Journal of the Korean Statistical Society (1999), **28**: *2, pp 211–223*

Mixed Linear Models with Censored Data†

Il-Do Ha[1], Youngjo Lee[2] and Jae-Kee Song[3]

ABSTRACT

We propose a simple estimation procedure in the mixed linear models with censored normal data, using both Buckly and James (1979) type pseudo random variables and Lee and Nelder's (1996) estimation procedure. The proposed method is illustrated with the matched pairs data in Pettitt (1986).

Keywords: Censored data; Dispersion components; Hierarchical generalized linear models; Hierarchical likehood; Mixed linear models; Random effects.

1. INTRODUCTION

When there are repeated measurements of survival on each individual (subject), the survival times may depend on observed covariates or unobserved random covariates called random effects (between-subjects effects). If the random effects are ignored, the inferences from linear models (LMs) (or Cox's (1972) proportional hazards models) with only observed covariates effects (fixed effects) may be neither reliable nor sufficient. Thus several authors (e.g. Pettitt, 1986; Carriquiry et al., 1987) have considered the mixed linear models (MLMs) under a normal distribution, where random effects act linearly on the survival times.

Pettitt (1986) obtained an estimation procedure for marginal maximum likelihood, using the EM algorithm introduced by Dempster et al. (1977). Also, Carriquiry et al. (1987) presented a Bayesian estimation procedure, using flat prior distributions for all parameters except the random effects and the Newton-Raphson algorithm. However, these methods become computationally expensive or intractable due to requiring multiple integrations: see Smith and Helms (1995, pp.426).

---

†This paper was partially supported by the Basic Science Research Institute Programs, Ministry of Education, 1998, Project No. BSRI-98-1403.

[1]Department of Statistics, Kyungsan University, Kyungpook, 712-240, Korea.

[2]Department of Statistics, Seoul National University, Seoul, 151-742, Korea.

[3]Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.

The purpose of this paper is to consider the MLMs with censored data and to show an expectation identity based on Buckly and James (1979) type pseudo random variables (rv's), as Buckly and James (1979) did in LMs, and then to propose a simple estimation procedure in the MLMs under a normal distribution, using both these pseudo rv's and Lee and Nelder's (1996) estimation procedure. The proposed method provides validity for the estimation and hence faciliates to obtain the estimation procedure because of no requiring integration.

In Section 2, we define some notations and describe the models considered. The expectation identity and the estimation procedure are shown and proposed in Section 3, respectively. Finally in Section 4, we illustrate our results with the matched pairs data in Pettitt(1986), an example of a repeated measures design producing censored data.

2. NOTATION AND MODEL STRUCTURES

Let T_{ij} $(i = 1, \cdots, q\ ;\ j = 1, \cdots, n_i\ ;\ n = \sum_i n_i)$ be the survival time on study of the jth repetition on the ith individual and C_{ij} be the censoring time associated with T_{ij}. However T_{ij}'s may not all be observable due to the censoring mechanism, i.e. the observable quantities are

$$Y_{ij} = \min(T_{ij}, C_{ij}), \quad \delta_{ij} = I(T_{ij} \leq C_{ij}), \quad x_{ij}^t = (x_{ij1}, \cdots, x_{ijp}),$$

where $I(\cdot)$ is the indicator function and x_{ij}^t is the $1 \times p$ vector of covariates associated with T_{ij}.

Let u_i be the unobserved random effect for ith individual. In this paper we assume that

$$T_{ij} \text{ and } C_{ij} \text{ are conditionally independent given } u_i.$$

We here consider the MLM which is described as the following (two-stage) hierachical model: for $i = 1, \cdots, q\ ;\ j = 1, \cdots, n_i$

(i) $T_{ij}|u_i \sim$ independent $N(\mu_{ij}', \sigma_\epsilon^2)$ with $\mu_{ij}' = E(T_{ij}|u_i) = x_{ij}^t\beta + u_i$,

(ii) $u_i \sim$ i.i.d. $N(0,\ \sigma_u^2)$,

where β is the $p \times 1$ vector of fixed effects.

