end{aligned}$$

[10-7]

$$+ \text{tr}(2M'ABX - 2M'AK')$$

$$\frac{\partial L}{\partial a} = 2Aa + 2ABX\beta^* - 2AK'\beta^* = 0$$

$$\Rightarrow a = K'\beta^* - BX\beta^*$$

$$\frac{\partial L}{\partial M} = 2X'B'A - 2KA = 0$$

$$\Rightarrow BX = K'$$

$$\Rightarrow a = BX\beta^* - BX\beta^*$$

$$a = 0$$

$$\frac{\partial L}{\partial B} = 2ABV + 2ABX\beta^*\beta^{*'X} - 2AK'\beta^*\beta^{*'X} - 2AL'C' + 2AMX' = 0$$

But $K' = BX$, thus

$$BV + BX\beta^*\beta^{*'X} - BX\beta^*\beta^{*'X} - L'C' + MX' = 0$$

\Rightarrow The BLUP of $L'g$ will be invariant to β because β^* was eliminated from this equation due to the restriction $BX = K'$.

We now solve for B in the set of equations:

$$BV + MX' = L'C' \quad (1)$$

$$BX = K' \quad (2)$$

From (1),

$$B = L'C'V^{-1} - MX'V^{-1} \quad (3)$$

From (3) and (2),

$$L'C'V^{-1}X - MX'V^{-1}X = K'$$

$$\Rightarrow M = -K'(X'V^{-1}X)^{-1} + L'C'V^{-1}X(X'V^{-1}X)^{-1} \quad (4)$$

[10-8]

From (4) and (3),

$$\begin{aligned} \mathbf{B} &= \mathbf{L}'\mathbf{C}'\mathbf{V}^{-1} + \mathbf{K}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}' - \mathbf{L}'\mathbf{C}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1} \\ \Rightarrow \hat{\mathbf{w}} &= \mathbf{a} + \mathbf{B}\mathbf{y} \\ \hat{\mathbf{w}} &= \mathbf{K}'\beta^\circ + \mathbf{L}'\hat{\mathbf{g}}, \end{aligned}$$

where

$$\begin{aligned} \beta^\circ &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}, \quad \text{GLS of } \beta \text{ for } \mathbf{y} \sim (\mathbf{X}\beta^*, \mathbf{V}), \\ \mathbf{K}'\beta^\circ &= \text{BLUE of } \mathbf{K}'\beta^* \text{ (not invariant to the value of } \beta\text{), and} \\ \mathbf{L}'\hat{\mathbf{g}} &= \mathbf{L}'\mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta^\circ), \quad \text{the BLUP of } \mathbf{L}'\mathbf{g}. \end{aligned}$$

\Rightarrow The BLUP of $\mathbf{L}'\mathbf{g}$, $\mathbf{L}'\hat{\mathbf{g}}$, is invariant to the value of β , i.e., $\mathbf{L}'\hat{\mathbf{g}}$ is translation invariant.

The translation invariance of the BLUP of $\mathbf{L}'\mathbf{g}$ results as a consequence of the restriction $\mathbf{B}\mathbf{X} = \mathbf{K}'$.

Thus, **the constraint $\mathbf{B}\mathbf{X} = \mathbf{K}'$ causes the BLUP of $\mathbf{L}'\mathbf{g}$ to be translation invariant as well as it insures its unbiasedness.**

Notation:

Let

$$\hat{\mathbf{g}} = \mathbf{C}'\mathbf{P}\mathbf{y}$$

where

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$$

Remarks:

$$\begin{aligned} (1) \quad \mathbf{P}\mathbf{X} &= \mathbf{V}^{-1}\mathbf{X} - \mathbf{V}^{-1}\mathbf{X} \\ &= 0 \end{aligned}$$

$$(2) \quad \mathbf{X}'\mathbf{P} = \mathbf{X}'\mathbf{V}^{-1} - \mathbf{X}'\mathbf{V}^{-1}$$

