International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:9, No:12, 2015

An Estimation of Variance Components in Linear Mixed Model

Shuimiao Wan, Chao Yuan, Baoguang Tian

Abstract—In this paper, a linear mixed model which has two random effects is broken up into two models. This thesis gets the parameter estimation of the original model and an estimation's statistical qualities based on these two models. Then many important properties are given by comparing this estimation with other general estimations. At the same time, this paper proves the analysis of variance estimate (ANOVAE) about σ^2 of the original model is equal to the least-squares estimation (LSE) about σ^2 of these two models. Finally, it also proves that this estimation is better than ANOVAE under Stein function and special condition in some degree.

Keywords—Linear mixed model, Random effects, Parameter estimation, Stein function.

I. Introduction

THE model we treat here is linear mixed model which has two variance components, described by

$$y = X\beta + U\xi + \varepsilon, \tag{1}$$

where Y is the $n \times 1$ observation vector, X is a $n \times t$ design matrix which we are known, U is a $n \times s$ known design matrix, β is a $t \times 1$ fixed effect, ξ is a $s \times 1$ random effect, ε is a $n \times 1$ error of random effect, and the basic assumptions of (1) for ξ and ε are $\xi_{s\times 1} \sim N(0, {\sigma_1}^2 I)$ and $\varepsilon_{n\times 1} \sim N(0, \sigma^2 I)$ [1].

As is known to us all, the LSE of the estimable function $c'\beta$ in the model (1) is

$$c'\hat{\beta}_0 = c'(X'X)^+ X'y.$$

The ANOVAE of variance component σ^2 is

$$\hat{\sigma}^2(A) = \frac{y'(I - P_{X\vdots U})y}{n - r(X\vdots U)}.$$

And the ANOVAE of variance component ${\sigma_1}^2$ is

$$\hat{\sigma_1}^2(A) = \frac{y'(P_{X\vdots U} - P_X)y - (r(X\vdots U) - r(X))\hat{\sigma}^2(A)}{tr((P_{X\vdots U} - P_X)UU')},$$

where $M(A)$ is a space composed by the column vectors of matrix A, P_A represents the rectangular projection of $M(A)$.

At first we give some conclusions which will play key roles in the following text [2]-[3].

Lemma 1. *Assume that r is the rank of matrix $A_{m\times n}$, P and Q are orthogonal matrices, then $n \times p$ matrix,then*

$$A = P \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} Q',$$

where $A_r = diag(\lambda_1, \lambda_2, \dots \lambda_r)$, $\lambda_1^2, \lambda_2^2, \dots \lambda_r^2$ are non-zero characteristic values of $A'A$, λ_i is singular value of matrix A and $\lambda_i > 0, i = 1, 2, \dots r$.

Lemma 2. *Assume that $y = X\beta + e$, $E(e) = 0$, $Cov(e) = \sigma^2\Sigma$, $\Sigma \geq 0$. For any estimable function $c'\beta$, its LSE $c'\tilde{\beta} = c'(X'X)^+X'y$ is BLUE $\Longleftrightarrow$ $P_X\Sigma$ is a symmetric matrix.*

It carries on the singular value decomposition of the design matrix U in model (1) by lemma 1 as follows:
Suppose that $r(U) = r$, P and Q are orthogonal matrices, then we have

$$U = P \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} Q'.$$

II. An Estimation of Fixed Effect $c'\hat{\beta}$

As we all know ξ is independent identically distributed to $Q'\xi$ and ε is independent identically distributed to $P'\varepsilon$ in former context, so we denote $\tilde{\xi} = Q'\xi$, $\tilde{\varepsilon} = P'\varepsilon$, then

$$y = X\beta + P \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} Q'\xi + \varepsilon$$

$$\Longleftrightarrow y = X\beta + P \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} \tilde{\xi} + \varepsilon$$

$$\Longleftrightarrow P'y = P'X\beta + \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} \tilde{\xi} + \tilde{\varepsilon}.$$

Let

$$P' = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \quad \begin{matrix} \dots r \\ \dots n - r \end{matrix},$$

$$\tilde{\xi} = (\tilde{\xi_1} \ \vdots \ \tilde{\xi_2}), \qquad \tilde{\xi} = \begin{pmatrix} \tilde{\xi_1} \\ \tilde{\xi_2} \end{pmatrix},$$

then we have two models that are equivalent to the original model

$$P_1 y = P_1 X\beta + \Lambda_r \tilde{\xi_1} + \tilde{\varepsilon_1}, \tag{2}$$
$$P_2 y = P_2 X\beta + \tilde{\varepsilon_2}. \tag{3}$$

S. Wan,C. Yuan and B. Tian are with the Department of Mathematics, Qingdao University of Science & Technology, Shandong 266061, China (e-mail: 675593912@qq.com).