In the stage (i), T_{ij} may be a suitable transformation of the survival time on study, e.g. $\log(T_{ij})$. Hereafter we call the above model Normal-MLM. Recently Lee and Nelder (1996) introduced hierarchical generalized linear models

(HGLMs) with arbitrary distribution for random effects. In fact, the Normal-MLM is exactly the same as the Normal-Normal HGLM since both $T_{ij}|u_i$ and u_i follow the Normal distribution.

Note that the Normal-MLM takes into account intra-subject correlation of the outcomes, i.e. the correlation between two outcomes T_{ij} and $T_{ij'}$ from the same subject is given by

$$\rho_{jj'} = \rho = \sigma_u^2/(\sigma_\epsilon^2 + \sigma_u^2), \quad j \neq j'. \tag{2.1}$$

Here, σ_ϵ^2 and σ_u^2 are called dispersion components (or variance components) and stand for variability within and between subjects, respectively.

In matrix notation the Normal-MLM is of the form

(i) $T|u \sim N(\mu' \ , \ \sigma_\epsilon^2 I_n)$ with $\mu' = E(T|u) = X\beta + Zu$,

(ii) $u \sim N(0 \ , \ \sigma_u^2 I_q)$,

where T is the $n \times 1$ vector with the ijth element T_{ij}, u is the $q \times 1$ vector with the ith element u_i, μ' is the $n \times 1$ vector with the ijth element μ'_{ij}, X is the $n \times p$ model matrix whose ijth row vector is x_{ij}^t, Z is the $n \times q$ indicator model matrix whose the (ij,k)th element is $\partial\mu'_{ij}/\partial u_k$, and $I_n[I_q]$ is the $n \times n[q \times q]$ identity matrix.

3. ESTIMATION PROCEDURE

3.1. Buckley and James (1979) Type Pseudo rv's

Consider $n_i = 1$ for all i in notation of Section 2. Then under the LM, $E(T_i) = x_i^t\beta$ with unspecified distribution of T_i, Buckley and James (1979) proved the expectation identity $E(Y_i^*) = x_i^t\beta$ with $Y_i^* = Y_i\delta_i + E(T_i|T_i > Y_i)(1-\delta_i)$. This was proposed to overcome the inconsistency problem in Miller's (1976) approach. Also, Wolynetz (1979), Schmee and Hahn (1979), and Aitkin (1981) used the pseudo rv's Y_i^*'s to obtain an estimation procedure in the LM with censored normal data. Consequently, these estimation methods using Y_i^*'s provide not only validity for the estimation but also a simple estimation procedure.

Under the Normal-MLM, T_{ij} may be subject to censoring, so only Y_{ij}'s are observed, but

$$E(Y_{ij}|u_i) \neq x_{ij}^t\beta + u_i. \tag{3.1}$$

We now consider the following another pseudo rv (we call it Buckley and James type pseudo rv)

$$Y_{ij}^* = Y_{ij}\delta_{ij} + E(T_{ij}|T_{ij} > Y_{ij}, u_i)(1 - \delta_{ij}). \tag{3.2}$$

Then by the conditional independence of T_{ij} and C_{ij}, we have the following expectation identity

$$E(Y_{ij}^*|u_i) = x_{ij}^t\beta + u_i. \tag{3.3}$$

In fact, the equations (3.1),(3.2) and (3.3) hold without specifyng the associated distributions. For proof of (3.3) see Appendix A. It can be easily shown that if the distribution of $T_{ij}|u_i$ is normal, then (3.2) becomes

$$Y_{ij}^* = Y_{ij}\delta_{ij} + (\mu_{ij}' + \sigma_\epsilon V(M_{ij}'))(1 - \delta_{ij}), \tag{3.4}$$

where $V(\cdot)$ is the hazard function for the standard normal distribution, i.e. $V(\cdot) = \phi(\cdot)/\bar{\Phi}(\cdot)$, ϕ and $\Phi(= 1 - \bar{\Phi})$ are the standard normal density and the distribution function, respectively, and $M_{ij}' = (Y_{ij} - \mu_{ij}')/\sigma_\epsilon$.

