Bayesian Analysis of Multivariate Threshold Animal Models Using Gibbs Sampling[†]

Seung-Chun Lee[1] and Deukhwan Lee[2]

ABSTRACT

The estimation of variance components or variance ratios in linear model is an important issue in plant or animal breeding fields, and various estimation methods have been devised to estimate variance components or variance ratios. However, many traits of economic importance in those fields are observed as dichotomous or polychotomous outcomes. The usual estimation methods might not be appropriate for these cases. Recently threshold linear model is considered as an important tool to analyze discrete traits specially in animal breeding field.

In this note, we consider a hierarchical Bayesian method for the threshold animal model. Gibbs sampler for making full Bayesian inferences about random effects as well as fixed effects is described to analyze jointly discrete traits and continuous traits. Numerical example of the model with two discrete ordered categorical traits, calving ease of calves from born by heifer and calving ease of calf from born by cow, and one normally distributed trait, birth weight, is provided.

Keywords. Threshold animal model, liability, hierarchical Bayes method, Gibbs sampling.
AMS 2000 subject classifications. Primary 62F15, 62J10; Secondary 62F10.

1. Introduction

In an analysis of mixed effects linear model, one objective may be inference about the variance components or the ratio of variance components. For instance, in the context of animal breeding, variance components or ratio of variance components often represent the heritability of some trait and, under certain assumptions, heritability can be expressed as a function of them. However, dependent

variables in a linear model are often measured in discrete scale, and hence the usual variance component estimation methods under normality assumption is not appropriate. For example, in American Gelbvieh Association, calving ease is scored as 1 (natural calving, no assistance), 2 (easy pull), 3 (hard pull) or 4 (mechanical force or Cesarean). Because the calving ease is an important trait to determine the economic loss for heifer and cow as well as calf in reproduction, various methodologies have been devised to analysis calving ease score.

One of the appealing methodology for genetic analysis of ordered categorical data is based on the threshold liability concept, which was originated by Wright (1934) in studies of the number of digits in guinea pigs. In the threshold model, it is postulated that there exists a latent or underlying variable (liability) which has a continuous distribution. A set of thresholds divides this continuous variable into the discrete scores. Thus the observed value of categorical variable is a representation of the liability falling between the thresholds. Applications of this model can be found in Gianola (1982), Foulley et al. (1987), Wang et al. (1994), Sorensen et al. (1994), Jensen et al. (1994) and Berger et al. (1995).

With the advent of inferential algorithms based simulations, Bayesian methods are being increasingly applied to genetic inference in animal breeding. This is partly due to the fact that the complex analytic solutions of Bayesian method are now feasible with the help of the inferential algorithms. The most popular family of such algorithm is Markov Chain Monte Carlo (MCMC) including Gibbs sampling (Geman and Geman, 1984; Gelfand et al., 1990). On the other hand certain genetic inference requires a super population model or Bayes model in the sense the number of parameters to be estimated is greater than or equal to the number of observations, which often arises in animal breeding, and is referred to as the 'animal model' (Quass et al., 1980).

Gianola and Foulley (1983) described Bayesian model of single trait thresholds model assuming known genetic variance. Foulley et al. (1983) developed a method to deal with a binary trait and two continuous traits. This result was generalized to the situation of one multiple ordered categorical trait and several continuous traits. Sorensen et al. (1995) published the Bayesian analysis of univariate thresholds model via Gibbs sampling. This result further extended by Wang et al. (1997) to analysis one multiple ordered categorical trait and one continuous trait. However, we noted that the number of traits does not play much roles in analytic models as long as the discrete traits have multiple ordered categories. In this note we will describe a general hierarchical Bayes threshold animal model via Gibbs sampling.

The main purpose of this paper is to analysis the American Gelbvieh Association's calving ease data. The edited data is consist of 322,956 records from calves born from 1981 to 2000. 3 traits, one continuous traits, birth weight (BWT), and two ordered categorial traits, calving ease from born by heifer (CEH) and calving ease from born by cow (CEC), are being analyzed by calf's sex and age of dam group, which are assumed to be fixed effects, and herd-year-season effect, direct and maternal additive genetic effects, which are considered as random effects. In particular the direct and the maternal genetic effects are considered to be correlated random effects. Thus the model is consist of various types of effect and it might be useful for further applications in animal breeding field to set up a general methodology for this model because the methodology can be easily modified to other threshold animal models.

2. Model

Let $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k$ be $n \times 1$ vectors of continuous or ordered categorical observations. A typical member of $\mathbf{y}_i$ will be denoted by Y_{ij}. We assume that the first $m(\geq 1)$ traits, $\mathbf{y}_1, \ldots, \mathbf{y}_m$, are multiple ordered categorical observations and are the expression of underlying continuous random vectors $\mathbf{U}_1, \ldots, \mathbf{U}_m$, respectively with unknown threshold values $\mathbf{t}_1, \ldots, \mathbf{t}_m$ where $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{iC-1})'$ satisfying $-\infty = t_{i0} < t_{i1} < t_{i2} < \cdots < t_{iC-1} < t_{iC} = \infty$ for $i = 1, \ldots, m$. That is, if we denote the j-th element of $\mathbf{U}_i$ as U_{ij}, then Y_{ij} is the value recorded by

$$Y_{ij} = \ell, \quad \text{if } t_{i\ell-1} < U_{ij} \leq t_{i\ell} \tag{2.1}$$

for $i = 1, \ldots, m$ and $\ell = 1, 2, \ldots, C$. The remaining $k - m$ traits, $\mathbf{y}_{m+1}, \ldots, \mathbf{y}_k$, are assumed to be continuous observations. In what follows, we will use $(\mathbf{y}_{m+1}, \ldots, \mathbf{y}_k)$ and $(\mathbf{U}_{m+1}, \ldots, \mathbf{U}_k)$ interchangeably for notational convenience.

In analysis of calving ease score, some of Y_{ij}'s can be missing. For instance, CEH and CEC can not be observed simultaneously because of the nature of the traits. If CEH is observable, then CEC must be a missing value, and vice versa. Thus we should allow some missing values in the observations. The observed random vectors will be denoted by $\mathbf{y}_i^o, i = 1, 2, \ldots, k$.

Consider the following linear model for calving ease data :

$$\mathbf{U}_i = \mathbf{X}_i \boldsymbol{\beta}_i + \mathbf{Z}_h \mathbf{h}_i + \mathbf{Z}_d \mathbf{v}_{d_i} + \mathbf{Z}_m \mathbf{v}_{m_i} + \boldsymbol{\epsilon}_i, \quad \text{for } i = 1, \ldots, k, \tag{2.2}$$

where $\boldsymbol{\beta}_i$ is the fixed effect associated with intercept, sex and age of dam, $\mathbf{h}_i$ is the random herd-year-season effect, $\mathbf{v}_{d_i}$ is the direct additive genetic effect, and

$\mathbf{v}_{m_i}$ is the maternal genetic effect. It is assumed that $\mathbf{X}_i, \mathbf{Z}_h, \mathbf{Z}_d$ and $\mathbf{Z}_m$ are known matrices with dimension n by d_i, d_h, d_g and $d_m(= d_g)$, respectively, and $\mathbf{X}_i$ has full column rank. Let $\mathbf{W}_i = (\mathbf{X}_i, \mathbf{Z}_h, \mathbf{Z}_d, \mathbf{Z}_m)$, $\boldsymbol{\theta}_i = (\boldsymbol{\beta}_i', \mathbf{h}_i', \mathbf{v}_{d_i}', \mathbf{v}_{m_i}')'$ for $i = 1, 2, \ldots, k$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \ldots, \boldsymbol{\theta}_k')'$. Then a hierarchical Bayes model for (2.2) might be set up as :

(I) Conditional on $\boldsymbol{\theta}$ and $\mathbf{R}$

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 \\ \vdots \\ \mathbf{U}_k \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{W}_1 \boldsymbol{\theta}_1 \\ \vdots \\ \mathbf{W}_k \boldsymbol{\theta}_k \end{pmatrix}, \ \mathbf{R} \otimes \mathbf{I}_n \right),$$

where $\mathbf{R}$ is the $k \times k$ variance-covariance matrix of $(U_{1j}, \ldots, U_{kj}), j = 1, 2, \ldots, n$.

(II) $\boldsymbol{\theta}$ and $\mathbf{R}$ have a certain joint prior distribution proper or improper.

