Original article

Restricted best linear unbiased prediction using canonical transformation

Y Itoh, H Iwaisaki

Kyoto University, Faculty of Agriculture
Department of Animal Science, Kyoto 606, Japan

Summary – The restricted BLUP procedure requires the solution of high order simultaneous equations if there are many traits and a large number of animals to be evaluated. In this paper, a canonical transformation technique through which new independent traits are introduced is presented. Thus only equations of relatively low order for each transformed trait have to be solved. Furthermore, it is shown that the number of independent transformed traits is reduced by the number of restrictions imposed. The technique is applicable when a multiple-trait animal model is assumed.

restricted BLUP / canonical transformation / multiple trait animal model

Résumé – **Utilisation de la transformation canonique pour calculer le meilleur prédicteur linéaire sans biais avec restrictions.** *La procédure du BLUP restreint demande la résolution d'un système d'équations simultanées d'ordre élevé, s'il y a beaucoup de caractères et un grand nombre d'animaux évalués. Dans cette étude, la technique de la transformation canonique est présentée pour obtenir des caractères transformés indépendants. Il suffit alors de résoudre un nombre moins élevé d'équations pour chaque caractère transformé. De plus, le nombre des caractères indépendants est réduit du nombre des restrictions imposées. Cette technique est applicable quand un modèle multicaractère animal est posé.*

BLUP restreint / transformation canonique / modèle multicaractère animal

INTRODUCTION

Kempthorne and Nordskog (1959) proposed the restricted selection index which is a modification of the usual selection index for predicting genetic merits. Selection based on this index can change population means of some traits, but holds some linear functions of other traits unchanged. Quaas and Henderson (1976a) extended the restricted selection index and proposed the restricted best linear unbiased prediction (restricted BLUP) procedure which could include observations with unknown means, missing observations and related animals. They suggested that it provides a useful selection criterion, for example, for altering the growth curve of beef cattle in a favorable manner (Quaas and Henderson, 1976b). However, this

method requires the solution of high order simultaneous equations if there are many traits and a large number of individuals to be evaluated.

Of course, this difficulty holds true for the ordinary multiple-trait BLUP evaluation (Henderson and Quaas, 1976). Such a computational difficulty can be overcome in several ways. One of them is an application of the canonical transformation technique (Thompson, 1977; Lee, 1979; Arnason, 1982) through which new independent traits are introduced, and consequently only mixed model equations of relatively smaller order for each trait need to be solved.

The objective of this paper is to discuss the application of canonical transformation to the restricted BLUP.

THEORY

A multiple-trait animal model with the number of traits expressed as q is assumed. The model for the i-th trait is written as:

$$\mathbf{y}_i = \mathbf{X}_0\boldsymbol{\beta}_i + \mathbf{Z}_0\mathbf{u}_i + \mathbf{e}_i$$

where:
   $\mathbf{y}_i$ is a vector of observations for the i-th trait,
   $\mathbf{X}_0$ is an incidence matrix relating fixed effects to observations,
   $\boldsymbol{\beta}_i$ is a vector of unknown fixed effects,
   $\mathbf{Z}_0$ is an incidence matrix relating $\mathbf{u}_i$ to observations,
   $\mathbf{u}_i$ is a vector of unknown additive genotypic values of animals and
   $\mathbf{e}_i$ is a vector of errors.

It is assumed that $\mathbf{X}_0$ and $\mathbf{Z}_0$ are the same for all traits. Only the records of individuals who have records on all the traits or who have no record on any trait are used, but the records of individuals whose records are partially missing on some traits are not used. The number of individuals to be evaluated is denoted by p, the number of individuals with records by n and the number of the columns of $\mathbf{X}_0$ by f.

When the records are ordered by individuals within traits, the model for all the traits is written as:

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

where

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_q \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_q \end{bmatrix}, \mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_q \end{bmatrix}, \mathbf{e} = \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_q \end{bmatrix},$$

$$\mathbf{X} = \mathbf{I}_q * \mathbf{X}_0 \text{ and } \mathbf{Z} = \mathbf{I}_q * \mathbf{Z}_0,$$

$\mathbf{I}_q$ denotes the identity matrix of order $q \times q$ and $*$ denotes the direct product operation. It is assumed that

$$E\begin{bmatrix} u \\ e \end{bmatrix} = O \text{ and } Var\begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} G & O \\ O & R \end{bmatrix}$$

where

$$G = G_0 * A, \quad R = R_0 * I_n,$$

A is the numerator relationship matrix among individuals to be evaluated, G_0 and R_0 are the additive genotypic and error variance-covariance matrices among traits, and I_n is the identity matrix of order $n \times n$.