After all, owing to the expectation identity (3.3), we use the Y_{ij}^*'s in order to estimate the parameters in the Normal-MLM. Actually, since we can't observe all of the Y_{ij}^*'s, we use estimates of the Y_{ij}^*'s as follows:

$$\widehat{Y_{ij}^*} = Y_{ij}\delta_{ij} + (\widehat{\mu_{ij}'} + \widehat{\sigma}_\epsilon V(\widehat{M_{ij}'}))(1 - \delta_{ij}), \tag{3.5}$$

where $\widehat{M_{ij}'} = (Y_{ij} - \widehat{\mu_{ij}'})/\widehat{\sigma}_\epsilon$ and $\widehat{\mu_{ij}'} = x_{ij}^t\widehat{\beta} + \widehat{u}_i$.

3.2. Hierarchical Likelihood

Lee and Nelder (1996) proposed the use of the hierarchical likehood (h-likelihood) for inferences from HGLMs. The h-likelihood is defined by the logarithm of the joint density function for the response and v, where $v = v(u)$ is the strictly monotone function of the random effect u. In fact, it is a generalization of Henderson's (1975) joint likelihood. Note that the Normal-MLM is the Normal-Normal HGLM with identity link, $v(u) = u$.

We now construct the h-likelihood for the Normal-MLM in order to obtain the corresponding estimation procedure. Let y_{ij} be the observed value of Y_{ij} and y be the $n \times 1$ vector with the ijth element y_{ij}. Then by the conditional independence of T_{ij} and C_{ij}, the h-likelihood, denoted by h, is given by

$$h = h(\beta, \sigma_\epsilon^2, \sigma_u^2; y, u) = \ell(\beta, \sigma_\epsilon^2; y|u) + \ell(\sigma_u^2; u), \tag{3.6}$$

where

$$\begin{aligned}\ell(\beta,\sigma_\epsilon^2;y|u) &= \log\{\prod_{ij\in D} p_1(y_{ij}|u_i)\prod_{ij\in C} S_1(y_{ij}|u_i)\} \\ &\propto -\frac{r}{2}\log\sigma_\epsilon^2 - \frac{1}{2\sigma_\epsilon^2}\sum_{ij\in D}(y_{ij}-\mu'_{ij})^2 \\ &+ \sum_{ij\in C}\log\bar{\Phi}((y_{ij}-\mu'_{ij})/\sigma_\epsilon) \end{aligned} \tag{3.7}$$

and

$$\ell(\sigma_u^2;u) = \log\{\prod_{i=1}^{q} p_2(u_i)\} \propto -\frac{q}{2}\log\sigma_u^2 - \frac{1}{2\sigma_u^2}\sum_{i=1}^{q} u_i^2. \tag{3.8}$$

Here p_1 and S_1 are the normal density and survival function of $T_{ij}|u_i$, respectively, p_2 is the normal density function of u_i, D[C] are the index set of individuals for uncensored[censored] observations, and r is the number of uncensored observations.

3.3. Estimation of Fixed and Random Effects

According to Lee and Nelder's (1996) h-likelihood approach, the maximum h-likelihood estimates (MHLEs) which are derived from the h-likelihood (3.6) can be obtained by solving

$$\partial h/\partial\beta_k = 0 \ (k=1,\cdots,p) \ \text{ and } \ \partial h/\partial u_i = 0 \ (i=1,\cdots,q). \tag{3.9}$$

For details, we have

$$\partial h/\partial\beta_k = \frac{1}{\sigma_\epsilon}\{\sum_{ij\in D} x_{ijk}m'_{ij} + \sum_{ij\in C} x_{ijk}V(m'_{ij})\} \ (k=1,\cdots,p), \tag{3.10}$$

$$\partial h/\partial u_i = \frac{1}{\sigma_\epsilon}\{\sum_{j\in D} m'_{ij} + \sum_{j\in C} V(m'_{ij})\} - \frac{1}{\sigma_u^2}u_i \ (i=1,\cdots,q), \tag{3.11}$$

where $m'_{ij} = (y_{ij}-\mu'_{ij})/\sigma_\epsilon$. By plugging (3.4) in equations (3.10) and (3.11), they are reduced to

$$\partial h/\partial\beta_k = \frac{1}{\sigma_\epsilon^2}\sum_{ij\in A}(y^*_{ij}-\mu'_{ij})x_{ijk} \ (k=1,\cdots,p), \tag{3.12}$$

