J. Dairy Sci. 97:5880–5884
http://dx.doi.org/10.3168/jds.2014-8155

Short communication: Alteration of priors for random effects in Gaussian linear mixed models

J. Vandenplas,*†[1] O. F. Christensen,‡ and N. Gengler*
*Animal Science Unit, Gembloux Agro Bio-Tech, University of Liege, 5030 Gembloux, Belgium
†National Fund for Scientific Research, 1000 Brussels, Belgium
‡Center for Quantitative Genetics and Genomics, Dept. Molecular Biology and Genetics, Aarhus University, 8830 Tjele, Denmark

ABSTRACT

Linear mixed models, for which the prior multivariate normal distributions of random effects are assumed to have a mean equal to 0, are commonly used in animal breeding. However, some statistical analyses (e.g., the consideration of a population under selection into a genomic scheme breeding, multiple-trait predictions of lactation yields, and Bayesian approaches integrating external information into genetic evaluations) need to alter both the mean and (co)variance of the prior distributions and, to our knowledge, most software packages available in the animal breeding community do not permit such alterations. Therefore, the aim of this study was to propose a method to alter both the mean and (co)variance of the prior multivariate normal distributions of random effects of linear mixed models while using currently available software packages. The proposed method was tested on simulated examples with 3 different software packages available in animal breeding. The examples showed the possibility of the proposed method to alter both the mean and (co)variance of the prior distributions with currently available software packages through the use of an extended data file and a user-supplied (co)variance matrix.
Key words: prior distribution, Bayesian, linear mixed model

Short Communication

Currently, Henderson's mixed models methods and BLUP (Henderson, 1975) are commonly used in animal breeding. The typical linear mixed model is written as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \qquad [1]$$

where $\mathbf{y}$ is the vector of records; $\boldsymbol{\beta}$ and $\mathbf{u}$ are vectors of fixed and random effects related to the records through the incidence matrices $\mathbf{X}$ and $\mathbf{Z}$, respectively; and $\mathbf{e}$ is the vector of residuals.

It is assumed that the expectations are

$$E\begin{bmatrix}\mathbf{u}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}$$

and the (co)variance matrices are

$$Var\begin{bmatrix}\mathbf{u}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{G} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}\end{bmatrix},$$

where $\mathbf{G}$ is the (co)variance matrix associated with $\mathbf{u}$ and $\mathbf{R}$ is the (co)variance matrix associated with $\mathbf{e}$. The estimates of $\boldsymbol{\beta}$; that is, $\hat{\boldsymbol{\beta}}$, and the predictions of $\mathbf{u}$; that is, $\hat{\mathbf{u}}$, can be obtained solving the mixed-model equations written as follows (Henderson, 1950):

$$\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1}\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}. \qquad [2]$$

In the case of BLUP, the (co)variance matrices $\mathbf{G}$ and $\mathbf{R}$ are assumed to be known.

Although $\mathbf{u}$ is assumed to have an expectation equal to $\mathbf{0}$, a need exists to alter this expectation in some statistical analyses. For example, Bayesian approaches integrating external information (i.e., EBV and associated reliabilities obtained from a foreign genetic evaluation) into a local genetic evaluation alter both $E[\mathbf{u}]$ and $Var[\mathbf{u}]$ (e.g., Gianola and Fernando, 1986; Quaas and Zhang, 2006; Legarra et al., 2007). For such approaches, $E[\mathbf{u}]$ is equal to the foreign EBV and $Var[\mathbf{u}]$ represents the associated matrix of prediction error (co)variances. Another example is the consideration of a population under selection into a genomic scheme breeding by assuming $E[\mathbf{u}] \neq \mathbf{0}$ for the genotyped animals. Indeed, they may have an expectation different from 0 if selection occurred (Vitezica et al., 2011). Also, Schaeffer and Jamrozik (1996) proposed a multiple-trait procedure for predicting lactation yields for dairy cows based on an alteration of $E[\mathbf{u}]$ with information from groups

Received March 21, 2014.
Accepted June 9, 2014.
[1]Corresponding author: jvandenplas@ulg.ac.be

of cows sharing the same production characteristics. However, to our knowledge, most software packages currently available in animal breeding do not permit alterations of expectations of random effects, whereas they may allow the use of a user supplied covariance matrix. Therefore, the aim of this study was to propose a method to alter both the expectations and (co)variances of random effects while using software packages currently available in animal breeding. The development of the proposed method was based on a Bayesian view of linear mixed models.

Bayesian View of Linear Mixed Models

Bayes estimators for linear (mixed) models and their relations with BLUP were discussed by several authors (e.g., Lindley and Smith, 1972; Dempfle, 1977; Gianola and Fernando, 1986; Sorensen and Gianola, 2002). From a Bayesian view, all fixed and random effects are considered as random. However, the terms "fixed" and "random" will still be used below to differentiate $\boldsymbol{\beta}$ from $\mathbf{u}$, respectively.