Consider the restricted BLUP proposed by Quaas and Henderson (1976a). In that method, a linear predictor $b'y$ is used which is uncorrelated with some linear function of u, say $C'u$. This is expressed algebraically as $Cov(b'y, C'u) = b'ZGC = O$. If the same constraints are imposed on the additive genotypic values of all animals, then C is expressed as $C = C_0 * I_n$ where C_0 is a matrix of order $q \times r$ and the same as that used by Kempthorne and Nordskog (1959). The columns of C_0 are assumed to be linearly independent. Subject to this additional constraint, the best linear unbiased predictor can be derived. Such a predictor of u, denoted by $\hat{u}$, is obtained by solving the following equations:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z & X'R^{-1}ZGC \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} & Z'R^{-1}ZGC \\ C'GZ'R^{-1}X & C'GZ'R^{-1}Z & C'GZ'R^{-1}ZGC \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \\ \hat{t} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \\ C'GZ'R^{-1}y \end{bmatrix} .(1)$$

Eliminating $\hat{t}$ by absorption gives:

$$\begin{bmatrix} X'SX & X'SZ \\ Z'SX & Z'SZ + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'Sy \\ Z'Sy \end{bmatrix} \tag{2}$$

where

$$\begin{aligned} S &= R^{-1} - R^{-1}ZGC(C'GZ'R^{-1}ZGC)^-C'GZ'R^{-1} \\ &= R_0^{-1} * I_n - [R_0^{-1}G_0C_0(C_0'G_0R_0^{-1}G_0C_0)^{-1}C_0'G_0R_0^{-1}] \\ &\quad * [Z_0A(AZ_0'Z_0A)^-AZ_0']. \end{aligned}$$

Note that:

$$Z_0A(AZ_0'Z_0A)^-AZ_0' = Z_0(Z_0'Z_0)^-Z_0' = I_n \tag{3}$$

in animal models. Using this, S can be rewritten as:

$$S = S_0 * I_n$$

where

$$S_0 = R_0^{-1} - R_0^{-1}G_0C_0(C_0'G_0R_0^{-1}G_0C_0)^{-1}C_0'G_0R_0^{-1}. \tag{4}$$

Note that S_0 has rank $q - r$.

Because G_0 is positive definite and S_0 is positive semidefinite, there exists a non-singular matrix Q such that $Q'G_0^{-1}Q = I_q$ and $Q'S_0Q = D$ where D is a diagonal matrix whose diagonal elements, denoted by $\lambda_1 \geq \ldots \geq \lambda_{q-r}(> 0)$ and $\lambda_{q-r+1} = \ldots = \lambda_q(= 0)$, are the roots of the equation:

$$|S_0 - \lambda G_0^{-1}| = 0.$$

Such a matrix Q and λ_i's are easily obtainable through general program packages.

Premultiplying $\begin{bmatrix} Q' * I_f & O \\ O & Q' * I_p \end{bmatrix}$ where I_f and I_p are identity matrices of order $f \times f$ and $p \times p$, equations (2) can be modified as:

$$\begin{bmatrix} Q' * I_f & O \\ O & Q' * I_p \end{bmatrix} \begin{bmatrix} X'SX & X'SZ \\ Z'SX & Z'SZ + G^{-1} \end{bmatrix} \begin{bmatrix} Q * I_f & O \\ O & Q * I_p \end{bmatrix}$$

$$\begin{bmatrix} Q^{-1} * I_f & O \\ O & Q^{-1} * I_p \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix}$$

$$= \begin{bmatrix} Q' * I_f & O \\ O & Q' * I_p \end{bmatrix} \begin{bmatrix} X'S(Q * I_n)(Q^{-1} * I_n)y \\ Z'S(Q * I_n)(Q^{-1} * I_n)y \end{bmatrix}$$

or

$$\begin{bmatrix} D * (X_0'X_0) & D * (X_0'Z_0) \\ D * (Z_0'X_0) & D * (Z_0'Z_0) + I_q * A^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta}^* \\ \hat{u}^* \end{bmatrix} = \begin{bmatrix} (D * X_0')y^* \\ (D * X_0')y^* \end{bmatrix} \tag{5}$$

where

$$\hat{\beta}^* = (Q^{-1} * I_f)\hat{\beta} = \begin{bmatrix} \hat{\beta}_1^* \\ \vdots \\ \hat{\beta}_q^* \end{bmatrix}, \quad \hat{u}^* = (Q^{-1} * I_p)\hat{u} = \begin{bmatrix} \hat{u}_1^* \\ \vdots \\ \hat{u}_q^* \end{bmatrix},$$

$$y^* = (Q^{-1} * I_n)y = \begin{bmatrix} y_1^* \\ \vdots \\ y_q^* \end{bmatrix}. \tag{6}$$