$$\partial h/\partial u_i = \frac{1}{\sigma_\epsilon^2}\sum_{j=1}^{n_i}(y^*_{ij}-\mu'_{ij}) - \frac{1}{\sigma_u^2}u_i \ (i=1,\cdots,q), \tag{3.13}$$

where A is the set of full observations which contains both D and C, and y^*_{ij} is the observed value of Y^*_{ij}. In matrix notation, from (3.12) and (3.13) the equation (3.9) can be written as

$$\sigma_\epsilon^2 X^t(Y^* - X\beta - Zu) = 0, \tag{3.14}$$

$$\sigma_\epsilon^2 Z^t(Y^* - X\beta - Zu) - \frac{1}{\sigma_u^2}u = 0, \tag{3.15}$$

where Y^* is the $n \times 1$ vector with the ijth element y^*_{ij}.

Thus, if Y^* and $(\sigma_\epsilon^2, \sigma_u^2)$ are known, the MHLEs $(\hat\beta, \hat u)$ can be obtained by solving iteratively the following equations

$$\begin{pmatrix} X^tX & X^tZ \\ Z^tX & Z^tZ + \lambda I_q \end{pmatrix} \begin{pmatrix} \hat\beta \\ \hat u \end{pmatrix} = \begin{pmatrix} X^tY^* \\ Z^tY^* \end{pmatrix}, \tag{3.16}$$

where $\lambda = \sigma_\epsilon^2/\sigma_u^2$. In fact, in order to obtain the solutions of (3.16), we use $\hat Y^*$ instead of Y^*.

Note that if there is no censoring then the equations (3.16) reduce to Henderson's (1975) mixed-model or best linear unbiased predictor (BLUP) equations. This fact is due to both the expectation identity (3.3) and the equation (3.9). Moreover, the equations (3.16) are rewritten as

$$\begin{aligned} (X^t\Sigma^{-1}X)\hat\beta &= X^t\Sigma^{-1}Y^*, \\ \hat u &= \sigma_u^2 Z^t\Sigma^{-1}(Y^* - X\hat\beta), \end{aligned}$$

where $\Sigma = \Sigma(\sigma_\epsilon^2, \sigma_u^2) = \text{Var}(T) = ZZ^t\sigma_u^2 + I_n\sigma_\epsilon^2$.

Next, if X has full column rank, then the estimator of the covariance matrix of $\hat\beta$ and $\hat u - u$ is approximately (e.g. Carriquiry et al. (1987) and Lee and Nelder (1996))

$$H^{*-1}|_{\Theta=\hat\Theta} \cdot \sigma_\epsilon^2, \tag{3.17}$$

where

$$H^* = -\,\sigma_\epsilon^2\left(\frac{\partial^2 h}{\partial(\beta, u)^2}\right) = -\,\sigma_\epsilon^2\begin{pmatrix} A_1 & A_2 \\ A_2^t & A_3 \end{pmatrix}, \tag{3.18}$$

and A_1, A_2 and A_3 are matrices such that the ijth element of A_1 is $\partial^2 h/\partial\beta_i\partial\beta_j$, the jkth element of A_2 is $\partial^2 h/\partial\beta_j\partial u_k$, and the ijth element of A_3 is $\partial^2 h/\partial u_i\partial u_j$, and $\hat\Theta = (\hat\beta, \hat u)$ is the vector of solutions after convergence has occurred. In Appendix B we show that (3.18) becomes

$$H^* = \begin{pmatrix} X^tWX & X^tWZ \\ Z^tWX & Z^tWZ + \lambda I_q \end{pmatrix}, \tag{3.19}$$

where

$$W = \begin{pmatrix} I_D & 0_D \\ 0_C & G_C \end{pmatrix},$$

I_D is the identity matrix whose ijth element belongs to C, $0_D[0_C]$ is the zero matrix whose ijth element belongs to $D[C]$, and G_C is the diagonal matrix with the ijth element $\tau(m'_{ij}) = V(m'_{ij})[V(m'_{ij}) - m'_{ij}]$ for $ij \in C$. Note that W in (3.19) is the weight matrix which takes into account loss of information due to censoring and in case of no censoring H^* reduces to the square matrix on the left hand side of (3.16). In addition, we can see that Y^* and W depend on censoring patterns.

