Lecture 6: Rao's MINQUE for Variance Components Estimation, Non-Gaussian Adjustment (09-13-2007)

1 Content

1) (Rao, 1970) **MINQUE** for variance components estimation. MINQUE=minimum norm quadratic unbiased estimation.

2) Non-Gaussian adjustment for variance-covariance of estimated parameters.

2 MINQUE (Rao, 1970)

2.1 Sample Model

$$Y = X\beta + Z_1\alpha_1 + Z_2\alpha_2 + \cdots + Z_m\alpha_m + \epsilon \tag{1}$$

i.e., we have multiple (m) random effects. Where,

1) α_i ~N(0,$\sigma_i^2 I_{q_i}$),($i = 1, ..., m$), with q_i levels.

2) ϵ ~N(0,$\sigma_0^2 I_N$), with $q_0 = N$ levels. N is number of total observations.

3) α_is and ϵ are all independent.

Then we have

1) $V(y) = \sum_{i=0}^m \sigma_i^2 Z_i Z_i'$ ($Z_0 = I_N$)

2) $\theta = (\sigma_0^2, \sigma_1^2, \ldots, \sigma_m^2)'$

2.2 Objectives

Summarization in (Ravishanker and Dey, 2002): MINQUE is a criterion for estimating a **linear** function of variance components

$$F = \sum_{j=0}^m b_j \sigma_j^2 \tag{2}$$

by a quadratic function of y

$$y'Ay \tag{3}$$

with requirements

1) Translation invariant: If $\beta \rightarrow \beta - \beta_0$ (any β_0), i.e., a different model (responses $y \rightarrow y_d$=?) under translation, then

$$y' A y = y_d' A y_d \tag{4}$$

For which $Ax = 0$ suffices.

2) Unbiased:

$$E(y' A y) = F \tag{5}$$

2.3 Estimation Elicitation

If random effects and random errors are known, then **reasonable** estimator for F:

$$\hat{F} = \sum_{i=1}^{m} b_i \frac{\alpha_i' \alpha_i}{q_i} + b_0 \frac{\epsilon' \epsilon}{q_0} = \eta' \Delta \eta \tag{6}$$

where,

$$\eta' = (\alpha_1', \alpha_2', \ldots, \alpha_m', \epsilon'), \text{ and}$$

$$\Delta = diag((b_1/q_1) 1_{q_1}, \ldots, (b_m/q_m) 1_{q_m}), (b_0/N) 1_N)$$

The **required** estimator is of form

$$y' A y = \eta' (Z' A Z) \eta \tag{7}$$

compare it with

$$\hat{F} = \eta' \Delta \eta \tag{8}$$

and make them close, we want to minimize

$$||Z' A Z - \Delta||. \tag{9}$$

2.4 Solution for MINQUE

Denote $T = \sum_{j=0}^{m} Z_i Z_i' = ZZ'$, Rao (1971) showed this is a question of

1) minimizing: Trace$[(AT)^2]$ subject to $AX = 0$.

2) Solution: $A = \sum_{j=0}^{m} \lambda_j S T_j S$.

3) where,

$$S = T^{-1} - T^{-1} X (X' T^{-1} X)^{-} X' T^{-1}$$

and λ_is are solutions of

$$\sum_{j=0}^{m} \lambda_j \text{trace}[S T_j S T_l] = b_l, \ (l = 0, 1, \ldots, m).$$

3 Misspecifying Random Effects and Adjustment

3.1 Model

(Verbeke and Lesaffre, 1997): longitudinal data (Hight$_{ij}$ of Girl i at age j=6 to 10) with bivariate random effects.

$$y_i = X_i\alpha + Z_i b_i + \epsilon_i \tag{10}$$

where,

1) Var(b_i)=$D(\psi)$

2) Var(ϵ_{ij})=σ^2

3) All parameters to estimate: (β, ψ and σ^2)

e.g.,

$$\text{Height}_{ij}\text{=}(b_{1i}\text{+}\beta_1\text{Group A}_i\text{+}\beta_2\text{Group B}_i \text{ +}\beta_3\text{Group C}_i)$$
$$+ \ (b_{2i}\text{+}\beta_4\text{Group A}_i\text{+}\beta_5\text{Group B}_i \text{ +}\beta_6\text{Group C}_i)\text{Age}_{ij}\text{+}\epsilon_{ij}$$

3.2 Recall: MLE Theory

z When the model is correct, the maximum likelihood estimators are consistent and asymptotically normally distributed with inverse Fisher information matrix as asymptotic variance-covariance matrix. How about parameter inference under **wrong** model?

