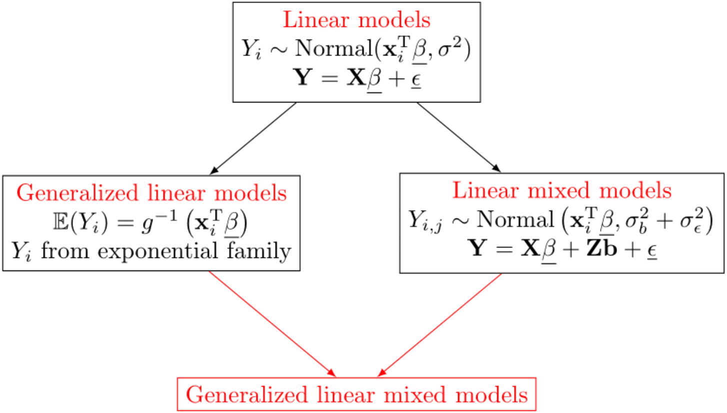
MA50260 Statistical Modelling

Lecture 18: GLMMs - Estimation

Ilaria Bussoli

April 19, 2024

Generalised Linear Mixed Models (I)



Generalised Linear Mixed Models (II)

Suppose we have I groups and J observations per group.

A GLMM generally comprises three components:

- **Linear predictor** $\eta_{i,j} = \mathbf{x}_{i,j}^{\mathrm{T}} \underline{\beta} + \mathbf{z}_{i,j}^{\mathrm{T}} \mathbf{b}$.
- **Link function** $g(\mu_{i,j}) = \eta_{i,j}$ linking $\eta_{i,j}$ to $\mu_{i,j} = \mathbb{E}(Y_{i,j})$.
- **Probability distribution** $Y_{i,j} \sim F\left(\mu_{i,j}\right)$ from the exponential family.

A forth component specifies the distribution of the random effects $\mathbf{b}$.

Exercise - Number of children

- Let $Y_{i,j}$ be the number of children of woman j in country i
- We further have knowledge on the woman's age $x_{i,j}$

How would we define a GLMM model for $Y_{i,j}$?

Recap: Estimation of GLMs and LMMs

For GLMs, the log-likelihood function is

$$\ell(\underline{\beta}, \phi) = \frac{1}{\phi} \left\{ \sum_{i=1}^{n} w_i [y_i \theta_i - b(\theta_i)] \right\} + \sum_{i=1}^{n} c(y_i, \phi).$$

We introduced the IRWLS method to derive estimates for $\underline{\beta}$.

Recap: Estimation of GLMs and LMMs

For GLMs, the log-likelihood function is

$$\ell(\underline{\beta}, \phi) = \frac{1}{\phi} \left\{ \sum_{i=1}^{n} w_i [y_i \theta_i - b(\theta_i)] \right\} + \sum_{i=1}^{n} c(y_i, \phi).$$

We introduced the IRWLS method to derive estimates for $\underline{\beta}$.

For LMMs, we considered the log-likelihood function

$$\ell(\underline{\beta}, \underline{\theta}) = -\frac{1}{2} \left\{ \log |\mathbf{V}(\underline{\theta})| + \left(\mathbf{y} - \mathbf{X}\underline{\beta} \right)^{\mathrm{T}} \mathbf{V}(\underline{\theta})^{-1} \left(\mathbf{y} - \mathbf{X}\underline{\beta} \right) \right\}.$$

We obtained expressions for the estimates of $\underline{\beta}$ and we introduced REML for estimating σ_b^2 and σ_ϵ^2.

Estimation of GLMMs

We specify

$$\mathbf{b} \sim \mathrm{MVN}_I(0, \sigma_b^2 \mathbf{I}_I).$$

with density $h(\mathbf{b} \mid \underline{\gamma})$.

Let's derive the likelihood function for the GLMM.

Estimation of GLMMs

We specify

$$\mathbf{b} \sim \mathrm{MVN}_I(0, \sigma_b^2 \mathbf{I}_I).$$

with density $h(\mathbf{b} \mid \underline{\gamma})$.

Let's derive the likelihood function for the GLMM.

We derive the likelihood function as

$$L(\underline{\beta}, \phi, \underline{\gamma}) = \prod_{i=1}^{n} \int f(y_i \mid \underline{\beta}, \phi, \mathbf{b}) h(\mathbf{b} \mid \underline{\gamma}) \mathrm{d}\mathbf{b},$$

Penalized Quasi-Likelihood (I)

Estimate the parameters via an iterative numerical algorithm.