3.4. Estimation of Dispersion Components

In estimating dispersion components, the (marginal) maximum likelihood estimator may be biased due to failing to take into account the loss in degree of freedom resulting from estimating fixed effects.

Thus for estimation of dispersion components (ϕ, α), Lee and Nelder(1996) considered an adjusted h-likelihood

$$h_A = h + \frac{1}{2}\log\{\det(2\pi\phi H^{-1})\}, \tag{3.20}$$

where $H = -\ \phi E(\frac{\partial^2 h}{\partial(\beta,v)^2}|v)$. The maximum adjusted profile h-likelihood estimators (MAPHLEs) for dispersion parameters are obtained by solving the followings iteratively

$$\partial h_A/\partial\phi|_{\beta=\hat\beta, v=\hat v} = 0 \ \text{ and } \ \partial h_A/\partial\alpha|_{\beta=\hat\beta, v=\hat v} = 0, \tag{3.21}$$

where $\hat\beta$ and $\hat v$ are re-evaluated in each iteration. Note that in Normal-Normal HGLM the MAPHLEs become the restricted maximum likelihood estimators (REMLEs) proposed by Patterson and Thompson (1971).

Now we want to estimate the dispersion components $(\sigma_\epsilon^2, \sigma_u^2)$ in the Normal-MLM. Here $(\sigma_\epsilon^2, \sigma_u^2)$ corresponds to (ϕ, α). Since it is difficult to calculate H in (3.20) due to censoring mechanism, we use H^* instead of H. In Appendix C we show that the MAPHLEs for $(\sigma_\epsilon^2, \sigma_u^2)$ given Y^* are

$$\widehat{\sigma_\epsilon^2} = (Y^* - \widehat{\mu'})^t(Y^* - \widehat{\mu'})/\{n_1 - (p + q - \gamma_1)\}, \tag{3.22}$$

$$\widehat{\sigma_u^2} = \widehat{u}^t\widehat{u}/(q - \gamma_2), \tag{3.23}$$

where n_1, γ_1, and γ_2 are described in the Appendix C.

Note that in case of no censoring (3.22) and (3.23) reduce to Schall's (1991, pp.721) REMLEs for $(\sigma_\epsilon^2, \sigma_u^2)$.

3.5. Algorithm

We provide an efficient algorithm for the estimation procedure of the Normal-MLM as follows:

Step 0 : Obtain the initial estimates $(\hat{\beta}, \hat{u})$ and $(\widehat{\sigma_\epsilon^2}, \widehat{\sigma_u^2})$ of (β, u) and $(\sigma_\epsilon^2, \sigma_u^2)$, respectively.

Step 1 : Calculate the followings for $ij \in C$

$$\begin{aligned}\widehat{m'_{ij}} &= (y_{ij} - \widehat{\mu'_{ij}})/\widehat{\sigma_\epsilon},\\ y^*_{ij} &= \widehat{\mu'_{ij}} + \widehat{\sigma_\epsilon} V(\widehat{m'_{ij}}),\\ \tau(m'_{ij}) &= [V(\widehat{m'_{ij}})]^2 - \widehat{m'_{ij}} V(\widehat{m'_{ij}}).\end{aligned}$$

Step 2 : Calculate the new estimates $(\hat{\beta}, \hat{u})$ and $(\widehat{\sigma_\epsilon^2}, \widehat{\sigma_u^2})$ using (3.16), (3.22) and (3.23).

Step 3 : Repeat Steps 1 and 2 until the necessary convergence criterion is satisfied.

Note that the initial estimates in Step 0 can be usually obtained by regarding censored observations as uncensored ones. In details, the first iterative solutions of (3.16) with arbitrary starting value $\lambda(> 0)$ and replacing Y^* by Y are used as initial estimates of (β, u), and then the first iteration for the Schall's REML estimates as initial estimates of $(\sigma_\epsilon^2, \sigma_u^2)$. Further as a convergence criterion in Step 3 we use the maximum absolute difference (upper bound 10^{-5}) of the previous and current estimates for (β, u).