In what follow we will use a minus subscript to delete an appropriate elements from a matrix or a vector. For example, $\mathbf{U}_{-i}$ will denote the vector $\mathbf{U}$ with $\mathbf{U}_i$ deleted. Similarly and $\mathbf{U}_{-ij}$ denotes $\mathbf{U}$ with U_{ij} deleted. We also denote $\mathbf{B}_{i \cdot 21}$ and $\mathbf{B}_{i \cdot 22}$ as the lower-left $(p-1) \times 1$ and lower-right $(p-1) \times (p-1)$ of a partitioned matrix $\mathbf{L}_i \mathbf{B} \mathbf{L}_i$ where $\mathbf{B}$ is a $p \times p$ symmetric matrix and $\mathbf{L}_i$ is a $p \times p$ identity matrix with the i-th column being moved to the first column. For example,

$$\mathbf{L}_2 \mathbf{R} \mathbf{L}_2 = \left(\begin{array}{c|ccccc} r_{22} & r_{21} & r_{23} & r_{24} & \cdots & r_{2k} \\ \hline r_{12} & r_{11} & r_{13} & r_{14} & \cdots & r_{1k} \\ r_{32} & r_{31} & r_{33} & r_{34} & \cdots & r_{3k} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{k2} & r_{k1} & r_{k3} & r_{k4} & \cdots & r_{kk} \end{array} \right) = \begin{pmatrix} r_{22} & \mathbf{R}_{2 \cdot 21}' \\ \mathbf{R}_{2 \cdot 21} & \mathbf{R}_{2 \cdot 22} \end{pmatrix}.$$

Then under model (I), it is easy to check that, given $\mathbf{U}_{-ij}, \boldsymbol{\theta}$ and $\mathbf{R}$, the conditional density of U_{ij} is given by :

$$f(u_{ij} | \mathbf{U}_{-ij}, \boldsymbol{\theta}, \mathbf{R}) = \phi \left(\frac{u_{ij} - \eta_{ij}}{\sigma_{i \cdot}} \right), \tag{2.3}$$

where $\phi(\cdot)$ is the density function of a standard normal distribution, and

$$\eta_{ij} = \mathbf{w}_{ij}' \boldsymbol{\theta}_i + \mathbf{R}_{i \cdot 21}' \mathbf{R}_{i \cdot 22}^{-1} \begin{pmatrix} U_{1j} - \mathbf{w}_{1j}' \boldsymbol{\theta}_1 \\ U_{2j} - \mathbf{w}_{2j}' \boldsymbol{\theta}_2 \\ \vdots \\ U_{kj} - \mathbf{w}_{kj}' \boldsymbol{\theta}_k \end{pmatrix}_{-i},$$

$$\sigma_{i \cdot}^2 = r_{ii} - \mathbf{R}_{i \cdot 21}' \mathbf{R}_{i \cdot 22}^{-1} \mathbf{R}_{i \cdot 21},$$

with $\mathbf{w}'_{ij}$ denoting the j-th row of $\mathbf{W}_i$.

Consider the conditional distribution of U_{ij} given $\boldsymbol{\theta}, \mathbf{R}, \mathbf{t}, \mathbf{U}_{-ij}$ and $\mathbf{y}^o$ where $\mathbf{y}^o = (\mathbf{y}_1^{o\prime}, \ldots, \mathbf{y}_k^{o\prime})'$. Suppose Y_{ij} is missing. It is clear that given $\mathbf{U}_{-ij}$, U_{ij} is independent of $\mathbf{y}^o$ and the density is given by (2.3). Suppose now $Y_{ij} \in \mathbf{y}^o$. If $\mathbf{y}_i$ is an order categorical trait, then

$$\Pr\left[Y_{ij} = \ell \mid \mathbf{U}_{-ij}, \boldsymbol{\theta}, \mathbf{R}, \mathbf{t}_i\right] = \Pr\left[t_{i\ell-1} < U_{ij} \le t_{i\ell} \mid \mathbf{U}_{-ij}, \boldsymbol{\theta}, \mathbf{R}, \mathbf{t}_i\right]$$

$$= \Phi\left(\frac{t_{i\ell} - \eta_{ij}}{\sigma_{i\cdot}}\right) - \Phi\left(\frac{t_{i\ell-1} - \eta_{ij}}{\sigma_{i\cdot}}\right),$$

where $\Phi(\cdot)$ is the distribution function of a standard normal distribution. Thus the conditional distribution of U_{ij}, given $\boldsymbol{\theta}, \mathbf{R}, \mathbf{t}_i, \mathbf{U}_{-ij}$ and $\mathbf{y}^o$, is a truncated normal with density

$$f(u_{ij}|Y_{ij} = \ell, \mathbf{U}_{-ij}, \boldsymbol{\theta}, \mathbf{R}, \mathbf{t}) = \frac{\phi\left(\frac{u_{ij} - \eta_{ij}}{\sigma_{i\cdot}}\right)}{\Phi\left(\frac{t_{i\ell} - \eta_{ij}}{\sigma_{i\cdot}}\right) - \Phi\left(\frac{t_{i\ell-1} - \eta_{ij}}{\sigma_{i\cdot}}\right)} \mathbf{1}(u_{ij} \in (t_{i\ell-1}, t_{i\ell}]),$$

where $\mathbf{1}(\cdot)$ is an indicator function. On the other hand if $\mathbf{y}_i$ is a continuous trait, then $U_{ij}|\boldsymbol{\theta}, \mathbf{R}, \mathbf{t}, \mathbf{U}_{-ij}, \mathbf{y}^o$ have a degenerating distribution at Y_{ij} because $\mathbf{y}_i \equiv \mathbf{U}_i$.

Theorem 2.1. *Let ϕ and Φ be the density and the distribution function of a standard normal distribution. Then under model (I), the conditional density or conditional probability mass function of U_{ij} given $\mathbf{U}_{-ij}, \boldsymbol{\theta}, \mathbf{R}, \mathbf{t}, \mathbf{y}^o$ can be obtained according to cases, (1) $Y_{ij} \notin \mathbf{y}^o$, (2) $Y_{ij} \in \mathbf{y}^o$ and the i-th trait is continuous, and (3) $Y_{ij} \in \mathbf{y}^o$, i-th trait is ordered categorical and $Y_{ij} = \ell$, and is given by*

$$f(u_{ij}|\mathbf{U}_{-ij}, \boldsymbol{\theta}, \mathbf{R}, \mathbf{t}, \mathbf{y}^o) = \begin{cases} \phi\left(\frac{u_{ij} - \eta_{ij}}{\sigma_{i\cdot}}\right) & \text{case (1)} \\ \mathbf{1}(u_{ij} = Y_{ij}) & \text{case (2)} \\ \frac{\phi\left(\frac{u_{ij} - \eta_{ij}}{\sigma_{i\cdot}}\right)}{\Phi\left(\frac{t_{i\ell} - \eta_{ij}}{\sigma_{i\cdot}}\right) - \Phi\left(\frac{t_{i\ell-1} - \eta_{ij}}{\sigma_{i\cdot}}\right)} \mathbf{1}(u_{ij} \in (t_{i\ell-1}, t_{i\ell}]) & \text{case (3)}. \end{cases}$$

Theorem 2.1 provides the full posterior distribution of $\mathbf{U}$ to implement the Gibbs sampler. Next consider the prior distributions of parameters. As we mentioned before $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \ldots, \boldsymbol{\beta}'_k)'$ is a fixed effect while $\mathbf{h} = (\mathbf{h}'_1, \ldots, \mathbf{h}'_k)'$ and $\mathbf{v} = (\mathbf{v}'_{d_1}, \ldots, \mathbf{v}'_{d_k}, \mathbf{v}'_{m_1}, \ldots, \mathbf{v}'_{m_k})'$ are random. The traditional model assumptions can be achieved in Bayesian model by putting the prior distributions appropriately. We consider the following prior distributions :

1. $f(\boldsymbol{\beta}) \propto constant.$

2. $f(\mathbf{h}|\mathbf{Q}) \sim N(\mathbf{0}, \mathbf{Q} \otimes \mathbf{I})$ where $\mathbf{Q}$ is $k \times k$ variance-covariance matrix of $(h_{1j}, h_{2j}, \ldots, h_{kj})$.

3. $f(\mathbf{v}|\mathbf{G}) \sim N(\mathbf{0}, \mathbf{G} \otimes \mathbf{A})$ where $\mathbf{G}$ is $2k \times 2k$ variance-covariance matrix of $(v_{d_{1j}}, \ldots, v_{d_{kj}}, v_{m_{1j}}, \ldots, v_{m_{kj}})$ and $\mathbf{A}$ is a known matrix determined by genetic relationship.

Note that the prior distributions belong to a conjugate family or a noninformative prior. Because $\mathbf{Q}$ and $\mathbf{G}$ together with $\mathbf{R}$ are unknown, we rely on the hierarchical Bayes procedure and put noninformative Jeffrey priors for those variance-covariance matrices.

4. $f(\mathbf{R}) \propto |\mathbf{R}|^{-k/2}$.

5. $f(\mathbf{Q}) \propto |\mathbf{Q}|^{-k/2}$.