3.3 Misspecifying Random Effects

When the true random effects distribution is not normal, the maximum likelihood estimators for **fixed effects** and **variance components** in linear mixed models obtained under simple **normal** assumption are **consistent** and **asymptotically normally distributed** (Recall **ideal** MLE where likelihood is correctly assumed).

3.4 Asymptotic Σ Adjustment for ML

3.4.1 Verbeke and Lesaffre (1997)

1) If the true random effects are not normally distributed, while model is estimated under Gaussian assumption for random effects, then

   ◇ A sandwich-type correction to the inverse Fisher information matrix $I(\theta$=fixed effects, variance components) is needed for asymptotic variance-covariance matrix of estimated parameters.

◇ This correction may be more useful for the random-effects **covariance matrix**.

◇ Let $q_{ik}(\theta)$ and $q_{ikl}(\theta)$ denote the first and second order partial derivatives of $\log f_i(y_i|\theta)$ w.r.t θ_k and θ_l respectively, then we define

$$A_N(\theta) = \left[-\frac{1}{N}\sum_{i=1}^{N} q_{ikl}(\theta) \right]_{k,l}, B_N(\theta) = \left[\frac{1}{N}\sum_{i=1}^{N} q_{ik}(\theta) q_{il}(\theta) \right]_{k,l}, \tag{11}$$

with expectations $\bar{A}_N(\theta)$ and $\bar{B}_N(\theta)$.

◇ **(Theorem)** (Verbeke and Lesaffre, 1997)
Under general regularity conditions, we have that $\hat{\theta}_N$ is asymptotically normally distributed with mean θ^* and with asymptotic covariance matrix

$$\bar{A}_N^{-1}(\hat{\theta})\bar{B}_N(\hat{\theta})\bar{A}_N^{-1}(\hat{\theta})/N$$

where $\hat{\theta}$ is MLE.

3.4.2 Results by Jiang

1) (Jiang, 2007): Quasi-likelihood (including special normal case) leads to same REML equations as Gaussian model.

2) (Jiang, 1996): The asymptotic covariance matrix of the REML estimator (under Gaussian model) is given by

$$\Sigma_R = \left\{ E\left(\frac{\partial^2 l_R}{\partial\theta\partial\theta'} \right) \right\}^{-1} \mathrm{Var}\left(\frac{\partial l_R}{\partial\theta} \right) \left\{ E\left(\frac{\partial^2 l_R}{\partial\theta\partial\theta'} \right) \right\}^{-1} \tag{12}$$

where θ is variance components.

3) (Jiang, 1998): The asymptotic covariance matrix of the ML estimator (under Gaussian model) is given by

$$\Sigma = \left\{ E\left(\frac{\partial^2 l}{\partial\psi\partial\psi'} \right) \right\}^{-1} \mathrm{Var}\left(\frac{\partial l}{\partial\psi} \right) \left\{ E\left(\frac{\partial^2 l}{\partial\psi\partial\psi'} \right) \right\}^{-1} \tag{13}$$

where ψ is fixed effects and variance components: (β,θ).

4) Estimation: Textbook (Chapter 1.4.2).

4 References

◇ Jiming Jiang (1996), REML estimation: Asymptotic behavior and related topics. Ann.of Statistics, 24: 255-286.

◇ Jiming Jiang (1998), Asymptotic properties of the empirical BLUP and BLUE in linear mixed models, Statistica Sinica 8, 861-885. .

◇ Jiming Jiang (2007), Linear and Generalized Linear Mixed Models and Their Applications. Springer.

◇ Geert Verbeke and Emmanuel Lesaffre (1997), The effects of misspecifying the random-effects distribution in linear mixed models for longitudinal data. Computational Statistics and Data Analysis. 23: 541-556.

◇ Rao, C.R. (1970), Estimation of heteroscedastic variances in linear models. JASA 65(329), 161-172.

◇ Ravishanker, N. and Dey, K. D. (2002), A First Course in Linear Model Theory, Chapman & Hall/CRC.