4. EXAMPLE

We consider the matched pairs data in Pettitt (1986), adapted from the original data of Batchelor and Hackett (1970). The data shown in Table 4.1 consist of the survival times (days) of HLA (human lymphocyte antigen) closely and poorly matched skin grafts on the same burn patient. Pettitt (1986) analysed the data via the EM algorithm for the Normal-MLM with $p = 2$, where T_{ij} $(i = 1, \cdots, 11; j = 1, 2)$ is the logarithm of the jth survival time on the ith patient, $\mu'_{ij} = \beta_0 + \beta_1 x_{ij} + u_i$, x_{ij} is a covariate being -1 for a poor match and 1 for a close match, and u_i is the ith patient's random effect.

Table 4.1: Days of survival of skin grafts on burn patients.

Patient	1	2	3	4	5	6	7	8	9	10	11
Close	37	19	57^+	93	16	22	20	18	63	29	60^+
Poor	29	13	15	26	11	17	26	21	43	15	40

+: denotes right censored observation.

Table 4.2: Estimates of fixed effects and dispersion components. (S.E. in parentheses)

Method	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}^2_\epsilon$	$\hat{\sigma}^2_u$	$t = \hat{\beta}_1/SE(\hat{\beta}_1)$	$\hat{\rho}$
Pettitt	3.305	0.253	0.148	0.167	3.09	0.53
	(0.150)	(0.082)				
Proposed	3.300	0.246	0.155	0.178	2.89	0.53
	(0.153)	(0.085)				

Here we want to fit the above model via the proposed method. For the fitting we use SAS/IML and the estimation algorithm begins with arbitrarily selected starting value $\lambda = 1$. As results it gives the convergent solutions after the 12th iteration.

The obtained results are compared with those of Pettitt: see Table 4.2. Table 4.2 indicates that two methods lead to very similar parameter estimates and S.E.s. Moreover, the approximate t-value to test $\beta_1 = 0$ is 2.89 and the estimated correlation $\hat{\rho}$ $(= \hat{\sigma}^2_u/(\hat{\sigma}^2_\epsilon + \hat{\sigma}^2_u))$ based on (2.1) is 0.53, and Pettitt's method gives 3.09 and 0.53. In other words, two methods lead to similar significance for the effect of β_1, which the closely matched grafts have significanly longer survival than the poorly matched grafts and also to almost the same positive value for the correlation. Therefore our method gives very reasonable agreement with that of Pettitt, but is conceptually simple and easy to implement.

APPENDIX

Appendix A: Proof of Expectation Identity (3.3)

From (3.2), we can obtain the following equation :

$$E(Y_{ij}^*|u_i) = E\{T_{ij}I(T_{ij} \le C_{ij})|u_i\} + E\{E(T_{ij}|T_{ij} > C_{ij}, u_i)I(T_{ij} > C_{ij})|u_i\}$$

Now by the conditional independence of T_{ij} and C_{ij}, the first term of the right hand side (RHS) of the above equation is as follows

$$\begin{aligned} E(T_{ij}I(T_{ij} \le C_{ij})|u_i) &= E(E(T_{ij}I(T_{ij} \le C_{ij})|T_{ij})|u_i) \\ &= E(T_{ij}|u_i) - \int_0^\infty t_{ij}G(t_{ij})dF_1(t_{ij}|u_i) \end{aligned}$$

and the second term of the RHS is also given by

$$E\{E(T_{ij}|T_{ij} > C_{ij}, u_i)I(T_{ij} > C_{ij})|u_i\} = \int_0^\infty t_{ij}\, G(t_{ij})dF_1\,(t_{ij}|u_i),$$

where $G[F_1]$ are arbitary continuous[conditional] distribution function of $C_{ij}[T_{ij}|u_i]$, respectively. Thus by combining two equations we obtain

$$\begin{aligned} E(Y_{ij}^*|u_i) &= E(T_{ij}|u_i) \\ &= x_{ij}^t\beta + u_i. \end{aligned}$$

Appendix B: Proof of (3.19)