6. $f(\mathbf{G}) \propto |\mathbf{G}|^{-k}$.

Other noninformative priors for the variance-covariance might be possible, see, for example, Wang $et\ al.$ (1997).

For the prior distribution of $\mathbf{t}$, we refer to Sorensen $et\ al.$ (1995). Because the thresholds are ordered values, a reasonable noninformative prior for thresholds might be the one suggested by Sorensen $et\ al.$, which is the distribution of order statistics from uniform($t_{\min}, t_{\max}$). However, as many authors pointed out, there is an identifiability problem. Because the location parameter and the scale parameter of underlying variable are arbitrary, one must set the values of the parameters or two constrains should be specified. Usually one threshold and error variance of underlying variable are set to 0 and 1, respectively (Harville and Mee, 1984). However, other parameterization is also possible. For example, we might set two thresholds values. We will follow the latter because it allows easy specification of the prior distribution for $\mathbf{R}$ as given in 4. Although the two parameterizations do not yield the same posterior distributions, the equivalence of two parameterizations were illustrated by Sorensen $et\ al.$ (1995).

Setting $t_{11} = \cdots = t_{m1} = 0$ and $t_{12} = \cdots = t_{m2} = 1$, noninformative priors for the thresholds are the distributions of order statistics from uniform($1, t_{\max}$) and are given by

7. $f(\mathbf{t}_i) = (C_i - 3)! \left(\frac{1}{t_{\max}-1} \right)^{C_i-3} \mathbf{1}(\mathbf{t}_i \in \mathcal{T}_i)$ where $\mathcal{T}_i = \{(t_{i1}, t_{i2}, \ldots, t_{iC_i-1}) \mid t_{i1} = 0 < t_{i2} = 1 < t_{i3} < \cdots < t_{iC_i-1} \leq t_{\max}\}$, for $i = 1, \ldots, m$.

3. Gibbs Sampler

The Gibbs sampler consists of a set of fully conditional posterior distributions of unknown parameters in the model. Since the posterior distributions can be obtained from joint posterior distribution of parameters, we start from the joint posterior distribution.

Bayes theorem gives joint posterior distribution of parameters :

$$
\begin{aligned}
f(\boldsymbol{\theta}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}|\mathbf{y}^o) &\propto f(\mathbf{U}, \mathbf{y}^o|\boldsymbol{\theta}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}) f(\boldsymbol{\theta}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}) \\
&= f(\mathbf{y}^o|\mathbf{U}, \mathbf{t}) f(\mathbf{U}|\boldsymbol{\theta}, \mathbf{R}) f(\boldsymbol{\beta}) f(\mathbf{h}|\mathbf{Q}) f(\mathbf{Q}) \qquad (3.1) \\
&\quad \times f(\mathbf{v}|\mathbf{G}) f(\mathbf{G}) f(\mathbf{R}) f(\mathbf{t}).
\end{aligned}
$$

Among the 9 terms in (3.1), the second and the third terms are the only ones that are functions of $\boldsymbol{\beta}$. Thus the posterior distribution of $\boldsymbol{\beta}$ is proportional to the product of these two conditional distributions. Then

$$
f(\boldsymbol{\beta}_i|\boldsymbol{\theta}_{-\beta_i}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o) \propto f(\mathbf{U}|\boldsymbol{\theta}, \mathbf{R}) f(\boldsymbol{\beta}) \propto f(\mathbf{U}_i|\mathbf{U}_{-i}, \boldsymbol{\theta}, \mathbf{R}).
$$

Here $\boldsymbol{\theta}_{-\beta_i}$ represents the $\boldsymbol{\theta}$ with β_i deleted. Note that $\mathbf{U}_i|\mathbf{U}_{-i}, \boldsymbol{\theta}, \mathbf{R}$ has a multivariate normal distribution with mean $\mathbf{X}_i\beta_i + \mathcal{M}_{\beta_i}$ and variance $\sigma_i^2 \mathbf{I}_n$ where

$$
\mathcal{M}_{\beta_i} = \mathbf{Z}_h \mathbf{h}_i + \mathbf{Z}_d \mathbf{v}_{d_i} + \mathbf{Z}_m \mathbf{v}_{m_i} + \boldsymbol{\xi}_i,
$$

$$
\boldsymbol{\xi}_i = \left((\mathbf{R}'_{i\cdot 21} \mathbf{R}^{-1}_{i\cdot 22}) \otimes \mathbf{I}_n \right) \begin{pmatrix} \mathbf{U}_1 - \mathbf{W}_1\boldsymbol{\theta}_1 \\ \vdots \\ \mathbf{U}_k - \mathbf{W}_k\boldsymbol{\theta}_k \end{pmatrix}_{-i}
$$

for $i = 1, \ldots, k$. Hence we have

$$
\begin{aligned}
& f(\boldsymbol{\beta}_i|\boldsymbol{\theta}_{-\beta_i}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o) \\
&\propto \exp\left[-\frac{1}{2\sigma_i^2} (\mathbf{U}_i - \mathbf{X}_i\beta_i - \mathcal{M}_{\beta_i})' (\mathbf{U}_i - \mathbf{X}_i\beta_i - \mathcal{M}_{\beta_i})' \right] \\
&\propto \exp\left[-\frac{1}{2\sigma_i^2} \left\{ \beta_i' \mathbf{X}_i' \mathbf{X}_i \beta_i - 2\beta_i' \mathbf{X}_i' (\mathbf{U}_i - \mathcal{M}_{\beta_i}) \right\} \right] \\
&\propto \exp\left[-\frac{1}{2\sigma_i^2} \left\{ \beta_i - (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' (\mathbf{U}_i - \mathcal{M}_{\beta_i}) \right\}' \mathbf{X}_i' \mathbf{X}_i \right. \\
&\quad \left. \times \left\{ \beta_i - (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' (\mathbf{U}_i - \mathcal{M}_{\beta_i}) \right\} \right].
\end{aligned}
$$

This shows that, under model (I) and (II), the full posterior distribution of β_i is given by :

$$\beta_i | \theta_{-\beta_i}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o \sim N \left((\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' (\mathbf{U}_i - \mathcal{M}_{\beta_i}) , \sigma_i^2 (\mathbf{X}_i' \mathbf{X}_i)^{-1} \right).$$

As shown in Wang *et al.* (1994), the scalar form of the Gibbs sampler for the fixed effect also can be obtainable. Let β_{ij} be the j-th element of β_i and $\beta_{i,-j}$ be the β_i with the element β_{ij} deleted. $\mathbf{x}_{i,j}$ and $\mathbf{X}_{i,-j}$ represent the j-th column of $\mathbf{X}_i$ and $\mathbf{X}_i$ itself with the j-th column deleted. Then

$$f(\beta_{ij} | \theta_{-\beta_{ij}}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o)$$

$$\propto \exp \left[-\frac{1}{2\sigma_{i\cdot}^2} \left\{ \left(\sum_\ell \beta_{i\ell} \mathbf{x}_{i,\ell}' \right) \left(\sum_\ell \beta_{i\ell} \mathbf{x}_{i,\ell} \right) - 2 \left(\sum_\ell \beta_{i\ell} \mathbf{x}_{i,\ell}' \right) (\mathbf{U}_i - \mathcal{M}_{\beta_i}) \right\} \right]$$

$$\propto \exp \left[-\frac{1}{2\sigma_{i\cdot}^2} \left\{ \beta_{ij}^2 \mathbf{x}_{i,j}' \mathbf{x}_{i,j} - 2\beta_{ij} \mathbf{x}_{i,j}' (\mathbf{U}_i - \mathbf{X}_{i,-j}\beta_{i,-j} - \mathcal{M}_{\beta_i}) \right\} \right]$$

$$\propto \exp \left[-\frac{\mathbf{x}_{i,j}' \mathbf{x}_{i,j}}{2\sigma_{i\cdot}^2} \left\{ \beta_{ij} - (\mathbf{x}_{i,j}' \mathbf{x}_{i,j})^{-1} \mathbf{x}_{i,j}' (\mathbf{U}_i - \mathbf{X}_{i,-j}\beta_{i,-j} - \mathcal{M}_{\beta_i}) \right\}^2 \right],$$

and we have following result.

