Robust Estimation in GLMM

Jiming Jiang and Weihong Zhang

University of California, Davis and San Francisco

1. Generalized Linear Mixed Models

• Examples

McCullagh and Nelder (1989), Breslow and Clayton (1993), Lee and Nelder (1996), Malec et. al. (1997), Ghosh et. al. (1998), etc.

• Definition

a. Given $\alpha_1, \ldots, \alpha_m$, responses $y_1, \ldots, y_N$ indep.;

b. the conditional density

$$f(y_i|\alpha) = \text{exp}\left\{w_i \frac{y_i\theta_i - b(\theta_i)}{\phi} + c_i(y_i, \phi)\right\};$$

c. with $\mu_i = E(y_i|\alpha)$,

$$g(\mu_i) = x_i^{\top}\beta + z_i^{\top}\alpha.$$

2. Inference about GLMM

• Likelihood based: McCulloch (1997), Booth and Hobert (1999).

• Non-likehood based: Breslow and Clayton (1993), Lin and Breslow (1996), Lee and Nelder (1996), Jiang (1998).

3. Robust estimation

• A general idea

θ - vector of unknown parameters under an assumed model, S - a vector of 'base statistics'(typically of larger dimension than θ):

(i) the mean of S is a known function of θ;

(ii) the covariance matrix of S is a known function of θ, or, at least is consistently estimable; and

(iii) certain smoothness and regularity conditions hold.

- First-step estimation

$$BS = Bu(\theta).$$

Best B?

The optimal B is given by $B = U^{\top}V^{-1}$, where U has entries $U_{ij} = \partial u_i / \partial \theta_j$, $V = \text{Cov}(S)$.

However, the optimal B depends on θ.

- Second-step estimation

$\tilde{B}$ - estimated optimal B by the first-step,

$$\tilde{B}S = \tilde{B}u(\theta).$$

- How does it work (for GLMM)?

• Choice of base statistics under GLMM

Jiang (1998):

$$
\begin{aligned}
S_j &= \sum_{i=1}^{N} w_i x_{ij} y_i, \quad 1 \le j \le p, \\
S_{p+j} &= (\sum_{i=1}^{N} w_i z_{ij} y_i)^2 - \sum_{i=1}^{N} (w_i z_{ij} y_i)^2 \\
&= \sum_{i \ne i'} w_i w_{i'} z_{ij} z_{i'j} y_i y_{i'}, \quad 1 \le j \le m.
\end{aligned}
$$

Example 1. A mixed logistic model: Given $\alpha_1, \dots, \alpha_m$, y_{ij}, $1 \le i \le m$, $1 \le j \le n_i$ cond. indep. binary,

$$\text{logit}\{\text{P}(y_{ij} = 1|\alpha)\} = \mu + \alpha_i.$$

Furthermore, α_i's indep. $\sim N(0, \sigma^2)$. The base statistics are $y_{\cdot\cdot}$ and $y_{i\cdot}^2 - y_{i\cdot}$, $1 \le i \le m$.

•Connection with GEE

1) GEE has beeen used mostly in cases where there are indep. groups; e.g. Diggle et. al., 1996,

$$\sum_{i=1}^{m} \left(\frac{\partial \mu_i}{\partial \beta}\right)^{\mathsf{T}} V_i^{-1}(Y_i - \mu_i) = 0.$$

However, in GLMM there may not be indep. groups at all.

2) In GEE the dispersion parameters are often treated as nuisance, while in GLMM they are of main interest (e.g., Prentice, 1988).

3) The two-steps are closely connected here.

4. Performance of robust estimation

- Two simulated examples

1) Mixed logistic model revisited

If $n_i = n$, $1 \le i \le m$, there is zero improvement of 2nd-step over 1st-step.

Let $m = 100$, $n_i = 2$, $1 \le i \le 50$, and $n_i = 6$, $51 \le i \le 100$; $\mu = 0.2$ and $\sigma = 1.0$; simulation size = 1000.

μ			σ		
Mean	Bias	SD	Mean	Bias	SD
.205	.005	.164	.984	-.016	.343
.194	-.006	.157	.976	-.024	.239

Overall MSE: .145 for 1st-step and .082 for 2nd-step, a 43% improvement.