Because D is diagonal, equations (5) can be subdivided into q independent equation systems which have different forms depending on λ_i. When λ_i is nonzero ($i = 1$ to $q - r$), the equations for the i-th transformed trait become:

$$\begin{bmatrix} X_0'X_0 & X_0'Z_0 \\ Z_0'X_0 & Z_0'Z_0 + \lambda_i^{-1}A^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta}_i^* \\ \hat{u}_i^* \end{bmatrix} = \begin{bmatrix} X_0'y_i^* \\ X_0'y_i^* \end{bmatrix}. \tag{7}$$

These equations can be solved with a computing program for single trait BLUP. On the other hand, when λ_i is zero ($i = q - r + 1$ to q), the equations are reduced to:

$$\begin{bmatrix} O & O \\ O & A^{-1} \end{bmatrix} \begin{bmatrix} \widehat{\beta}_i^* \\ \widehat{u}_i^* \end{bmatrix} = \begin{bmatrix} O \\ O \end{bmatrix} ; \tag{8}$$

thus in every case $\widehat{u}_i^* = 0$ and $\widehat{\beta}_i^*$ is indefinite.

From the facts stated above, the restricted BLUP can be computed easily as follows without directly solving equations (1) of high order. First, transform the observed records by (6). Then compute $\widehat{u}_i^*$ by solving equations (7) for each of the first $q - r$ transformed traits, and set $\widehat{u}_i^* = O$ for the remaining r traits. Finally, obtain $\widehat{u}$ by the inverse transformation:

$$\widehat{u} = (Q * I_p)\widehat{u}^*. \tag{9}$$

The solution $\widehat{u}$ derived in this way is identical to the solution $\widehat{u}$ given by solving equations (1). However, the fixed effects are not estimable because some β_i^*'s are indefinite.

NUMERICAL EXAMPLE

The following example including 2 traits, birth weight and weaning weight, illustrates the use of the method outlined above:

Animal	Birth wt	Weaning wt
1	61	362
2	72	401
3	68	350
4	—	—

There are 4 animals to be evaluated, but animal no 4 has no record. The numerator relationship matrix among them is:

$$A = \begin{bmatrix} 1.00 & 0.25 & 0.25 & 0.25 \\ 0.25 & 1.00 & 0.25 & 0.25 \\ 0.25 & 0.25 & 1.00 & 0.25 \\ 0.25 & 0.25 & 0.25 & 1.00 \end{bmatrix}$$

The genetic and error variance-covariance matrices are:

$$G_0 = \begin{bmatrix} 28.60 & 73.77 \\ 73.77 & 566.00 \end{bmatrix}, \quad R_0 = \begin{bmatrix} 36.30 & 67.43 \\ 67.43 & 1\,454.00 \end{bmatrix}.$$

The fixed effects in the model are only common means. Thus the matrices and vectors included in the model are expressed as:

$$y' = [61 \quad 72 \quad 68 \quad 362 \quad 401 \quad 350],$$

$$\mathbf{X}_0' = [1 \quad 1 \quad 1],$$

$$\boldsymbol{\beta}' = [\mu_1 \quad \mu_2],$$

$$\mathbf{Z}_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Suppose that it is desirable to improve weaning weight but to keep birth weight unchanged, then $\mathbf{C}_0$ is:

$$\mathbf{C}_0' = [1 \quad 0].$$

First, the direct solution of restricted BLUP will be shown. Equations (1) become (10). Solving these equations gives:

[67.620 85   372.601 40   0.0   0.0   0.0   0.0   1.119 97   2.957 75

−4.077 72   0.0   − 0.227 86   0.162 73   0.0   0.0].