[10-9]

$$= 0$$

$$\begin{aligned}
 (3) \quad PVP &= (I - V^{-1}X(X'V^{-1}X)^{-1}X')(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}) \\
 &= V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} + V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} \\
 &\quad - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} \\
 &= V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} \\
 &= P
 \end{aligned}$$

$$(4) \quad PV = I - V^{-1}X(X'V^{-1}X)^{-1}X'$$

$$\begin{aligned}
 (5) \quad PVPV &= (I - V^{-1}X(X'V^{-1}X)^{-1}X)(I - V^{-1}X(X'V^{-1}X)^{-1}X') \\
 &= I - V^{-1}(X'V^{-1}X)^{-1}X' - V^{-1}X(X'V^{-1}X)^{-1}X' \\
 &\quad + V^{-1}(X'V^{-1}X)^{-1}X'V^{-1}X(X'V^{-1}X)^{-1}X' \\
 &= PV, \text{ i.e., } \mathbf{PV \text{ is idempotent.}}
 \end{aligned}$$

Properties of the Best Linear Unbiased Predictor.

$$[1] \quad E[\hat{w}] = E[w] \text{ by definition (this was a requirement for } \hat{w} \text{)}$$

$$\begin{aligned}
 E[\hat{w}] &= E[K'\beta^\circ + L'C'V^{-1}(y - X\beta^\circ)] \\
 &= K'\beta + L'C'V^{-1}(X\beta - X\beta) \\
 &= K'\beta \\
 &= E[w]
 \end{aligned}$$

$$\begin{aligned}
 [2] \quad \text{var}(X\beta^\circ) &= X(X'V^{-1}X)^{-1}X'V^{-1} \text{ var}(y) V^{-1}X(X'V^{-1}X)^{-1}X' \\
 &= X(X'V^{-1}X)^{-1}X'V^{-1} VV^{-1}X(X'V^{-1}X)^{-1}X' \\
 &= X(X'V^{-1}X)^{-1}X'
 \end{aligned}$$

$$[3] \quad \text{var}(K'\beta^\circ) = K'(X'V^{-1}X)^{-1}X'V^{-1} \text{ var}(y) V^{-1}X(X'V^{-1}X)^{-1}K$$

[10-10]

$$= K'(X'V^{-1}X)^{-1}K.$$

$$\begin{aligned}[4] \text{cov}(X\beta^\circ, y') &= X(X'V^{-1}X)^{-1}X'V^{-1} \text{var}(y) \\ &= X(X'V^{-1}X)^{-1}X' \\ &= \text{var}(X\beta^\circ) \end{aligned}$$

$$\begin{aligned}[5] \text{var}(L'\hat{g}) &= L'C'P \text{var}(y) PCL \\ &= L'C'PV PCL \\ &= L'C'PCL \\ &= L'C'V^{-1}CL - L'C'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}CL \end{aligned}$$

$$\begin{aligned}[6] \text{cov}(X\beta^\circ, \hat{g}'L) &= X(X'V^{-1}X)^{-1}X'V^{-1} \text{var}(y) PCL \\ &= X(X'V^{-1}X)^{-1}X'V^{-1}VPCL \\ &= X(X'V^{-1}X)^{-1}XV^{-1}CL - X(X'V^{-1}X)^{-1}X'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}CL \\ &= 0 \end{aligned}$$

$$\begin{aligned}[7] \text{var}(\hat{w}) &= \text{var}(K'\beta^\circ + L'\hat{g}) \\ &= \text{var}(K'\beta^\circ) + \text{var}(L'\hat{g}) \text{ by [6]} \end{aligned}$$

$$\begin{aligned}[8] \text{cov}(L'\hat{g}, g'L) &= L'C'P \text{cov}(y, g'L) \\ &= L'C'PCL \\ &= L'C'V^{-1}CL - L'C'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}CL \\ &= \text{var}(L'\hat{g}) \end{aligned}$$

$$\begin{aligned}[9] \text{var}(L'(\hat{g} - g)) &= \text{var}(L'\hat{g}) + \text{var}(L'g) - 2 \text{cov}(L'\hat{g}, g'L) \\ &= \text{var}(L'g) - \text{var}(L'\hat{g}) \\ &= L'GL - L'C'PCL \end{aligned}$$