Lemma 3.1. *Under model (I) and (II), the full posterior distribution of β_{ij}, $i = 1, \ldots, k; j = 1, 2, \ldots, d$ is given by :*

$$\beta_{ij} | \theta_{-\beta_{ij}}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}$$

$$\sim N \left((\mathbf{x}_{i,j}' \mathbf{x}_{i,j})^{-1} \mathbf{x}_{i,j}' (\mathbf{U}_i - \mathbf{X}_{i,-j}\beta_{i,-j} - \mathcal{M}_{\beta_i}) , \sigma_i^2 (\mathbf{x}_{i,j}' \mathbf{x}_{i,j})^{-1} \right).$$

Next consider the posterior distribution of herd-year-season effects $\mathbf{h}$. As before the posterior distribution is proportional to the product of $f(\mathbf{h}|\mathbf{Q})$ and $f(\mathbf{U}|\theta, \mathbf{R})$ because only the two terms are the functions of $\mathbf{h}$. Hence

$$f(\mathbf{h}_i | \theta_{-h_i}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o) \propto f(\mathbf{h}|\mathbf{Q}) f(\mathbf{U}|\theta, \mathbf{R})$$

$$\propto f(\mathbf{h}_i | \mathbf{h}_{-i}, \mathbf{Q}) f(\mathbf{U}_i | \mathbf{U}_{-i}, \theta, \mathbf{R}).$$

Let $\mathcal{M}_{h_i} = \mathbf{X}_i \beta_i + \mathbf{Z}_d \mathbf{v}_{d_i} + \mathbf{Z}_m \mathbf{v}_{m_i} + \xi_i$, $\sigma_{h_i}^2 = q_{ii} - \mathbf{Q}_{i\cdot21}' \mathbf{Q}_{i\cdot22}^{-1} \mathbf{Q}_{i\cdot21}$ and $\lambda_{h_i} = \sigma_{h_i}^2 / \sigma_{i\cdot}^2$ for $i = 1, \ldots, k$. Since $\mathbf{h}_i | \mathbf{h}_{-i}, \mathbf{Q}$ is a multivariate normal random vector with mean $\mu_{h_i} = \left((\mathbf{Q}_{i\cdot21}' \mathbf{Q}_{i\cdot22}^{-1}) \otimes \mathbf{I} \right) \mathbf{h}_{-i}$ and variance $\sigma_{h_i}^2 \mathbf{I}$, and

$f(\mathbf{h}_i|\boldsymbol{\theta}_{-h_i}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o)$

$$\propto \exp\left[-\frac{1}{2\sigma_{h_i}^2}(\mathbf{h}_i - \boldsymbol{\mu}_{h_i})'(\mathbf{h}_i - \boldsymbol{\mu}_{h_i}) - \frac{1}{2\sigma_{i\cdot}^2}\left\{\mathbf{h}_i'\mathbf{Z}_h'\mathbf{Z}_h\mathbf{h}_i - 2\mathbf{h}_i'\mathbf{Z}_h'\left(\mathbf{U}_i - \mathcal{M}_{h_i}\right)\right\}\right]$$

$$\propto \exp\left[-\frac{1}{2\sigma_{h_i}^2}\left\{\mathbf{h}_i'\left(\mathbf{I} + \lambda_{h_i}\mathbf{Z}_h'\mathbf{Z}_h\right)\mathbf{h}_i - 2\mathbf{h}_i'\left(\boldsymbol{\mu}_{h_i} + \lambda_{h_i}\mathbf{Z}_h'\left(\mathbf{U}_i - \mathcal{M}_{h_i}\right)\right)\right\}\right]$$

$$\propto \exp\left[-\frac{1}{2\sigma_{h_i}^2}\left\{\mathbf{h}_i - \left(\mathbf{I} + \lambda_{h_i}\mathbf{Z}_h'\mathbf{Z}_h\right)^{-1}\left(\boldsymbol{\mu}_{h_i} + \lambda_{h_i}\mathbf{Z}_h'\left(\mathbf{U}_i - \mathcal{M}_{h_i}\right)\right)\right\}'\right.$$
$$\left. \times \left(\mathbf{I} + \lambda_{h_i}\mathbf{Z}_h'\mathbf{Z}_h\right)\left\{\mathbf{h}_i - \left(\mathbf{I} + \lambda_{h_i}\mathbf{Z}_h'\mathbf{Z}_h\right)^{-1}\left(\boldsymbol{\mu}_{h_i} + \lambda_{h_i}\mathbf{Z}_h'\left(\mathbf{U}_i - \mathcal{M}_{h_i}\right)\right)\right\}\right],$$

we see that the full posterior distribution of $\mathbf{h}_i$ is

$$\mathbf{h}_i|\boldsymbol{\theta}_{-h_i}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o \sim$$
$$N\left(\left(\mathbf{I} + \lambda_{h_i}\mathbf{Z}_h'\mathbf{Z}_h\right)^{-1}\left(\boldsymbol{\mu}_{h_i} + \lambda_{h_i}\mathbf{Z}_h'\left(\mathbf{U}_i - \mathcal{M}_{h_i}\right)\right), \sigma_{h_i}^2\left(\mathbf{I} + \lambda_{h_i}\mathbf{Z}_h'\mathbf{Z}_h\right)^{-1}\right).$$

Similarly the scalar form of the posterior distribution of h_{ij} is obtainable as follow:

Lemma 3.2. *Under model (I) and (II), the full posterior distribution of $h_{ij}, i = 1, 2, \ldots, k; j = 1, 2, \ldots, d_h$ is given by :*

$$h_{ij}|\boldsymbol{\theta}_{-h_{ij}}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o \sim$$
$$N\left((1 + \lambda_{h_i}\mathbf{z}_{h_j}'\mathbf{z}_{h_j})^{-1}\left(\mu_{h_{ij}} + \lambda_{h_i}\mathbf{z}_{h_j}'(\mathbf{U}_i - \mathbf{Z}_{h_{-j}}\mathbf{h}_{i,-j} - \mathcal{M}_{h_i})\right), \frac{\sigma_{h_i}^2}{1 + \lambda_{h_i}\mathbf{z}_{h_j}'\mathbf{z}_{h_j}}\right)$$

where $\mathbf{z}_{h_j}$ is the j-th column of $\mathbf{Z}_h$ and $\mathbf{Z}_{h_{-j}}$ represents $\mathbf{Z}_h$ with the j-th column deleted.

The posterior distributions of the direct and the maternal genetic effects can be obtained similarly. The posterior distribution of $\mathbf{v}$ is proportional to the product of $f(\mathbf{U}|\boldsymbol{\theta}, \mathbf{R})$ and $f(\mathbf{v}|\mathbf{G})$. Let $\mathbf{v}_{d_1} = \mathbf{v}_1, \mathbf{v}_{d_2} = \mathbf{v}_2, \ldots, \mathbf{v}_{m_k} = \mathbf{v}_{2k}$. For notational convenience, we use both type of subscripts for genetic effect.

Let $\mathbf{v} = (\mathbf{v}_1', \ldots, \mathbf{v}_{2k}')'$. The conditional distribution of $\mathbf{v}_i$ given $\mathbf{v}_{-i}$ and $\mathbf{G}$ is a multivariate normal with mean $\mathcal{V}_i = \left((\mathbf{G}_{i\cdot21}'\mathbf{G}_{i\cdot22}^{-1}) \otimes \mathbf{I}\right)\mathbf{v}_{-i}$ and variance covariance matrix $\sigma_{v_i}^2\mathbf{A}$ where $\sigma_{v_i}^2 = g_{ii} - \mathbf{G}_{i\cdot21}'\mathbf{G}_{i\cdot22}^{-1}\mathbf{G}_{i\cdot21}$ for $i = 1, 2, \ldots, 2k$. We also will use the both type of subscripts for $\mathcal{V}$ and σ_v^2. For example $\mathcal{V}_1$ and $\sigma_{v_1}^2$

imply $\mathcal{V}_{d_1}$ and $\sigma^2_{v_{d_1}}$, and so on. Let $\mathcal{M}_{v_{d_i}} = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{Z}_h\mathbf{h}_i + \mathbf{Z}_m\mathbf{v}_{m_i} + \boldsymbol{\xi}_i$, $\mathcal{M}_{v_{m_i}} = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{Z}_h\mathbf{h}_i + \mathbf{Z}_d\mathbf{v}_{d_i} + \boldsymbol{\xi}_i$ and $\lambda_{v_{c_i}} = \sigma^2_{v_{c_i}}/\sigma^2_i$ for $i = 1, 2, \ldots, k$ and $c = d, m$. Then we have

$$
f(\mathbf{v}_{c_i}|\boldsymbol{\theta}_{-v_{c_i}}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o)
$$

$$
\propto \exp\left[-\frac{1}{2\sigma^2_i}\{\mathbf{v}'_{c_i}\mathbf{Z}'_c\mathbf{Z}_c\mathbf{v}_{c_i} - 2\mathbf{v}'_{c_i}\mathbf{Z}'_c(\mathbf{U}_i - \mathcal{M}_{v_{c_i}})\} \right.
$$

$$
\left. - \frac{1}{2\sigma^2_{v_{c_i}}}\{\mathbf{v}'_{c_i}\mathbf{A}^{-1}\mathbf{v}_{c_i} - 2\mathbf{v}'_{c_i}\mathbf{A}^{-1}\mathcal{V}_{c_i}\} \right]
$$