It is readily to find the fixed effects and variance components of new models are consistent to them in original models. In the model of (2) and (3),

$$\tilde{\xi 1} \sim N(0, \sigma_1^2 I_r),$$

$$\tilde{\varepsilon_1} \sim N(0, \sigma^2 I_r),$$

$$\tilde{\varepsilon_2} \sim N(0, \sigma^2 I_{n-r}).$$

The LSE of the estimable function $c_1'\beta$ in model (2) is

$$c_1'\hat{\beta}_1 = c_1'(X'P_1'P_1X)^- X'P_1'P_1 y, c_1 \epsilon M(X'P_1').$$

The LSE of the estimable function $c_2'\beta$ in model (3) is

$$c_2'\hat{\beta}_2 = c_2'(X'P_2'P_2X)^- X'P_2'P_2 y, c_2 \epsilon M(X'P_2').$$

The LSE of σ^2 [4] in model (3) is

$$\hat{\sigma}^2(2) = \frac{[P_2y - P_2X\hat{\beta}_2]'[P_2y - P_2X\hat{\beta}_2]}{(n-r) - r(P_2X)}$$

$$= \frac{y'P_2'(I - P_{P_2X})P_2y}{(n-r) - r(P_2X)}.$$

Theorem 1. *In the model of (1), there is*

$$c'\hat{\beta} = c_1'\hat{\beta}_1 + c_2'\hat{\beta}_2$$

where $c_1'\hat{\beta}_1$ is irrelevant to $c_2'\hat{\beta}_2$, $c_1 \epsilon M(X'P_1')$, $c_2 \epsilon M(X'P_2')$.

Proof. Let A be a $m \times n$ matrix and B be a orthogonal matrix. For the partitioned matrix $B = (B_1 \vdots B_2)$, then we obtain

$$M(A) = M(AB_1) \oplus M(AB_2).$$

To prove this equality above-mentioned, the proof by contradiction is used here. As is known to us all,

$$M(AB_1) \cap M(AB_2) = \{0\} \Longleftrightarrow$$

$$M(A) = M(AB_1) \oplus M(AB_2).$$

We assume that $\exists \alpha \neq 0$, make $\alpha \epsilon M(AB_1)$ and $\alpha \epsilon M(AB_2)$, then $\exists t_1 \neq 0, t_2 \neq 0$, make $\alpha = AB_1 t_1$ and $\alpha = AB_2 t_2$, we get $B_1 t_1 = B_2 t_2$ due to the arbitrariness of A. Because B_1 and B_2 are two matrices which their column vectors are orthogonal mutually, the necessary and sufficient condition of equation validated is $t_1 = t_2 = 0$. This is contradictory to $t_1 \neq 0, t_2 \neq 0$. Hence the hypothesis is not valid. So $M(A) = M(AB_1) \oplus M(AB_2)$. Then for $\forall$ $c \epsilon M(X'), \exists c_1 \epsilon M(X'P_1'), c_2 \epsilon M(X'P_2')$, we have

$$c = c_1 + c_2,$$

then it is readily verified that

$$c'\hat{\beta} = c_1'\hat{\beta}_1 + c_2'\hat{\beta}_2.$$

Except the conclusion above-mentioned, we also get something new about estimable function $c'\hat{\beta}$ in model (1). $c'\hat{\beta}_1$ is BLUE of $c'\beta$ in model (2) and $c'\hat{\beta}_2$ is BLUE of $c'\beta$ in model (3). It is not difficult to prove the question based on Unified Theory of Least Squares.

III. The Properties of This Estimation

Before giving the properties of this estimation, we should prove the following lemmas first.