Thus the predicted additive genetic values are:

$$\hat{\mathbf{u}}_1' = [0.0 \quad 0.0 \quad 0.0 \quad 0.0], \tag{11}$$

$$\hat{\mathbf{u}}_2' = [1.119\,97 \quad 2.957\,75 \quad -4.077\,72 \quad 0.0]. \tag{12}$$

Next, the procedure using the canonical transformation will be shown. From (4), $\mathbf{S}_0$ becomes:

$$\mathbf{S}_0 = \begin{bmatrix} 0.004\,94 & -0.001\,91 \\ -0.001\,91 & 0.000\,74 \end{bmatrix}.$$

Using a program package, the following matrices can be computed from $\mathbf{S}_0$ and $\mathbf{G}_0^{-1}$:

$$\mathbf{Q} = \begin{bmatrix} 0.0 & 5.347\,90 \\ 19.383\,49 & 13.794\,21 \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} 0.278\,80 & 0.0 \\ 0.0 & 0.0 \end{bmatrix}.$$

Observations transformed by (6) are:

$$\mathbf{y}_1^{*'} = [10.558\,39 \quad 11.106\,63 \quad 9.007\,81],$$

$$\mathbf{y}_2^{*'} = [11.406\,35 \quad 13.463\,24 \quad 12.715\,28].$$

$\mathbf{y}_1^*$ is used in the next step, but $\mathbf{y}_2^*$ will not be used any more. Equations (7) for the first transformed trait where $\lambda_1 = 0.278\,80$ become:

$$
\begin{bmatrix}
0.09044 & -0.00419 & 0.03015 & 0.03015 & 0.03015 & 0.0 & -0.00140 & -0.00140 & -0.00140 & 0.0 & 0.75902 & 0.75902 & 0.75902 & 0.0 \\
-0.00419 & 0.00226 & -0.00140 & -0.00140 & -0.00140 & 0.0 & 0.00075 & 0.00075 & 0.00075 & 0.0 & 0.01554 & 0.01554 & 0.01554 & 0.0 \\
0.03015 & -0.00140 & 0.09034 & -0.01003 & -0.01003 & -0.01003 & -0.00924 & 0.00131 & 0.00131 & 0.00131 & 0.75902 & 0.0 & 0.0 & 0.0 \\
0.03015 & -0.00140 & -0.01003 & 0.09034 & -0.01003 & -0.01003 & 0.00131 & -0.00924 & 0.00131 & 0.00131 & 0.0 & 0.75902 & 0.0 & 0.0 \\
0.03015 & -0.00140 & -0.01003 & -0.01003 & 0.09034 & -0.01003 & 0.00131 & 0.00131 & -0.00924 & 0.00131 & 0.0 & 0.0 & 0.75902 & 0.0 \\
0.0 & 0.0 & -0.01003 & -0.01003 & -0.01003 & 0.06020 & 0.00131 & 0.00131 & 0.00131 & -0.00785 & 0.0 & 0.0 & 0.0 & 0.0 \\
-0.00140 & 0.00075 & -0.00924 & 0.00131 & 0.00131 & 0.00131 & 0.00379 & -0.00051 & -0.00051 & -0.00051 & 0.01554 & 0.0 & 0.0 & 0.0 \\
-0.00140 & 0.00075 & 0.00131 & -0.00924 & 0.00131 & 0.00131 & -0.00051 & 0.00379 & -0.00051 & -0.00051 & 0.0 & 0.01554 & 0.0 & 0.0 \\
-0.00140 & 0.00075 & 0.00131 & 0.00131 & -0.00924 & 0.00131 & -0.00051 & -0.00051 & 0.00379 & -0.00051 & 0.0 & 0.0 & 0.01554 & 0.0 \\
0.0 & 0.0 & 0.00131 & 0.00131 & 0.00131 & -0.00785 & -0.00051 & -0.00051 & -0.00051 & 0.00304 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.75902 & 0.01554 & 0.75902 & 0.0 & 0.0 & 0.0 & 0.01554 & 0.0 & 0.0 & 0.0 & 22.85405 & 0.0 & 0.0 & 0.0 \\
0.75902 & 0.01554 & 0.0 & 0.75902 & 0.0 & 0.0 & 0.0 & 0.01554 & 0.0 & 0.0 & 0.0 & 22.85405 & 0.0 & 0.0 \\
0.75902 & 0.01554 & 0.0 & 0.0 & 0.75902 & 0.0 & 0.0 & 0.0 & 0.01554 & 0.0 & 0.0 & 0.0 & 22.85405 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0
\end{bmatrix}
\begin{bmatrix}
\hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{u}_{11} \\ \hat{u}_{12} \\ \hat{u}_{13} \\ \hat{u}_{14} \\ \hat{u}_{21} \\ \hat{u}_{22} \\ \hat{u}_{23} \\ \hat{u}_{24} \\ \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \hat{\varepsilon}_3 \\ \hat{\varepsilon}_4
\end{bmatrix}
=
\begin{bmatrix}
4.50320 \\ 0.55664 \\ 1.33278 \\ 1.60985 \\ 1.56057 \\ 0.0 \\ 0.18716 \\ 0.20113 \\ 0.16834 \\ 0.0 \\ 51.92422 \\ 60.87933 \\ 57.05092 \\ 0.0
\end{bmatrix}
\tag{10}
$$

