Lecture 5: Henderson's Decomposition of SSR (09-11-2007)

1 Motivation

1) (Previous classes:) **MLE** and **REML** estimation for variance components estimation in linear mixed models may be **biased** even if the model is **correctly** specified.

2) (Previous classes:) **MLE** and **REML** estimation for variance components estimation in linear mixed models requires Gaussian model assumption.

3) We now introduce **Henderson's Decomposition of SSR** (Sum of Squares of Regression/Residue) method for **unbiased** variance components estimation in linear mixed models where **no** Gaussian assumption is needed.

2 Original Model

Data: "artificial breeding of diary cows" (Henderson, 1953).

$$y_{h,i,j,k} = \mu + a_h + h_i + s_j + (hs)_{ij} + e_{h,i,j,k} \tag{1}$$

where,

1) μ: common for all

2) a_h: common for h-th year (may be fixed effects)

3) h_i: i-th herd ($1 \to 2{,}000$, random effects, σ_h^2)

4) s_j: common for daughters of the j-th sire ($1 \to 100$, random effects, σ_s^2)

5) $(hs)_{ij}$: peculiar to all records by the daughters of j-th sire in the i-th herd (σ_{hs}^2).

6) $\epsilon_{h,i,j,k}$: peculiar to each record (σ_ϵ^2).

3 Sample Model for SSR Derivation

$$Y = X\beta + Z_1\alpha_1 + Z_2\alpha_2 + \epsilon = W\gamma + \epsilon, \tag{2}$$

Note

1) If fixed effects are mistakenly taken as random effects, then the variance components estimation for other random effects (α_1 and α_2) are biased.

2) Method III in (Hendenson, 1953) has advantages:

   1) "It gets around the difficulty of fixed effects in the model..."

   2) "It yields unbiased estimates even if certain elements are correlated..."

3) We are stimulated to study this method.

   1) Get **unbiased** variance components estimates ($\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$) for random effects (b_1,b_2) by incorporating fixed effects design matrix X (**Focus**)

   2) Get good estimate $\hat{\beta}$ of fixed effects (β) **accordingly**?

4 Matrix Algebra Notations

For full space W spanned by matrices (X,Z): W=(X,Z). We have (considering possible non-full-ranks)

1) Projection matrix: projecting y onto W:

$$P_W(y) = W(W'W)^{-}W'y \tag{3}$$

2) Denote (difference between **full** projection and **partial** projection):

$$P_{Z\ominus X} = P_{(X,Z)} - P_X \tag{4}$$

   i.e.,

$$P_{(X,Z)} = P_{Z\ominus X} + P_X \tag{5}$$

   where, $Z \ominus X = P_{X^\perp}(Z)$.
   (Graph example for perpendicular spaces?)

5 SSRs and Extra SSRs

5.1 Sample Model Continued

For linear mixed model (LMM)

$$Y = X\beta + Z_1 b_1 + Z_2 b_2 + \epsilon = W\gamma + \epsilon \tag{6}$$

where, $W = (X, Z_1, Z_2)$, $\gamma = \begin{pmatrix} \beta \\ b_1 \\ b_2 \end{pmatrix}$. Let $(Z = Z_1, Z_2)$. (which is fixed effects?, which is random effects?).

5.2 Projection

Project y onto space $W = (X, Z)$

$$P_W(y) = P_{(X,Z_1,Z_2)}(y).$$

5.3 Definitions: SSRs

- **SSR** (β,b_1,b_2):

$$\text{SSR}(\beta,\ b_1,\ b_2)\text{=(y)}'P_W(y).$$

- **SSR**(β,b_1):

$$\text{SSR}(\beta,\ b_1)\text{=(y)}'P_{X,Z_1}(y).$$

- **SSR**(β):

$$\text{SSR}(\beta)=(y)'P_X(y).$$

5.4 Definitions: Extra SSRs

- **SSR**($\alpha|\beta$):

$$\text{SSR}(\beta,\ \alpha)-\text{SSR}(\beta)=(y)'[P_{(X,Z)} - P_{(X)}](y) = (y)'P_{(Z\ominus X)}(y)$$

- **SSR**($\alpha_2|\beta, \alpha_1$):

$$\text{SSR}(\beta,\ \alpha_1,\alpha_2)-\text{SSR}(\beta,\alpha_1)=(y)'[P_{(X,Z_1,Z_2)} - P_{(X,Z_1)}](y) = (y)'P_{(Z_2\ominus(X,Z_1))}(y)$$

5.5 SSE

- **SSE**($\beta, \alpha_1, \alpha_2$):

$$y'I_Ny - SSR(\beta,\alpha_1,\alpha_2)=(y)'[I - P_{(X,Z_1,Z_2)}](y) = (y)'P_{(X,Z)^\perp}(y)$$

6 Expectations: SSRs and SSE

Any preceding SSR and/or SSE has an expectation function of three components (σ^2_ϵ,$\sigma^2_{b_1}$,$\sigma^2_{b_2}$).

1) E(**SSR**($\alpha|\beta$))=σ^2_0 (rank((X,Z))-rank(X)) +σ^2_1 trace($Z_1'P_{X^\perp}Z_1$)+σ^2_2 trace($Z_2'P_{X^\perp}Z_2$)
2) E(**SSR**($\alpha_2|\beta, \alpha_1$))=σ^2_0 (rank((X,Z))-rank(X,Z_1))+σ^2_2 trace($Z_2'P_{(X,Z_1)^\perp}Z_2$)
3) E(**SSE**)=σ^2_0 (N-rank(X,Z))

7 Unbiased Variance Components Estimation

The following expectation equations guarantee unbiased estimators for variance components.

1) SSR($\alpha|\beta$))=E(**SSR**($\alpha|\beta$))

2) SSR($\alpha_2|\beta, \alpha_1$))=E(**SSR**($\alpha_2|\beta, \alpha_1$))

3) SSE=E(**SSE**)

8 Expectation Derivations

- ⋄ Matrix trace property: Textbook, Lemma 1.1 (p 27).
- ⋄ Projection matrix algebra (orthogonality).

9 More Random Effects? b_1, b_2 and b_3

See Henderson 1953 original model, how to proceed within SSR framework?

10 Remark

We also kind of multiply the fixed effects design matrix X by matrix projection (similar to REML?)

11 References

- ⋄ Henderson, C. R. (1953), Estimation of variance and covariance components. Biometrics. Vol 9 (No 2). pp. 226-252.
- ⋄ Khuri, A.I., Thomas Matthew and B.K. Sinha (1998), Statistical tests for mixed linear models, New York: Wiley.
- ⋄ Jiming Jiang (2007), Linear and Generalized Linear Mixed Models and Their Applications. Springer.
- ⋄ Jiming Jiang (2001), previous notes.