$$
\propto \exp\left[-\frac{1}{2\sigma^2_{v_{c_i}}}\left\{ \mathbf{v}'_{c_i}(\lambda_{v_{c_i}}\mathbf{Z}'_c\mathbf{Z}_c + \mathbf{A}^{-1})\mathbf{v}_{c_i} \right.\right.
$$

$$
\left.\left. - 2\mathbf{v}'_{c_i}\left(\lambda_{v_{c_i}}\mathbf{Z}'_c(\mathbf{U}_i - \mathcal{M}_{c_i}) + \mathbf{A}^{-1}\mathcal{V}_{c_i}\right) \right\} \right]
$$

$$
\propto \exp\left[-\frac{1}{2\sigma^2_{v_{c_i}}}\left\{ \mathbf{v}_{c_i} - (\lambda_{v_{c_i}}\mathbf{Z}'_c\mathbf{Z}_c + \mathbf{A}^{-1})^{-1}(\lambda_{v_{c_i}}\mathbf{Z}'_c(\mathbf{U}_i - \mathcal{M}_{c_i}) + \mathbf{A}^{-1}\mathcal{V}_{c_i})\right\}' \right.
$$

$$
(\lambda_{v_{c_i}}\mathbf{Z}'_c\mathbf{Z}_c + \mathbf{A}^{-1})\left\{ \mathbf{v}_{c_i} - (\lambda_{v_{c_i}}\mathbf{Z}'_c\mathbf{Z}_c + \mathbf{A}^{-1})^{-1}(\lambda_{v_{c_i}}\mathbf{Z}'_c(\mathbf{U}_i - \mathcal{M}_{c_i}) \right.
$$

$$
\left.\left. + \mathbf{A}^{-1}\mathcal{V}_{c_i})\right\} \right].
$$

The posterior distribution of genetic effect is a multivariate normal with mean $(\lambda_{v_{c_i}}\mathbf{Z}'_c\mathbf{Z}_c + \mathbf{A}^{-1})^{-1}(\lambda_{v_{c_i}}\mathbf{Z}'_c(\mathbf{U}_i - \mathcal{M}_{c_i}) + \mathbf{A}^{-1}\mathcal{V}_{c_i})$ and variance-covariance matrix $\sigma^2_{v_{c_i}}(\lambda_{v_{c_i}}\mathbf{Z}'_c\mathbf{Z}_c + \mathbf{A}^{-1})^{-1}$. Since

$$
\sigma^2_{v_{c_i}}(\lambda_{v_{c_i}}\mathbf{Z}'_c\mathbf{Z}_c + \mathbf{A}^{-1})^{-1} = \mathbf{A} - \lambda_{v_{c_i}}\mathbf{A}\mathbf{Z}'_c(\mathbf{I} + \lambda_{v_{c_i}}\mathbf{Z}_c\mathbf{A}\mathbf{Z}'_c)^{-1}\mathbf{Z}_c\mathbf{A} = \boldsymbol{\Sigma}_c(\lambda_{v_{c_i}}),
$$

$$
(\lambda_{v_{c_i}}\mathbf{Z}'_c\mathbf{Z}_c + \mathbf{A}^{-1})^{-1}(\lambda_{v_{c_i}}\mathbf{Z}'_c(\mathbf{U}_i - \mathcal{M}_{c_i}) + \mathbf{A}^{-1}\mathcal{V}_{c_i})
$$

$$
= \mathcal{V}_{c_i} + \lambda_{v_{c_i}}\mathbf{A}\mathbf{Z}'_c(\mathbf{I} + \lambda_{v_{c_i}}\mathbf{Z}_c\mathbf{A}\mathbf{Z}'_c)^{-1}(\mathbf{U}_i - \mathcal{M}_{c_i} - \mathbf{Z}_c\mathcal{V}_{c_i}),
$$

we could alternatively write

$$
\mathbf{v}_{c_i}|\boldsymbol{\theta}_{-v_{c_i}}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o
$$

$$
\sim N\left(\mathcal{V}_{c_i} + \lambda_{v_{c_i}}\mathbf{A}\mathbf{Z}'_c(\mathbf{I} + \lambda_{c_i}\mathbf{Z}_c\mathbf{A}\mathbf{Z}'_c)^{-1}(\mathbf{U}_i - \mathcal{M}_{c_i} - \mathbf{Z}_c\mathcal{V}_{c_i}), \boldsymbol{\Sigma}_c(\lambda_{v_{c_i}})\right). \quad (3.2)
$$

In a calving ease model, the genetic effects usually has very many levels. This is one of the reason we should rely on the hierarchical Bayesian model. In fact it is common that d_g, the number of levels of both direct and maternal genetic

effects, are much larger than the sample size n. Therefore the dimension of $\mathbf{A}$ is usually much bigger than n and hence (3.2) is preferable for computation. See, however, Henderson (1976) for the computation of $\mathbf{A}^{-1}$. The scalar form for the posterior distribution of genetic effect is given in Lemma 3.3.

Lemma 3.3. *Let* $\mathbf{B} = \{b_{ij}\} = \mathbf{A}^{-1}$ *and* $\mathbf{b}'_{i.} = (b_{i1}, \ldots, b_{ig})$ *be the* i-th *row of* $\mathbf{B}$. *Under model (I) and (II), the full posterior distribution of* $v_{c_{ij}}$, $i = 1, 2, \ldots, k; c = d, m; j = 1, \ldots, d_g$ *is a normal with mean,*

$$
\mathrm{E}\left(v_{c_{ij}} | \boldsymbol{\theta}_{-v_{c_{ij}}}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o\right) = \left(b_{jj} + \lambda_{v_{c_i}} \mathbf{z}'_{c_j} \mathbf{z}_{c_j}\right)^{-1}
$$

$$
\times \left[b_{jj} \mathcal{V}_{c_{ij}} + \mathbf{b}_{i,-i}(\mathcal{V}_{c_{i-j}} - \mathbf{v}_{c_{i-j}}) + \lambda_{v_{c_i}} \mathbf{z}'_{c_j}(\mathbf{U}_i - \mathbf{Z}_{c_{-j}} \mathbf{v}_{c_i,-j} - \mathcal{M}_{v_{c_i}})\right],
$$

and variance,

$$
\mathrm{Var}\left(v_{c_{ij}} | \boldsymbol{\theta}_{-v_{c_{ij}}}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o\right) = \frac{\sigma^2_{v_{c_i}}}{b_{jj} + \lambda_{v_{c_i}} \mathbf{z}'_{c_j} \mathbf{z}_{c_j}},
$$

where $\mathbf{z}_{c_j}$ *is the* j-th *column of* $\mathbf{Z}_c$ *and* $\mathbf{Z}_{c_{-j}}$ *represents* $\mathbf{Z}_c$ *with the* j-th *column deleted.*

Next we will consider the posterior distributions of the variance-covariance matrices $\mathbf{R}, \mathbf{Q}$ and $\mathbf{G}$. Note that

$$
f(\mathbf{U}|\boldsymbol{\theta}, \mathbf{R})
$$

$$
\propto |\mathbf{R} \otimes \mathbf{I}_n|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left\{\mathbf{U} - \left(\bigoplus_{i=1}^{k} \mathbf{W}_i\right)\boldsymbol{\theta}\right\}' (\mathbf{R} \otimes \mathbf{I}_n)^{-1} \left\{\mathbf{U} - \left(\bigoplus_{i=1}^{k} \mathbf{W}_i\right)\boldsymbol{\theta}\right\}\right]
$$

$$
\propto |\mathbf{R}|^{-n/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\mathbf{R}^{-1}\mathbf{S}_U\right)\right],
$$

where $\bigoplus$ denotes the direct sum operator and $\mathbf{S}_U = \{S_{ij}\}_{k \times k}$ with $S_{ij} = (\mathbf{U}_i - \mathbf{W}_i\boldsymbol{\theta}_i)'(\mathbf{U}_j - \mathbf{W}_i\boldsymbol{\theta}_j)$. Similarly we have

$$
f(\mathbf{h}|\mathbf{Q}) \propto |\mathbf{Q} \otimes \mathbf{I}_{d_h}|^{-1/2} \exp\left[-\frac{1}{2}\mathbf{h}'(\mathbf{Q} \otimes \mathbf{I}_{d_h})^{-1}\mathbf{h}\right]
$$

$$
\propto |\mathbf{Q}|^{-d_h/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\mathbf{Q}^{-1}\mathbf{S}_h\right)\right],
$$