[10-11]

$$\begin{aligned}
 &= L'[G - C'PC]L \\
 &= L'[G - C'V^{-1}C + C'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}C]L \\
 \text{var}(L'(\hat{g} - g)) &= L'[G - C'V^{-1}C]L + L'C'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}CL \\
 \text{var}(L'(\hat{g} - g)) &= \text{PEV of BLP} + \text{var}(L'C'V^{-1}X\beta^\circ) \\
 \text{var}(L'(\hat{g} - g)) &= \text{var}(BP - L'g) \quad \} \quad \text{PEV of } E[g|y] \text{ (lower bound)} \\
 &\quad + \text{var}(BLP - BP) \quad \} \quad \text{departure from linearity of } E[g|y] \\
 &\quad + \text{var}(L'C'V^{-1}X\beta^\circ) \quad \} \quad \text{variance due to the estimation of } \beta
 \end{aligned}$$

Remark:

$$\begin{aligned}
 \text{var}(L'(\hat{g} - g)) &= \text{var}(L'g) - \text{var}(L'\hat{g}) \\
 \Rightarrow \text{var}(L'\hat{g}) &= \text{var}(L'g) - \text{var}(L'(\hat{g} - g)) \\
 [10] \quad \text{var}(\hat{w} - w) &= \text{var}(K'\beta^\circ + L'\hat{g} - K'\beta - L'g) \\
 &= \text{var}(K'\beta^\circ + L'(\hat{g} - g)) \\
 &= \text{var}(K'\beta^\circ) + \text{var}(L'(\hat{g} - g)) - \text{cov}(K'\beta^\circ, g'L) - \text{cov}(L'g, \beta^\circ K) \\
 &= K'(X'V^{-1}X)^{-1}K + L'[G - C'PC]L - K'(X'V^{-1}X)^{-1}X'V^{-1}CL \\
 &\quad - L'C'V^{-1}X(X'V^{-1}X)^{-1}K
 \end{aligned}$$

[11] The BLUP maximizes the correlation between \hat{w} and w in the class of linear predictors ($a + By$) invariant to β (i.e., translation invariant) with $BX = K' \equiv 0$.

Proof:

$$\begin{aligned}
 r(\hat{w}, w') &= [\text{cov}(\hat{w}, w')][\text{var}(\hat{w}) \text{ var}(w')]^{-1/2} \\
 \text{cov}(\hat{w}, w') &= \text{cov}(K'\beta^\circ + L'\hat{g}, \beta'K + g'L) \\
 &= \text{cov}(K'\beta^\circ, g'L) + \text{cov}(L'\hat{g}, g'L)
 \end{aligned}$$

[10-12]

$$\begin{aligned}
 &= \text{cov}(\mathbf{K}'\beta^\circ, \mathbf{g}'\mathbf{L}) + \text{var}(\mathbf{L}'\hat{\mathbf{g}}) \\
 &= \text{var}(\mathbf{L}'\hat{\mathbf{g}}) \quad \text{because } \mathbf{K}' \equiv \mathbf{0} \\
 \text{var}(\hat{\mathbf{w}}) &= \text{var}(\mathbf{K}'\beta^\circ) + \text{var}(\mathbf{L}'\hat{\mathbf{g}}), \quad \text{from BLUP property [7]} \\
 &= \text{var}((\mathbf{L}'\hat{\mathbf{g}})) \quad \text{if } \mathbf{K} \equiv \mathbf{0}
 \end{aligned}$$