From (3.10), the klth element of A_1 in (3.18) is given by

$$\partial^2 h/\partial\beta_k\partial\beta_l = -\frac{1}{\sigma_\epsilon^2}\sum_{ij\in D} x_{ijk}x_{ijl} - \frac{1}{\sigma_\epsilon^2}\sum_{ij\in C} x_{ijk}\tau(m_{ij}')x_{ijl} \quad (k, l = 1, \cdots, p). \tag{B.1}$$

Let X and Z be partition as follows:

$$X = (X_D^t, X_C^t)^t, \quad Z = (Z_D^t, Z_C^t)^t,$$

where $X_D[X_C]$ and $Z_D[Z_C]$ denote X and Z which depend on $D[C]$, respectively. Then $(B.1)$ can be written as

$$A_1 = -\frac{1}{\sigma_\epsilon^2}X^tWX. \tag{B.2}$$

Similary, A_2 and A_3 in (3.18) can be written as

$$A_2 = -\frac{1}{\sigma_\epsilon^2} X^t W Z \qquad (B.3)$$

and

$$A_3 = -\frac{1}{\sigma_\epsilon^2} Z^t W Z - \frac{1}{\sigma_u^2} I_q. \qquad (B.4)$$

Plugging $(B.2)$ – $(B.4)$ in (3.18) completes the proof of the equation (3.19).

Appendix C : Proofs of MAPHLEs (3.22) and (3.23)

Let us denote that h_A^* is h_A with H being replaced by H^*, i.e.,

$$h_A^* = h + \frac{1}{2}\log\det(2\pi\sigma_\epsilon^2 H^{*^{-1}}),$$

where h and H^* are given by (3.6) and (3.18), respectively.
Then we have

$$\partial h_A^*/\partial\sigma_\epsilon^2 = \partial\ell(\beta,\sigma_\epsilon^2;y|u)/\partial\sigma_\epsilon^2 + \frac{1}{2}[(p+q) - \sigma_\epsilon^2\,\mathrm{trace}\{H^{*^{-1}}(\partial H^*/\partial\sigma_\epsilon^2)\}]/\sigma_\epsilon^2, \qquad (C.1)$$

and

$$\partial h_A^*/\partial\sigma_u^2 = \partial\ell(\sigma_u^2;u)/\partial\sigma_u^2 - \frac{1}{2}\mathrm{trace}\{H^{*^{-1}}(\partial H^*/\partial\sigma_u^2)\}. \qquad (C.2)$$

Now, we want to calculate the equations $(C.1)$ and $(C.2)$. The first term of the RHS of $(C.1)$ is obtained from (3.4) and (3.7) as follows:

$$\partial\ell(\beta,\sigma_\epsilon^2;y|u)/\partial\sigma_\epsilon^2 = -\frac{r}{2\sigma_\epsilon^2} + \frac{1}{2\sigma_\epsilon^2}\sum_{ij\in A}\frac{(y_{ij}^* - \mu_{ij}')^2}{\sigma_\epsilon^2} - \frac{1}{2\sigma_\epsilon^2}\sum_{ij\in C}\tau(m_{ij}'). \qquad (C.3)$$

On the other hand, we have

$$\mathrm{trace}\{H^{*^{-1}}(\partial H^*/\partial\sigma_\epsilon^2)\} = -\gamma_1/\sigma_\epsilon^2, \qquad (C.4)$$

where $\gamma_1 = \sigma_\epsilon^2\mathrm{trace}\{H^{*^{-1}}(\partial H^*/\partial\sigma_\epsilon^2)\}$.
Here

$$\partial H^*/\partial\sigma_\epsilon^2 = \begin{pmatrix} X^t W^* X & X^t W^* Z \\ Z^t W^* X & Z^t W^* Z + \frac{1}{\sigma_u^2} I_q \end{pmatrix},$$

$$W^* = \partial W/\partial\sigma_\epsilon^2 = \begin{pmatrix} 0_D & 0_D \\ 0_C & G_C^* \end{pmatrix},$$

$$G_C^* = \partial G_C / \partial \sigma_\epsilon^2 = \text{diag}\{dm'_{ij}\} \text{ for } ij \in C,$$

and

$$\begin{aligned} dm'_{ij} &= \partial \tau(m'_{ij}) / \partial \sigma_\epsilon^2 \\ &= \frac{1}{2\sigma_\epsilon^2} m'_{ij} V(m'_{ij}) [3m'_{ij} V(m'_{ij}) - 2(V(m'_{ij}))^2 - (m'_{ij})^2 + 1]. \end{aligned}$$