Penalized Quasi-Likelihood (I)

Estimate the parameters via an iterative numerical algorithm.

Define an adjusted variable

$$\mathbf{v}_k = \underline{\hat{\eta}_k} + (\mathbf{y} - \underline{\hat{\mu}_k})^{\mathrm{T}} \left. \frac{\mathrm{d}\underline{\eta}}{\mathrm{d}\underline{\mu}} \right|_{\underline{\eta_k}} .$$

This is used to approximate the likelihood, and is deemed a **quasi-likelihood** approach.

Penalized Quasi-Likelihood (I)

Estimate the parameters via an iterative numerical algorithm.

Define an adjusted variable

$$\mathbf{v}_k = \underline{\hat{\eta}_k} + (\mathbf{y} - \underline{\hat{\mu}_k})^{\mathrm{T}} \left. \frac{\mathrm{d}\underline{\eta}}{\mathrm{d}\underline{\mu}} \right|_{\underline{\eta_k}} .$$

This is used to approximate the likelihood, and is deemed a **quasi-likelihood** approach.

A (linear) mixed model is fitted for the adjusted variable

$$\mathbf{v} \mid \mathbf{b} \sim \mathrm{MVN}_n \left(\mathbf{X}\underline{\beta} + \mathbf{Zb} \, , \, \mathbf{W}^{-1}\phi \right) , \qquad \mathbf{b} \sim \mathrm{MVN}_I \left(0 \, , \, \mathcal{D} \right)$$

Penalized Quasi-Likelihood (II)

Perform the following algorithm until convergence (up to a tolerance):

1. Set initial estimates $\hat{\mathbf{b}}_0$ and $\underline{\hat{\beta}_0}$.
2. Compute the adjusted variable $\mathbf{v}_0$.
3. Compute the weight matrix $\mathbf{W}_0$.
4. Fit the **linear mixed model**

$$\mathbf{v}_0 = \mathbf{X}\underline{\beta} + \mathbf{Z}\mathbf{b} + \underline{\epsilon},$$

   with $\underline{\epsilon} \sim \mathrm{MVN}_n\left(0, \mathbf{W}^{-1}\phi\right)$ and $\mathbf{b} \sim \mathrm{MVN}_l(0, \mathcal{D})$.

5. Iterate steps 2-4 until convergence (subject to some tolerance).

Gauss-Hermite Quadrature

- Approximate the (true) likelihood via numerical methods.
- Gauss-Hermite numerical approaches are used for integrals of the form $\int m(x) \exp(-x^2) \mathrm{d}x$
- Evaluate $m(x) \exp(-x^2)$ at a number of points (the more the better)
- Estimates are more reliable than PQL, but also more computationally expensive.

Example - Ohio wheeze data (I)

- 536 children were studied for four years from age seven to ten
- The observed response y_i is whether child i wheezed or not
- Explanatory variables:
  - Age of the child
  - Whether the parents smoked when the child was seven

How would we define a GLMM to analyse the data?

Example - Ohio wheeze data (II)

```
library( faraway, warn.conflicts = F )
data( "ohio" )
head( ohio )
```

```
##   resp id age smoke
## 1    0  0  -2     0
## 2    0  0  -1     0
## 3    0  0   0     0
## 4    0  0   1     0
## 5    0  1  -2     0
## 6    0  1  -1     0
```

Example - Ohio wheeze data (III) - PQL Estimates

```r
library( MASS )
estimPQL <- glmmPQL( resp~age+smoke, random= ~1|id,
                     data=ohio, family = binomial,
                     verbose=F)
estimPQL$coefficients$fixed
```

```
## (Intercept)         age       smoke
##  -2.7658365  -0.1815756   0.3251839
```

We find $\underline{\hat{\beta}} = (-2.77, -0.18, 0.33)^T$.

Example - Ohio wheeze data (IV) - Gauss-Hermite Quadrature

```r
library( lme4 )
```

```
## Loading required package: Matrix
```

```r
estimGH <- glmer( resp~age+smoke + (1|id), data=ohio,
                  nAGQ = 25, family = binomial )
estimGH@beta
```

```
## [1] -3.1015338 -0.1756312  0.3985708
```

We obtain $\underline{\hat{\beta}} = (-3.1, -0.18, 0.40)^T$.