$$f(\mathbf{v}|\mathbf{G}) \propto |\mathbf{G} \otimes \mathbf{A}|^{-1/2} \exp\left[-\frac{1}{2}\mathbf{v}'(\mathbf{G} \otimes \mathbf{A})^{-1}\mathbf{v}\right]$$

$$\propto |\mathbf{G}|^{-d_g/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\mathbf{G}^{-1}\mathbf{S}_v\right)\right],$$

where $\mathbf{S}_h = \{\ \mathbf{h}_i'\mathbf{h}_j\}$ and $\mathbf{S}_v = \{\ \mathbf{v}_i'\mathbf{A}^{-1}\mathbf{v}_j\}$. Therefore

$$f(\mathbf{R}|\boldsymbol{\theta}, \mathbf{Q}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o) \propto f(\mathbf{U}|\boldsymbol{\theta}, \mathbf{R})f(\mathbf{R})$$

$$\propto |\mathbf{R}|^{-(n+k)/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\mathbf{R}^{-1}\mathbf{S}_U\right)\right],$$

$$f(\mathbf{Q}|\boldsymbol{\theta}, \mathbf{R}, \mathbf{G}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o) \propto f(\mathbf{h}|\mathbf{Q})f(\mathbf{Q})$$

$$\propto |\mathbf{Q}|^{-(d_h+k)/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\mathbf{Q}^{-1}\mathbf{S}_h\right)\right],$$

$$f(\mathbf{G}|\boldsymbol{\theta}, \mathbf{R}, \mathbf{Q}, \mathbf{t}, \mathbf{U}, \mathbf{y}^o) \propto f(\mathbf{v}|\mathbf{G})f(\mathbf{G})$$

$$\propto |\mathbf{G}|^{-(d_g+2k)/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\mathbf{G}^{-1}\mathbf{S}_v\right)\right].$$

These show the posterior distributions of $\mathbf{R}$, $\mathbf{Q}$ and $\mathbf{G}$ are Inv-Wishart$_{n-1}(\mathbf{S}_y^{-1})$, Inv-Wishart$_{d_h-1}(\mathbf{S}_h^{-1})$ and Inv-Wishart$_{d_g-1}(\mathbf{S}_u^{-1})$, respectively.

The noninformative priors for the thresholds $\mathbf{t}_i$'s are assumed to be independent and to be distributed as order statistics from a uniform distributions. However, it seems that some thresholds should not be independent for certain cases. For example, consider the 3 traits model with BWT, CEH and CEC. Because two traits are ordered categorical variable, we need two set of thresholds for CEH and CEC. However, it is believed that the model should have the common thresholds for both traits in the sense that both underlying variables measure the same phenomenon, calving ease, and the calving ease scores should be determined by the same criteria. In our preliminary study, if we do not restrict such condition to the thresholds, the estimates of two sets of thresholds tend to be quite different, which may be unexpected. We believe that this might happen mainly due to fact that the location and the scale parameters of underlying variables are adjusted by the thresholds separately. In other words, two calving ease scores are measured by different scales as well as standard (location). In this note we assumed that some ordered categorical traits may share a set of thresholds.

Let a thresholds $\mathbf{t}^* = (t_1^*, \ldots, t_C^*)'$ be accompanied to $\mathbf{y}_1, \ldots, \mathbf{y}_p$. Then we

have

$$f(t_\ell^* | \boldsymbol{\theta}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}_{-t_\ell^*}, \mathbf{U}, \mathbf{y}^o) = \frac{1}{\min_{\substack{1 \le i \le p \\ 1 \le j \le n_i}} (U_{ij} | Y_{ij} = \ell + 1) - \max_{\substack{1 \le i \le p \\ 1 \le j \le n_i}} (U_{ij} | Y_{ij} = \ell)},$$

where n_i is the dimension of $\mathbf{y}_i^o$. This is due to the fact that given $\mathbf{U}$, each $\mathbf{y}_i^o, i = 1, \ldots, p$ has a degenerating distribution,

$$f(\mathbf{y}_1^o, \ldots, \mathbf{y}_p^o | \mathbf{U}, \mathbf{t}) f(\mathbf{t}) = \prod_{i=1}^p \prod_{j=1}^{n_i} \left\{ \sum_{\ell=1}^C \mathbf{1}(t_{\ell-1}^* < U_{ij} \le t_\ell^*) \mathbf{1}(Y_{ij} = \ell) \right\} \mathbf{1}(\mathbf{t}^* \in \mathcal{T}^*),$$

where $\mathcal{T}^* = \{(t_1, t_2, \ldots, t_{C-1}) \mid t_1 = 0 < t_2 = 1 < t_3 < \cdots < t_{C-1} \le t_{\max}\}$, and

$$f(t_{i\ell} | \boldsymbol{\theta}, \mathbf{R}, \mathbf{Q}, \mathbf{G}, \mathbf{t}_{-i\ell}, \mathbf{U}, \mathbf{y}^o) \propto f(\mathbf{y}^o | \mathbf{U}, \mathbf{t}) f(\mathbf{t}).$$

It shows that given $\mathbf{U}$, the upper bound of the threshold $t_{i\ell}^*$ is $\min_{i,j}(U_{ij} | Y_{ij} = \ell + 1)$ and the lower bound is $\max_{i,j}(U_{ij} | Y_{ij} = \ell)$. Hence t_{ij}^* has a uniform distribution in this interval.

Remark 1. As we mentioned before, to prevent an identifiability problem, it is necessary to impose two restrictions on the parameters of underlying variable or on the thresholds. Many authors set one threshold and the variance of underlying variable to 0 and 1, respectively. Under this setting, the values of thresholds are limited in a finite interval. In our experiment, this setting has an advantage in speed of convergence compared with other restrictions. However, in a multiple traits model, the restrictions on underlying variables might cause a difficulty in specifying noninformative prior for $\mathbf{R}$. We could not find a literature to solve this problem.

Although, in an original threshold model, it is assumed that an observed categorical variable is determined by (2.1), we conducted an experiment under the assumption that a categorical variable Y_{ij} is recorded by

$$Y_{ij} = \ell, \quad \text{if } t_{i\ell-1} < U_{ij}/r_{ii} \le t_{i\ell} \tag{3.3}$$

with $t_{i1} = 0, t_{i2} = 1$ and some obvious modification on Theorem 2.1, because it is believed that the speed up is mainly due to the limited range of threshold value. The experiment is successful. Because the thresholds in (3.3) have one-to-one relationship with those in (2.1), we would like to recommend (3.3) for a multiple traits threshold model. However, there may be a difficulty in interpretation of

thresholds in a common threshold model. For instance, CEH and CEC have a common threshold under the assumption that the underlying variables have different location and scale parameters but the calving ease scores are determined by the same criterion. If we defined the calving ease score by (3.3), then the scores are affected by the scale. Thus it violates the assumptions. Thus we do not follow (3.3) for our numerical examples.

4. Numerical Examples

4.1. Data and model description

The data of birth weight (BWT) and calving ease (CE) from calves born from 1981 to 2000 is consisted of 322,956 from American Gelbvieh Association (AGA). Those data were selected by herd in which records were greater than 99 and by herd-year-season (HYS) in which records were greater than 1. HYS is defined to associate with percentage gelbvieh and working order group as well as herd, year and season like definition of contemporary group in AGA. We assume that calving ease of calves born by heifer (CEH) (assumed age of dam to 930 days) and calving ease of calves born by cow (CEC) (assumed age of dam > 930). The CE scores are 1, 2, 3 and 4 for no assistance, minor assistance, major assistance and caesarian, respectively. Calf's sex and age of dam (AOD) are grouped and assumed fixed effects. AOD for CEH are grouped with 4 categories (AOD < 675d, 675d to 750d, 750d to 825d, and > 825d) and AOD for CEC are grouped with 5 categories (AOD=3 yrs, 4 yrs, 5–6 yrs, 7–8 yrs, and > 9yrs). Three data sets were prepared with 57,643, 265,313 and 322,956 records on CEH, CEC and CEH+CEC with BWT for all data sets. Overall and subclass mean and proportion of each categories for CEH and CEC are shown in Table 1. Total number of sire and average number of progeny records per sires are 4,975 and 11.59 heads on CEH and 8,685 and 30.55 heads on CEC. 728 herds are retained for this study (Table 1).