Thus,

$$\begin{aligned}
 r(\hat{\mathbf{w}}, \mathbf{w}') &= r(\mathbf{L}'\hat{\mathbf{g}}, \mathbf{g}'\mathbf{L}), \quad \text{if } \mathbf{K} \equiv \mathbf{0} \\
 &= [\text{var}(\mathbf{L}'\hat{\mathbf{g}})]^{1/2} [\text{var}(\mathbf{L}'\mathbf{g})]^{-1/2} \\
 &= [\text{var}(\mathbf{L}'\hat{\mathbf{g}}) / \text{var}(\mathbf{L}'\mathbf{g})]^{1/2} \\
 &= [[\text{var}(\mathbf{L}'\mathbf{g}) - \text{var}(\mathbf{L}'(\hat{\mathbf{g}} - \mathbf{g}))] / \text{var}(\mathbf{L}'\mathbf{g})]^{1/2}, \quad \text{by BLUP property [9]} \\
 &= [\mathbf{I} - [\text{var}(\mathbf{L}'(\hat{\mathbf{g}} - \mathbf{g})) / \text{var}(\mathbf{L}'\mathbf{g})]]^{1/2}
 \end{aligned}$$

Thus, because the BLUP of \mathbf{w} **minimizes** $\text{var}(\mathbf{L}'(\hat{\mathbf{g}} - \mathbf{g}))$ when $\mathbf{K} \equiv \mathbf{0}$, it also **maximizes** $r(\hat{\mathbf{w}}, \mathbf{w}') = r(\mathbf{L}'\hat{\mathbf{g}}, \mathbf{g}'\mathbf{L})$ when $\mathbf{K} \equiv \mathbf{0}$, i.e.,

$$\begin{aligned}
 \text{as } \text{var}(\mathbf{L}'(\hat{\mathbf{g}} - \mathbf{g})) &\rightarrow 0 \\
 r(\hat{\mathbf{w}}, \mathbf{w}') &\rightarrow \mathbf{I}
 \end{aligned}$$

[12] **Assuming normality,**

$$\begin{bmatrix} \hat{w} \\ w \end{bmatrix} \sim MVN \left\{ \begin{bmatrix} K'\beta \\ K'\beta \end{bmatrix}, \begin{bmatrix} K'(X'V^{-1}X)^{-1}K + L'C'PCL & K'(X'V^{-1}X)^{-1}X'V^{-1}CL + L'C'PCL \\ L'C'V^{-1}X(X'V^{-1}X)^{-1}K + L'C'PCL & L'GL \end{bmatrix} \right\}$$

If $\mathbf{K} \equiv \mathbf{0}$,

$$\begin{bmatrix} L'\hat{g} \\ L'g \end{bmatrix} \sim MVN \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} L'C'PCL & L'C'PCL \\ L'C'PCL & L'GL \end{bmatrix} \right\}$$

[10-13]

[12.1] $X\beta^\circ$ is the maximum likelihood estimator (MLE) of $X\beta$, thus, \hat{w} , the BLUP of w under **normality**, is the MLE of $E[w | y]$.

Proof:

$$\begin{aligned} w &= K'\beta + L'g \\ \begin{bmatrix} y \\ w \end{bmatrix} &\sim MVN \left\{ \begin{bmatrix} X\beta \\ K'\beta \end{bmatrix}, \begin{bmatrix} V & C \\ C' & L'GL \end{bmatrix} \right\} \\ \Rightarrow E[w | y] &= K'\beta + L'C'V^{-1}(y - X\beta) \end{aligned}$$

But, $K' = BX$,

$$\Rightarrow E[w | y] = BX\beta + L'C'V^{-1}(y - X\beta).$$

Because $X\beta^\circ$ is the MLE of $X\beta$, by the **invariance** property of MLE,

$$\begin{aligned} \hat{w} &= BX\beta^\circ + L'C'V^{-1}(y - X\beta^\circ) \\ \hat{w} &= K'\beta^\circ + L'C'V^{-1}(y - X\beta^\circ) \end{aligned}$$

is the MLE of $E[w | y]$.