By combining (C.3) and (C.4), the equation (C.1) reduces to

$$\partial h_A^* / \partial \sigma_\epsilon^2 = -\frac{1}{2\sigma_\epsilon^2}[n_1 - (p + q - \gamma_1)] + \frac{1}{2\sigma_\epsilon^2} \sum_{ij \in A} \frac{1}{\sigma_\epsilon^2} (y_{ij}^* - \mu'_{ij})^2, \qquad (C.5)$$

where $n_1 = r + \sum_{ij \in C} \tau(m'_{ij})$.

Next, the first term of the RHS of $(C.2)$ is

$$\partial \ell(\sigma_u^2; u) / \partial \sigma_u^2 = -\frac{q}{2\sigma_u^2} + \frac{1}{2\sigma_u^2} \sum_{i=1}^{q} \frac{u_i^2}{\sigma_u^2} \qquad (C.6)$$

and also

$$\text{trace}\{H^{*-1} (\partial H^* / \partial \sigma_u^2)\} = -\gamma_2 / \sigma_u^2, \qquad (C.7)$$

where $\gamma_2 = \lambda \text{ trace}\{H_{22}^*\}$ and the matrix H_{22}^* is given by bottom right-hand coner of H^{*-1}.

By combining (C.6) and (C.7), the equation (C.2) reduces to

$$\partial h_A^* / \partial \sigma_u^2 = -\frac{1}{2\sigma_u^2}(q - \gamma_2) + \frac{1}{2\sigma_u^2} \sum_{i=1}^{q} \frac{u_i^2}{\sigma_u^2}. \qquad (C.8)$$

Thus by applying $(C.5)$ and $(C.8)$ to (3.21), the proofs of (3.22) and (3.23) are completed.

References

Aitkin, M. A. (1981). "A note on the regression analysis of censored data," *Technometrics*, **23**, 161-163.

Batchelor, J. R. and Hackett, M. (1970). "HLA matching in treatment of burned patients with skin allografts," *Lancet*, **2**, 581-583.

Buckley, J. and James, I. (1979). "Linear regression with censored data," *Biometrika*, **66**, 429-436.

Carriquiry, A. L., Gianola, D., and Fernando, R. L. (1987). "Mixed-model analysis of a censored normal distribution with reference to animal breeding," *Biometrics*, **43**, 929-939.

Cox, D. R. (1972). "Regression models and life tables," *Journal of the Royal Statistical Society B*, **74**, 187-220.

Dempster, A. P., Laird, N. M., and Rubin. D. B. (1977). "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society B*, **39**, 1-38.

Henderson, C. R. (1975). "Best linear unbiased estimation and prediction under a selection model," *Biometrics*, **31**, 423-447.

Lee, Y. and Nelder, J. A. (1996). "Hierarchical generalized linear models(with discussions)," *Journal of the Royal Statistical Society B*, **58**, 619-678.

Miller, R. G. (1976). "Least squares regression with censored data," *Biometrika*, **63**, 449-464.

Patterson, H. D. and Thompson, R. (1971). "Recovery of interblock information when block sizes are unequal," *Biometrika*. **58**, 545-554.

Pettitt, A. N. (1986). "Censored observations, repeated measures and mixed effects models: An approach using the EM algorithm and normal errors," *Biometrika*. **73**, 635-643.

Schall, R. (1991). "Estimation in generalized linear models with random effects," *Biometrika*. **78**, 719-727.

Schmee, J. and Hahn, G. J. (1979). "A simple method for regression analysis with censored data,"" *Technometrics*, **21**, 417-423.

Smith, F. B. and Helms, R, W. (1995). "EM mixed model analysis of data from informatively censored normal distributions," *Biometrics*. **51**, 425-436.

Wolynetz, M.S.(1979). "Maximum likelihood estimation in a linear model from confined and censored normal data," Algorithm AS139, *Applied Statistics*, **28**, 195-206.