$$\begin{bmatrix} 3.0 & 1.0 & 1.0 & 1.0 & 0.0 \\ 1.0 & 5.099\,27 & -0.683\,21 & -0.683\,21 & -0.683\,21 \\ 1.0 & -0.683\,21 & 5.099\,27 & -0.683\,21 & -0.683\,21 \\ 1.0 & -0.683\,21 & -0.683\,21 & 5.099\,27 & -0.683\,21 \\ 0.0 & -0.683\,21 & -0.683\,21 & -0.683\,21 & 4.099\,27 \end{bmatrix} \begin{bmatrix} \hat{\mu}_1^* \\ \hat{u}_{11}^* \\ \hat{u}_{12}^* \\ \hat{u}_{13}^* \\ \hat{u}_{14}^* \end{bmatrix} = \begin{bmatrix} 30.672\,82 \\ 10.558\,39 \\ 11.106\,63 \\ 9.007\,81 \\ 0.0 \end{bmatrix}$$

The solution is:

$$[10.224\,27 \quad 0.057\,78 \quad 0.152\,59 \quad -0.210\,37 \quad 0.0],$$

thus:

$$\hat{u}_1^{*\prime} = [0.057\,78 \quad 0.152\,59 \quad -0.210\,37 \quad 0.0].$$

As to the second transformed trait, set $\hat{u}_2^* = \mathbf{O}$. Thus $\hat{u}^*$ is obtained. Finally, transforming $\hat{u}^*$ by (9) gives the predicted additive genetic values which are identical to (11) and (12).

DISCUSSION

It has been shown that a canonical transformation technique is applicable to the restricted BLUP evaluation when an animal model is assumed. It is not necessary to solve the independent equation systems for all the transformed traits. The number of equation systems to be solved becomes $q - r$, the number of traits minus the number of constraints. Furthermore, a standard computing program for single trait BLUP is applicable to the restricted BLUP evaluation, and no special program is needed. Therefore, the computational task to obtain the solution is much reduced.

However, this method is applicable only to animal models. When the model used is not an animal model, equality (3) does not hold. Consequently, no simple formula can be obtained after that. This method also has another limitation: it assumes that the models are the same for all traits and there are no partially missing observations. Hence, the method is not applicable to, for example, multiple lactation records in dairy cows where the number of lactations is not constant for all cows generally.

It was shown that there is no estimable function of the fixed effects in general. This also holds true when the solutions are obtained directly using (1), as pointed out by Quaas and Henderson (1976a), who expressed it as "the linear dependencies among the fixed effect equations". Therefore, the restricted BLUP is available only when animals are evaluated on some functions of $\hat{u}$ that do not include any element of $\hat{\beta}$.

REFERENCES

Arnason T (1982) Prediction of breeding values for multiple traits in small non-random mating (horse) populations. *Acta Agric Scand* 32, 171-176

Henderson CR, Quaas RL (1976) Multiple trait evaluation using relatives' records. *J Anim Sci* 43, 1188-1197

Kempthorne O, Nordskog AW (1959) Restricted selection index. *Biometrics* 15, 10-19

Lee AJ (1979) Mixed model, multiple evaluation of related sires when all traits are recorded. *J Anim Sci* 48, 1079-1088

Quaas RL, Henderson CR (1976a) Restricted best linear unbiased prediction of breeding values. Mimeo, Cornell University

Quaas RL, Henderson CR (1976b) Selection criteria for altering the growth curve. *J Anim Sci* (abstr) 43, 221

Thompson R (1977) Estimation of quantitative genetic parameters, p 639-658. *In· Proc Int Conf Quantitative Genetics* (Pollak E, Kempthorne O, Bailey TB Jr, eds) August 16-21, 1976. Iowa State University Press, Ames, IA