[12.2] **Under normality, and with $K' \equiv 0$,**

$$(a) E[w | \hat{w}] = E[L'g | L'\hat{g}] = L'\hat{g}$$

Proof:

$$\begin{aligned} E[L'g | L'\hat{g}] &= 0 + L'C'PCL(L'C'PCL)^{-1}(L'\hat{g} - E[L'\hat{g}]) \\ &= I(L'\hat{g} - L[0]) \\ &= L'\hat{g} \end{aligned}$$

$$(b) \text{var}(w | \hat{w}) = \text{var}(L'g | L'\hat{g}) = \text{var}(L'(\hat{g} - g))$$

Proof:

[10-14]

$$\begin{aligned}
 \text{var}(L'g | L' \hat{g}) &= L'GL - L'C'PCL(L'C'PCL)^{-1}L'C'PCL \\
 &= L'GL - L'C'PCL \\
 &= \text{var}(L'g) - \text{var}(L' \hat{g}) \\
 &= \text{var}(L'(\hat{g} - g)) \quad \text{from BLUP property [9]}
 \end{aligned}$$

(c) The ranking on \hat{w} when $K \equiv 0$, i.e., the ranking on $L' \hat{g}$, maximizes the probability of correct pairwise ranking for all pairs $\{L'g_i, L'g_{i'}\}$ in the class of translation invariant linear predictors with mean zero, i.e., with $E[L' \hat{g}] = 0$.

Proof:

Let $t'(L'g)$ be a contrast between two sets of g 's, i.e., $L'g_i - L'g_{i'}$. Then,

$$P\{\text{correct ranking}\} = P\{t'(L'g) > 0 | t'(L' \hat{g}) > 0\} + P\{t'(L'g) < 0 | t'(L' \hat{g}) < 0\}$$

But $E[t'L'g] = 0$. Thus, to maximize $P\{\text{correct pairwise ranking}\}$ is equivalent to maximizing the correlation between $t'(L'g)$ and $t'(L' \hat{g})$, i.e.,

Maximizing

$$P \begin{Bmatrix} \text{correct} \\ \text{pairwise} \\ \text{ranking} \end{Bmatrix} \Leftrightarrow P \begin{Bmatrix} r(t'(L'g), \hat{g}'L't) \\ = \\ t'r(L'g, \hat{g}'L)t \end{Bmatrix}$$

But, by BLUP property [11], $L' \hat{g}$ maximizes $r(L'g, \hat{g}'L)$. Thus, $L' \hat{g}$ also maximizes the probability of correct pairwise ranking in the class of translation invariant linear predictors with zero mean.

[12.3] **Assuming normality, $K \equiv 0$ and $L' = I$** , the BLUP of $w = g$ is equal to the BP of g based on $T'y$, for T' chosen such that $E[T'y] = 0$.

Proof:

Consider a matrix of linear contrasts, T , of rank $n - \text{rank}(X)$, satisfying:

- (i) $T'X = 0 \Rightarrow T'$ is in the null space of X ,
- $\Rightarrow T$ is in the orthogonal complement of the column space of X ,
- (ii) $TT' = T$, and
- (iii) $T'T = I$.

One such T' is:

$$T' = I - X(X'V^{-1}X)^{-1}X'V^{-1}$$

Under **normality**,

$$\begin{bmatrix} y \\ g \end{bmatrix} \sim MVN \left\{ \begin{bmatrix} X\beta \\ 0 \end{bmatrix}, \begin{bmatrix} V & C \\ C' & G \end{bmatrix} \right\}$$