TABLE 1   *Data information for birth weight and calving ease by heifer and cow in Gelbvieh*

	Data from Heifer				Data from Cow			
	No. Obs	Mean	SD	Min-Max	No. Obs	Mean	SD	Min-Max
no. progenys/sire	4975	12.41	52.02	1–1646	8685	30.67	72.76	1–2609
BWT	57643	80.16	10.87	35–140	265313	86.82	11.27	35–160
Frequency for calving ease								
Calving ease	1	2	3	4	1	2	3	4
Overall %	75.93	17.54	5.05	1.47	96.75	2.21	0.56	0.48
By sex								
Female	82.88	13.21	3.10	0.80	98.14	1.28	0.23	0.34
Male	68.34	22.26	7.18	2.21	95.31	3.18	0.89	0.63
By Age of dam[†]								
550-674(3)	70.55	21.47	5.77	2.22	94.24	4.06	1.03	0.67
675-749(4)	73.97	18.88	5.53	1.62	96.75	2.23	0.59	0.43
750-824(5 6)	81.77	13.51	3.77	0.95	97.58	1.57	0.40	0.44
825-930(7 8)	86.52	10.18	2.80	0.50	97.77	1.44	0.36	0.43
(Over 8)					97.96	1.37	0.27	0.40
Proportion of HYS having category[‡]								
No. HYS	5107	2997	1282	563	10466	2792	967	869
Proportion	98.56	57.89	24.76	10.88	100.00	26.67	9.24	8.30

NOTE : [†] *Unit of age of dam on data from heifer(cow) is days(yrs).* [‡] *Total HYS on data from heifer(cow) is 5177(10467).*

Three models considered in this analysis are as follow :

$$y_{BWT} = X\beta_{BWT} + Z_h h_{BWT} + Z_d v_{d_{BWT}} + Z_m v_{m_{BWT}} + \epsilon_{BWT}$$
$$U_{CEH} = X\beta_{CEH} + Z_h h_{CEH} + Z_d v_{d_{CEH}} + Z_m v_{m_{CEH}} + \epsilon_{CEH}$$
(Model 1)

$$y_{BWT} = X\beta_{BWT} + Z_h h_{BWT} + Z_d v_{d_{BWT}} + Z_m v_{m_{BWT}} + Z_p p_{BWT} + \epsilon_{BWT}$$
$$U_{CEC} = X\beta_{CEC} + Z_h h_{CEC} + Z_d v_{d_{CEC}} + Z_m v_{m_{CEC}} + Z_p p_{CEC} + \epsilon_{CEC}$$
(Model 2)

$$y_{BWT} = X\beta_{BWT} + Z_h h_{BWT} + Z_d v_{d_{BWT}} + Z_m v_{m_{BWT}} + Z_p p_{BWT} + \epsilon_{BWT}$$
$$U_{CEH} = X\beta_{CEH} + Z_h h_{CEH} + Z_d v_{d_{CEH}} + Z_m v_{m_{CEH}} + \epsilon_{CEH}$$ (Model 3)
$$U_{CEC} = X\beta_{CEC} + Z_h h_{CEC} + Z_d v_{d_{CEC}} + Z_m v_{m_{CEC}} + Z_p p_{CEC} + \epsilon_{CEC}$$

In particular Model 3 was analyzed with and without the restriction of common threshold. Permanent environmental effects (PE) were included in BWT and CEC in Model 2 and Model 3. The covariances of PE between traits were also taken into account in these model.

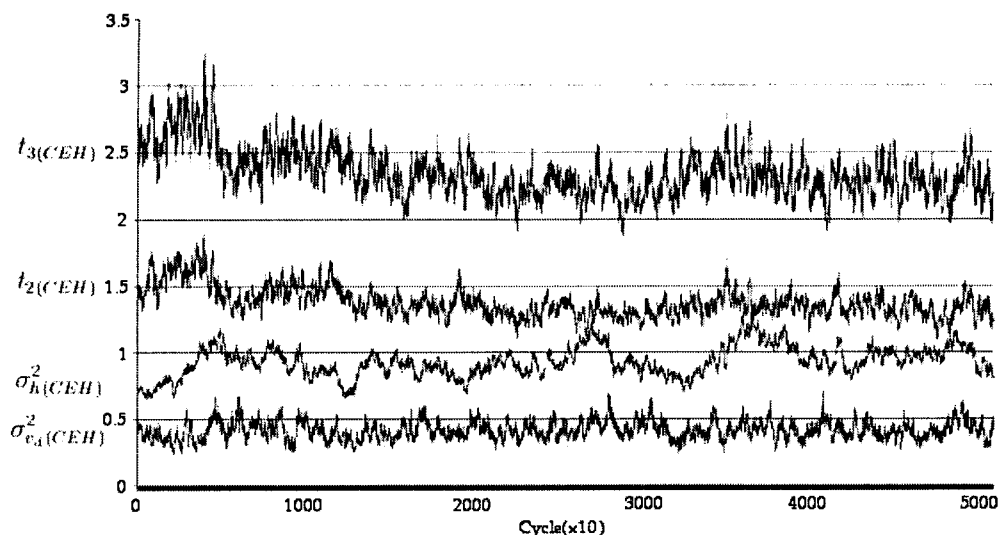Gibbs sampling analysis was carried out five times for each of the assumed

FIGURE 1 *Gibbs samples versus Gibbs cycle for thresholds(mathit$_{CEH}$), variance of herd effects ($\sigma^2_{h(CEH)}$) and variance of animal genetic effects($\sigma^2_{v_d(CEH)}$) on calving ease at first parity(CEH) in Model 3*

models. Each sample was consisted of 50,100 iterations, with the first 100 samples discarded and retained every 10-th estimates. To find out the convergent point and effective sample sizes, we followed the algorithm given by Raftery and Lewis (1992), which is known as a batching method. Figure 1 shows a typical convergence pattern. Posterior means and Monte Carlo errors for (co)variance components and threshold estimates from Gibbs sampler were calculated. These logics for threshold model with Gibbs sampling and Post–Gibbs analysis were implemented with FORTRAN90 under LINUX platform.

4.2. Results and discussion

It is well known that data associated with a particular fixed effect fall nearly all in the same category, known as the extreme category problem (ECP), can lead to biased estimates (Misztal *et al.*, 1989; Moreno *et al.*, 1997). In our study, total number of sire was 4,975 from heifers and average number of progeny per sire was 12.41 records. At the observations from cows, total number of sire was 8,685 and average number of progeny per sire was 30.67 (Table 1). Mean (SD) for birth weight was 80.16 (10.87)lb on 57,643 observations from heifers and 86.82 (11.27)lb on 265,313 observations from cows. The proportions of calving ease

were 75.93% on score1 (unassisted), 17.54% on score 2 (easy calving), 5.05% on score 3 (difficult calving), and 1.47% on score 4 (Caesarean) from heifers. Furthermore, those from cows were 96.75, 2.21, 0.56 and 0.48 on score 1, 2, 3 and 4, respectively. As shown Table 1, most births were unassisted especially for cows to calve. These proportions were in good agreement with several literatures (Varona *et al.*, 1999; Ramirez–Valverde, 2000).

The observations of CE on sex effects as fixed from heifers and cows were evenly distributed. So we do not need to worry about ECP on sex effects. However, there is a risk of ECP on age of dam effect which is assumed to be fixed effect. Also the distributions of CEH and CEC were skewed and might had ECP problems. 98.56% of HYS have the observations on CE score 1, otherwise, 10.88% of HYS only have the observations on CE score 4 from heifers. Likewise, all HYS have the observations on CE score 1 and 9.24% and 8.30% of HYS have the observations on CE score 3 and 4, respectively. This effect assumed to random to reduce bias estimates recommended by Moreno *et al.* (1997).

Table 2 and Table 3 show the estimated posterior means and SD of various variance components and correlations. Variance and covariance components of HYS effects for BWT and CEH were quickly stabilized and shown in good mixing rates in Model 3. Furthermore, direct and maternal genetic variances as well as residual variance for BWT were also shown to stabilize quickly and posterior means of these variances were close to estimates using by REML in single trait animal model.

As we mentioned before the estimates of threshold of CEH and CEC are quite different in Model 3. We believe that this might happen mainly due to fact that the thresholds adjust the location and the scale parameters of underlying variables separately. Thus if the estimation of variance components is a main issue for the analysis, this might cause a problem. However, the scaling for variance components did not affected to variance ratio.