For the linear model 1, the following prior multivariate normal (MVN) distributions are assumed:

$$[\boldsymbol{\beta}|\mathbf{B}] \sim MVN(\mathbf{b},\mathbf{B}),$$

where $\mathbf{b}$ is a mean vector and $\mathbf{B}$ is a (co)variance matrix,

$$[\mathbf{u}|\mathbf{G}] \sim MVN(\mathbf{g},\mathbf{G}),$$

where $\mathbf{g}$ is a mean vector, and

$$[\mathbf{e}|\mathbf{R}] \sim MVN(\mathbf{0},\mathbf{R}).$$

Assuming that all the (co)variance matrices are known, the joint posterior density of $\boldsymbol{\beta}$ and $\mathbf{u}$ can be written as follows:

$$f(\boldsymbol{\beta},\mathbf{u} \mid \mathbf{y},\mathbf{B},\mathbf{G},\mathbf{R}) \propto \exp\left\{-\frac{1}{2}\begin{bmatrix}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\mathbf{u})'\mathbf{R}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\mathbf{u})\\ +(\boldsymbol{\beta}-\mathbf{b})'\mathbf{B}^{-1}(\boldsymbol{\beta}-\mathbf{b})+(\mathbf{u}-\mathbf{g})'\mathbf{G}^{-1}(\mathbf{u}-\mathbf{g})\end{bmatrix}\right\}.$$

Because this joint posterior distribution is multivariate normal, its mean equals its mode, and $\boldsymbol{\beta}$ and $\mathbf{u}$ can be estimated by differentiating the joint posterior distribution with respect to $\boldsymbol{\beta}$ and $\mathbf{u}$ and setting the derivatives equal to zero. From this, as shown by Gianola and Fernando (1986), the following equation is obtained:

$$\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}+\mathbf{B}^{-1} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1}\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{u}}\end{bmatrix}=\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}+\mathbf{B}^{-1}\mathbf{b}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}+\mathbf{G}^{-1}\mathbf{g}\end{bmatrix}. \quad [3]$$

If a noninformative prior is considered for $\boldsymbol{\beta}$ (i.e., $\mathbf{B}^{-1}\rightarrow\mathbf{0}$) and $\mathbf{g}=\mathbf{0}$, the system of equations [3] simplifies to traditional mixed-model equations [2].

Alteration of Priors for Random Effects

In the following development, it is assumed that $\mathbf{g}$ and $\mathbf{G}^{-1}$ are known, $\mathbf{g}\neq\mathbf{0}$, and a noninformative prior for $\boldsymbol{\beta}$ is considered. Therefore, the system of equations [3] is written as follows:

$$\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1}\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{u}}\end{bmatrix}=\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}+\mathbf{G}^{-1}\mathbf{g}\end{bmatrix}. \quad [4]$$

Although most available software packages allow the use of a user-supplied (co)variance matrix as $\mathbf{G}$, or its inverse $\mathbf{G}^{-1}$, most of them do not allow that $\mathbf{g}\neq\mathbf{0}$. Thereby, the system of equations [4], allowing an alteration of the default mean, cannot be solved with current software packages. A way to solve this issue is to develop a system of equations equivalent to the system of equations [4] that can be solved by current software packages. Therefore, below we define $\mathbf{y_P}$, a vector of pseudo-records (i.e., records corrected for all other effects than $\mathbf{u}$); $\mathbf{X_P}$ and $\mathbf{Z_P}$, 2 incidence matrices relating pseudo-records to $\boldsymbol{\beta}$ and $\mathbf{u}$, respectively; $\mathbf{R_P}$, a residual (co)variance matrix associated to the pseudo-records; and $\mathbf{G}^*$, a (co)variance matrix associated with $\mathbf{u}$ conditional on pseudo-records. Assuming that $\mathbf{X_P}=\mathbf{0}$, $\mathbf{Z}'_\mathbf{P}\mathbf{R}_\mathbf{P}^{-1}\mathbf{y_P}=\mathbf{G}^{-1}\mathbf{g}$, and $\mathbf{G}^{-1}=\mathbf{G}^{*-1}+\mathbf{Z}'_\mathbf{P}\mathbf{R}_\mathbf{P}^{-1}\mathbf{Z_P}$, the system of equations [4] is equivalent to

$$\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}+\mathbf{X}'_\mathbf{P}\mathbf{R}_\mathbf{P}^{-1}\mathbf{X_P} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{X}'_\mathbf{P}\mathbf{R}_\mathbf{P}^{-1}\mathbf{Z_P}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}+\mathbf{Z}'_\mathbf{P}\mathbf{R}_\mathbf{P}^{-1}\mathbf{X_P} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{*-1}+\mathbf{Z}'_\mathbf{P}\mathbf{R}_\mathbf{P}^{-1}\mathbf{Z_P}\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{u}}\end{bmatrix}=\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}+\mathbf{X}'_\mathbf{P}\mathbf{R}_\mathbf{P}^{-1}\mathbf{y_P}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}+\mathbf{Z}'_\mathbf{P}\mathbf{R}_\mathbf{P}^{-1}\mathbf{y_P}\end{bmatrix}. \quad [5]$$

The system of equations [5] can be written using compact notation as

$$\begin{bmatrix}\mathbf{X}^{*\prime}\mathbf{R}^{*-1}\mathbf{X}^* & \mathbf{X}^{*\prime}\mathbf{R}^{*-1}\mathbf{Z}^*\\ \mathbf{Z}^{*\prime}\mathbf{R}^{*-1}\mathbf{X}^* & \mathbf{Z}^{*\prime}\mathbf{R}^{*-1}\mathbf{Z}^*+\mathbf{G}^{*-1}\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{u}}\end{bmatrix}=\begin{bmatrix}\mathbf{X}^{*\prime}\mathbf{R}^{*-1}\mathbf{y}^*\\ \mathbf{Z}^{*\prime}\mathbf{R}^{*-1}\mathbf{y}^*\end{bmatrix},$$