2) Beta-binomial

Given p, $Y_1, \dots, Y_l$ indep. $\sim$ Bernoulli(p), and $p \sim$ Beta$(\pi, 1-\pi)$ $(0 < \pi < 1)$. The marginal distr., denoted by $\mathrm{B}^2(l, \pi)$,

$$
\begin{aligned}
& \mathrm{P}(Y = k) \\
= \; & \frac{1}{k!(l-k)!} \cdot \frac{\Gamma(k+\pi)}{\Gamma(\pi)} \cdot \frac{\Gamma(l-k+1-\pi)}{\Gamma(1-\pi)},
\end{aligned}
$$

$0 \le k \le l$.

$\alpha_1, \dots, \alpha_m$ generated from $N(0, \sigma^2)$. Then, for each i, y_{ij}, $j = 1, \dots, n_i$ generated independently from $\mathrm{B}^2(l, \pi_i)$, where $\pi_i = h(\mu + \alpha_i)$ with $h(x) = e^x/(1+e^x)$. It follows that

$$
\begin{aligned}
E(y_{ij}|\alpha) &= l\pi_i, \\
\mathrm{var}(y_{ij}|\alpha) &= \phi l \pi_i (1-\pi_i),
\end{aligned}
$$

where $\phi = (l+1)/2$.

$m = 40$; $n_i = 4$, $1 \le i \le 20$, and $n_i = 8$, $21 \le i \le 40$; $l = 2$; $\mu = 0.2$, $\sigma = 1.0$; simulation size = 1000.

μ			σ		
Mean	Bias	SD[†]	Mean	Bias	SD
.250	.050	.249	1.132	.132	.371
.252	.052	.256	1.094	.094	.252

Overall MSE: .220 for 1st-step and .141 for 2nd-step, a 36% improvement.

- Theoretical properties

Theorem. Under regularity conditions, the 1st-step estimator is consistent. Under stronger regularity conditions, the 2nd-step estimator is consistent and has the optimal asymptotic covariance matrix.

See Jiang and Zhang (2001, *Biometrika*, to appear)

5. Some applications

- Fetal mortality in mouse litters

Source of data: Brooks et. al. (1997), Table 4: # of dead implants in 1328 litters of mice from untreated experimental animals.

Brooks et. al. (1997) used a Beta-binomial model.

We use a mixed logistic model similar to Example 1.

Estimates of Parameters

	μ		σ	
	Est.	Sd	Est.	Sd
1st-step	-2.276	.047	.644	.059
2nd-step	-2.296	.047	.698	.057

- The salamander mating data

McCullagh and Nelder, 1989, Ch. 14.5.

Three experiments: Summer 1986 (1), Fall 1986 (2).

The same group of 40 salamanders were used or the summer and first fall experiments. A new set of 40 animals was used in the second fall experiment.

McCullagh and Nelder (1989) propose a mixed logistic model with crossed random effects, a special case of GLMM. However, it is assumed that a different set of 40 animals is used in each of the three experiments. Also, see Drum and McCullagh, 1993; Lin and Breslow, 1996; Jiang, 1998; Booth and Hobert, 1999.

A robust estimation approach: assume only

$$E(y_{ijk}|u,v) = h(x_{ij}^{\mathsf{T}}\beta + u_{k,i} + v_{k,j}),$$

where x_{ij} involves an intercept; an indicator of Whiteside female, WS_f, an indicator of Whiteside male, WS_m, and $\mathsf{WS}_f \cdot \mathsf{WS}_m$; $u_{k,i}$ and $v_{k,j}$ are random effects corresponding to the ith female and jth male in group k; and $h(\cdot)$ is chosen as either logit or probit. Furthermore, assume that $u_{k,i}$'s and $v_{k,j}$'s are indep., $u_{k,i} \sim N(0,\sigma^2)$, $v_{k,j} \sim N(0,\tau^2)$.

Results

β_0	β_1	β_2	β_3	σ	τ
.95	-2.92	-.69	3.62	.99	1.28
(.55)	(.87)	(.60)	(1.02)	(.59)	(.57)
.56	-1.70	-.40	2.11	.57	.75
(.31)	(.48)	(.35)	(.55)	(.33)	(.32)