In animal breeding, "heritability" and "genetic correlation" are defined to be the proportion of genetic variation on total variation and the correlation of genetic effects between traits or correlated effects. Here heritability implies the potential for genetic ability. The posterior means of heritabilities (empirical SD) of direct (h^2_{dw}) and maternal genetic effects (h^2_{mw}) for BWT were 0.348(.066) and 0.279 (.050) in Model 1 (Table 2). These estimates are larger than those of Varona *et al.* (1999). In Model 2, h^2_{dw} and h^2_{mw} were estimated to be 0.417(.058) and 0.076(.013). So the estimated value of h^2_{dw} in Model 2 is larger than that of Model 1. On the other hand, the estimates of h^2_{mw} are somewhat different from the case

TABLE 2 *Means and empirical SD of heritability and genetic correlation for birth weight (BWT), calving ease for heifers(CEH), and calving ease for cows(CEC) using marginal posterior means from Gibbs sampling in Model 1 and Model 2*

	Model 1		Model 2	
	Mean	SD	Mean	SD
Heritabilities				
h^2_{dw}	0.348	0.066	0.417	0.058
h^2_{dh}	0.246	0.002	-	-
h^2_{dc}	-	-	0.284	0.073
h^2_{mw}	0.279	0.050	0.076	0.013
h^2_{mh}	0.164	0.019	-	-
h^2_{mc}	-	-	0.146	0.072
Genetic Correlations				
$r_{g(dw-dh)}$	0.815	0.126	-	-
$r_{g(dw-dc)}$	-	-	0.756	0.189
$r_{g(dw-mw)}$	−0.178	0.218	−0.341	0.062
$r_{g(dh-mh)}$	0.033	0.008	-	-
$r_{g(dc-mc)}$	-	-	−0.533	0.221
$r_{g(mw-mh)}$	0.403	0.035	-	-
$r_{g(mw-mc)}$	-	-	0.278	0.291
Prop. PE				
c^2_w	-	-	0.050	0.012
c^2_c	-	-	0.139	0.052
Correlation for PE between traits				
$r_{pe(w-c)}$	-	-	0.276	0.159

of h^2_{dw}. In Model 3, these estimates of direct and maternal genetic effects were 0.456(.003) and 0.073(.003), respectively. These estimates are in good agreement with the results of other studies (Bennett and Gregory, 1996; Lee *et al.*, 2001). Other estimates can be referred from Table 2 and Table 3.

The variance components and genetic parameters for calving ease from heifers and cows and birth weight were successfully estimated in a multivariate threshold animal model. Calving ease from cows was assumed the different trait with calving ease from heifers. Genetic merit of calving ease for calves whatever was born from heifer or cow can be jointly evaluated with other traits correlated regardless of type of traits. This should be increased the precision of genetic merits. However, the research for the relationship of prior information for thresholds between different traits is needed to compare fixed effects between different traits and to implement the genetic evaluation. Heritability estimates for direct (maternal) genetic effects of calving ease from heifer and cow were 0.32 (0.18) and 0.28 (0.17),

TABLE 3 *Means and empirical SD of heritability and genetic correlation for birth weight (BWT), calving ease for heifers(CEH), and calving ease for cows(CEC) using marginal posterior means from Gibbs sampling in Model 3*

	Two Thresholds Model		Common Threshold Model	
	Mean	SD	Mean	SD
Heritabilities				
h^2_{dw}	0.448	0.003	0.456	0.003
h^2_{dh}	0.264	0.007	0.222	0.008
h^2_{dc}	0.277	0.021	0.274	0.012
h^2_{mw}	0.065	0.004	0.073	0.003
h^2_{mh}	0.134	0.006	0.141	0.009
h^2_{mc}	0.157	0.016	0.143	0.012
Genetic Correlations				
$r_{g(dw-dh)}$	0.762	0.035	0.874	0.013
$r_{g(dw-dc)}$	0.799	0.026	0.802	0.018
$r_{g(dw-mw)}$	−0.262	0.017	−0.283	0.008
$r_{g(dh-dc)}$	0.656	0.032	0.746	0.008
$r_{g(dh-mh)}$	0.042	0.045	0.047	0.017
$r_{g(dc-mc)}$	−0.365	0.068	−0.465	0.021
$r_{g(mw-mh)}$	0.659	0.032	0.686	0.023
$r_{g(mw-mc)}$	0.142	0.047	0.186	0.047
$r_{g(mh-mc)}$	0.186	0.090	0.264	0.061
Prop. PE				
c^2_w	0.060	0.001	0.057	0.001
c^2_c	0.118	0.007	0.086	0.005
Correlation for PE between traits				
$r_{pe(w-c)}$	0.186	0.027	0.088	0.055
Thresholds				
t_{13}	2.288	0.000	2.540	0.014
t_{23}	1.176	0.029	-	-

respectively, in three traits threshold animal model with two categorical traits. Genetic correlation estimate for direct genetic effects for calving ease between from heifer and from cow was 0.59 with marginal posterior mean.

As shown tables, conditional posterior distributions for variance-covariance components with conditioned common thresholds for between CEH and CEC traits were used different to different thresholds for between these traits. Nevertheless, variance ratios (heritabilities and genetic correlations) for categorical traits were similar with common thresholds and different thresholds. However, the scale for location parameters should be different with each other. If location parameters (breeding values) were interest, common thresholds were used should be recommended to estimate genetic merits.

REFERENCES

Bennett, G. L. and Gregory, K. E. (1996). "Genetic (co)variances among birth-weight, 200-day weight, and postweaning gain in composites and parental breeds of beef cattle", *Journal of Animal Science*, **74**, 2598-2611.

Berger, P. J., Lin, E. C., Van Arendonk, J. and Janss, L. (1995). "Properties of genetic parameter estimates from selection experiments by Gibbs sampling", *Journal of Dairy Science*, **78**, (Suppl. 1) 246.

Gelfand A. E., Hills, S. E., Racine-Poon, A. and Smith, A. F. M. (1990). "Illustration of Bayesian inference in normal data models using Gibbs sampling", *Journal of the American Statistical Association*, **85**, 972-985.

Geman, S. and Geman, D. (1984). "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721-741.

Foulley, J. L., Gianola, D. and Thompson, R. (1983). "Prediction of genetic merit from data on binary and quantitative variates with an application to calving difficulty, birth weight and pelvic opening", *Genetics Selection Evolution*, **15**, 407-424.

Foulley, J. L., Gianola, D. and Hoeschele, I. (1987). "Empirical Bayes estimation of parameters for n polygenic binary traits", *Genetics Selection Evolution*, **15**, 197-224.

Gianola, D. (1982). "Theory and analysis of threshold characters", *Journal of Animal Science*, **54**, 1079-1096.

Gianola, D. and Fouley, J. L. (1983). "Sire evaluation for ordered categorical data with a threshold model", *Genetics Selection Evolution*, **15**, 201-244.

Harville, D. A. and Mee, R. W. (1984). "A mixed-model procedure for analyzing ordered categorical data", *Biometrics*, **40**, 393-408.

Herderson, C. R. (1976). "A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values", *Biometrics*. **32**. 69-83.

Jensen, J., Wang, C. S., Sorensen, D. and Gianola, D. (1994). "Bayesian inference on variance and covariance components for traits influenced by maternal and direct genetic effects using the Gibbs sampler", *Acta Agriculturæ Scandinavica, Section A, Animal Science*, **44**, 193–201.

Lee, D. H. and Bertrand, J. K. (2001) "Genetic correlation estimates for weaning weight and postweaning gain for Hereford populations in different coultries". Submitted for publication.

Misztal, I., Gianola, D. and Fouley, J. L. (1989). "Computing aspects of nonlinear method of sire evaluation for categorical data", *Journal of Dairy Science*, **72**, 1557–1568.

Moreno, C., Sorensen, D., Garcia-Cortes, L. A., Varona, L. and Altarriba, J. (1997). "On biased inferences about variance components in the binary threshold model", *Genetics Selection Evolution*, **29**, 145–160.

Quass, R. L., Zhao, Y. and Pollak, E. J. (1988). "Describing interactions in dystocia scores with a threshold model", *Journal of Animal Science*, **66**, 396–399.

Raftery, A. E. and Lewis, S. M. (1992). "How many iterations in the Gibbs sampler", In *Bayesian Statistics IV*, Oxford University Press, UK, 763–773.

Ramirez-Valverde, R. (2000). "Studies to Implement Genetic Evaluation for Calving Difficulty in Beef Cattle", Ph.D. thesis, University of Georgia.

Sorensen, D., Wang, C. S., Jensen, J. and Gianola, D. (1994). "Bayesian analysis of genetic trend due to selection to beef cattle breeding", *Genetics Selection Evolution*, **25**, 3–30.

Sorensen, D. A., Andersen, S., Gianola, D. and Korsgaard, I. (1995). "Bayesian inference in threshold models using Gibbs sampling", *Genetics Selection Evolution*, **27**, 229–249.

Varona, L., Misztal, I. and Bertrand, J. K. (1999). "Threshold-linear versus linear–linear analysis of birth weight and calving ease using an animal model I. Variance components estimation", *Journal of Animal Science*, **77**, 1994–2002.

Wang, C. S., Quaas, R. L. and Pollak, E. J. (1997). "Bayesian analysis of calving ease scores and birth weights", *Genetics Selection Evolution*, **20**, 117–143.

Wang, C. S., Rutledge, J. J. and Gianola, D. (1994). "Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs", *Genetics Selection Evolution*, **26**, 91–115.

Wright, S. (1934). "An analysis of variability in number of digits in an inbred strain of guinea pigs", *Genetics*, **19**, 506–536.