Here we need to know β . However, we can consider linear contrasts of the elements of β , e.g., T^* , such that $T^*X\beta = 0$, which implies that T^*X must be zero because $\beta \neq 0$. Thus, consider T' ,

$$\begin{aligned} E[T'y] &= T'E[y] \\ &= T'X\beta \\ &= (I - X(X'V^{-1}X)^{-1}X'V^{-1})X\beta \\ &= X\beta - X(X'V^{-1}X)^{-1}X'V^{-1}X\beta \\ &= X\beta - X\beta \\ &= 0 \end{aligned}$$

Thus,

[10-16]

$$\begin{bmatrix} T'y \\ g \end{bmatrix} \sim MVN \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} T'VT & T'C \\ C'T & G \end{bmatrix} \right\}$$

and

$$E[g|T'y] = C'T(T'VT)^{-1}T'y$$

But the BLUP of g is:

$$\hat{g} = C'Py$$

where

$$P = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}$$

Then,

$$\begin{aligned} E[g|T'y] &= C'T(T'VT)^{-1}T'y \\ &= \hat{g}, \text{ the BLUP of } g, \end{aligned}$$

and

$$\begin{aligned} \text{var}(g|T'y) &= G - C'T(T'VT)^{-1}T'C \\ &= \text{var}(\hat{g} - g), \text{ the EVP of } g. \end{aligned}$$

Now suppose that:

$$P = T(T'VT)^{-1}T'$$

But P was defined to be:

$$P = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}$$

Thus,

$$\begin{aligned} T(T'VT)^{-1}T' &= V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} \\ \Rightarrow V^{-1} &= T(T'VT)^{-1}T' - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} \end{aligned}$$

[10-17]

$$\Rightarrow I = VT(T'VT)^{-1}T' - X(X'V^{-1}X)^{-1}X'V^{-1}$$

or,

$$I = H_1 + H_2$$

$$I = H$$

where

$$H_1 = VT(T'VT)^{-1}T'$$

$$H_2 = X(X'V^{-1}X)^{-1}X'V^{-1}$$

$$H = H_1 + H_2$$

But for H to be equal to I , it must be idempotent and nonsingular. Thus, it needs to be proven that

$H^2 = H$, and that H is nonsingular of rank n .

(a) Show that H is idempotent, i.e., show that:

$$(H_1 + H_2)^2 = (H_1H_1 + H_1H_2 + H_2H_1 + H_2H_2) = (H_1 + H_2)$$

Thus,

$$H_1H_1 = VT(T'VT)^{-1}T'VT(T'V^{-1}T)^{-1}T'$$

$$= H_1,$$

$$H_2H_2 = X(X'V^{-1}X)^{-1}X'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}$$

$$= H_2,$$

$$H_1H_2 = VT(T'VT)^{-1}T'X(X'V^{-1}X)^{-1}X'V^{-1}$$

$$= 0, \text{ because } T'X = 0, \text{ and}$$

$$H_2H_1 = X(X'V^{-1}X)^{-1}X'V^{-1}VT(T'VT)^{-1}T'X$$

$$= 0, \text{ because } X'V^{-1}VT = X'T = 0.$$

$$\Rightarrow (H_1 + H_2)^2 = (H_1 + 0 + 0 + H_2)$$

[10-18]

$$= H_1 + H_2$$

$$= H$$

\Rightarrow H is idempotent.

(b) Show that H is nonsingular and that $\text{rank}(H) = n$

$$\begin{aligned}\text{rank}(H) &= \text{tr}(H_1 + H_2), \text{ because } H \text{ is idempotent,} \\ &= \text{tr}(H_1) + \text{tr}(H_2) \\ &= \text{rank}(H_1) + \text{rank}(H_2) \\ &= \text{rank}(T) + \text{rank}(X) \\ &= (n - \text{rank}(X)) + \text{rank}(X) \\ &= n\end{aligned}$$

\Rightarrow H is full rank.

But the **only nonsingular idempotent matrix is the identity matrix**. Thus, $H = I$.

$$\Rightarrow T(T'VT)^{-1}T' = P$$

$$\begin{aligned}\Rightarrow E[g | T'y] &= C'T(T'VT)^{-1}T'y \\ &= C'Py \\ &= \hat{g} \\ &= \text{BLUP of } g \text{ when } T'X = 0\end{aligned}$$

$$\Rightarrow \text{var}(g | T'y) = G - C'T(T'VT)^{-1}T'C$$

$$= G - C'PC$$

$$= \text{var}(\hat{g} - g)$$

$$= \text{EVP of } g \text{ when } T'X = 0$$

\Rightarrow Under normality, the BLUP of g is equal to the BP of g based on $T'y$, for T' such that

$$E[T'y] = 0.$$

[12.4] **Under normality and $K' \equiv 0$,** the BLUP of $w = L'g$ is equal to the BP of $L'g$ based on

$L'T'y$, for T' chosen such that $E[T'y] = 0$, i.e.,

$$\text{BLUP of } L'g = \text{BP of } E[L'g | L'T'y], \text{ for } T' \text{ such that } E[T'y] = 0, \text{ by [12.3].}$$

[12.5] **Under normality, $K \equiv 0$ and $L' = I$,** the BLUP of $w = g$ maximizes the $E_s[g_s]$ when the selection rule is to pick out s out of n individuals.

Proof:

$$\begin{aligned} \text{(i) BLUP of } g &= E[g | T'y], \text{ for } T' \text{ such that } E[T'y] = 0, \text{ by [12.3]} \\ &= \text{BP of } g \text{ based on } T'y. \end{aligned}$$

(ii) The BP of g based on $T'y$ maximizes the $E_s[g_s]$ when s out of n individuals are chosen using $\hat{g} = E[g | T'y]$.

Thus, because

$$\hat{g} = E[g | T'y] = \text{BLUP of } g \text{ under normality}$$

→ **under normality,** the BLUP of g maximizes $E_s[g_s]$ when the rule is to select s out of n individuals based on \hat{g} .

[12.6] **Under normality, $K \equiv 0$, $L' = I$, and**

- (i) **animals have the same amount and type of information, and**
- (ii) **animals are unrelated,**

the BLUP of $w = g$ maximizes $E_s[g_s]$ when the selection rule is to select all individuals whose BLUP of g is larger than a truncation value t , i.e., $E_s[g_s]$ maximizes the mean genetic value of the animals in the selected fraction s , where $s = P\{\hat{g}_i \geq t\}$.

[10-20]

Proof:

$$\begin{aligned}\text{BLUP of } g &= E[g | T'y] \text{ for } T' \text{ such that } E[T'y] = 0, \text{ by BLUP property [12.3]} \\ &= \text{BP of } g \text{ given } T'y.\end{aligned}$$

But BP of g given $T'y$ maximizes $E_s[g_s]$ over the selected fraction $s \Rightarrow$ BLUP of g also does it because BLUP of $g = E[g | T'y]$.

References

- Henderson, C. R. 1963. Selection index and expected genetic advance. In: Statistical Genetics and Plant Breeding, NAS-NRC 982, pp 141-163.
- Henderson, C. R. 1973. Sire evaluation and genetic trends. In: Proc. Anim. Breed. and Genet. Symp. in Honor of J.L. Lush, ASAS-ADSA, pp 10-41.
- Quaas, R. L. 1986. Personal Communication. Animal Science 720, Cornell University, Ithaca, NY.
- Searle, S. R. 1971. Linear Models. John Wiley and Sons, Inc., NY.
- Searle, S. R. 1973. Derivation of prediction formulae. Mimeo BU-482-M, Biometrics Unit, Cornell University, Ithaca, NY.
- Searle, S. R. 1974. Prediction, mixed models and variance components. In: Proc. Conf. Reliability and Biometry, Tallahassee, Florida. F. Proschan and R.J. Serfling (Ed.), SIAM, Philadelphia, Pennsylvania. (Mimeo BU-468-M, Biometrics Unit, Cornell Univ, Ithaca, NY, 1973).