Lemma 3. *Suppose that $c_1 \epsilon M(X'P_1')$, then*

$$c_1'(X'P_1'P_1X)^+ X'UU'X(X'P_1'P_1X)^+ c_1$$

$$\leq c_1'(X'X)^+ X'UU'X(X'X)^+ c_1,$$

$$c_1'(X'P_1'P_1X)^+ c_1 \leq c_1'(X'X)^+ c_1.$$

Proof. By the lemma 2, $c_1'\hat{\beta}_1$ is BLUE of $c_1\beta$, $c_1 \epsilon M(X'P_1')$ in the model of (2). Hence

$$Var(c_1'\hat{\beta}_1) \leq Var(c_1'\hat{\beta}_0),$$

that is

$$c_1'(X'P_1'P_1X)^+ X'P_1'P_1(UU'\sigma_1^2 +$$

$$I\sigma^2)P_1'P_1X(X'P_1'P_1X)^+ c_1$$

$$\leq c_1'(X'X)^+ X'(UU'\sigma_1^2 + I\sigma^2)X(X'X)^+ c_1,$$

$$\sigma_1^2[c_1'(X'P_1'P_1X)^+ X'UU'X(X'P_1'P_1X)^+ c_1] +$$

$$\sigma^2[c_1'(X'P_1'P_1X)^+ c_1]$$

$$\leq \sigma_1^2[c_1'(X'X)^+ X'UU'X(X'X)^+ c_1] +$$

$$\sigma^2[c_1'(X'X)^+ c_1].$$

According to the arbitrariness of σ_1^2 and σ^2, lemma 3 is tenable.

Similarly to the following lemma:

Lemma 4. *As $c_2 \epsilon M(X'P_2')$, there is*

$$c_2'(X'P_2'P_2X)^+ c_2 \leq c_2'(X'X)^+ c_2.$$

Theorem 2. *$c'\hat{\beta}$ is the unbiased estimation of estimable function $c'\beta$ in the model of (1). Let $P_1 P_X P_2' = 0$, then*

$$Var(c'\hat{\beta}) = Var(c'\hat{\beta}_0).$$

Proof.

$$Var(c'\hat{\beta}) = Var(c_1'\hat{\beta}_1) + Var(c_2'\hat{\beta}_2)$$

$$= c_1'(X'P_1'P_1X)^+ X'P_1'P_1(UU'\sigma_1^2 + I\sigma^2)P_1'P_1$$

$$* X(X'P_1'P_1X)^+ c_1 + c_2'(X'P_2'P_2X)^+ c_2 \sigma^2 = \sigma_1^2$$

$$* [c_1'(X'P_1'P_1X)^+ X'P_1'P_1UU'P_1'P_1X(X'P_1'P_1X)^+ c_1]$$

$$+ \sigma^2[c_1'(X'P_1'P_1X)^+ c_1 + c_2'(X'P_2'P_2X)^+ c_2].$$

$$Var(c'\hat{\beta}_0) = (c_1 + c_2)'(X'X)^+X'$$

$$*(UU'\sigma_1^2 + I\sigma^2)X(X'X)^+(c_1 + c_2)$$

$$= \sigma_1^2[c_1'(X'X)^+X'UU'X(X'X)^+c_1 + c_2'(X'X)^+X'$$

$$*UU'X(X'X)^+c_2 + 2c_1'(X'X)^+X'UU'X(X'X)^+c_2]$$

$$+\sigma^2[c_1'(X'X)^+c_1 + c_2'(X'X)^+c_2 + 2c_1'(X'X)^+c_2].$$

As $P_1P_XP_2' = 0$, we get

$$c_1'(X'X)^+c_2 = 0,$$
$$c_1'(X'X)^+X'UU'X(X'X)^+c_2 = 0,$$
$$c_2'(X'X)^+X'UU'X(X'X)^+c_2 = 0.$$

By lemma 2 and theorem 1,

$$Var(c'\hat{\beta}) \le Var(c'\hat{\beta}_0).$$

Thus $P_1P_XP_2' = 0 \Longleftrightarrow P_XP_U$ is exchangeable [5]. While $P_1P_XP_2' = 0$, $c'\hat{\beta}_0$ is BLUE of $c'\hat{\beta}$, we have

$$Var(c'\hat{\beta}) \ge Var(c'\hat{\beta}_0).$$

Above all,

$$Var(c'\hat{\beta}) = Var(c'\hat{\beta}_0)$$

is testified completely.

When $P_1P_XP_2' = 0$, the new estimation of $c'\hat{\beta}$ and the LSE of $c'\hat{\beta}_0$ achieve the best estimation simultaneously under mean square error by theorem 2.

Theorem 3. $\hat{\sigma}^2(2) = \hat{\sigma}^2(A)$.

Proof.

$$r(U \vdots X) = r(U \vdots (I - P_UX))$$

$$= r(U) + r(I - P_UX),$$

$$P_U = I - P_{P_2'}.$$

P_2' is a matrix with full column rank, then

$$r(U \vdots X) = r + r(P_2X).$$

Notice that

$$I - P_{X \vdots U} = I - P_U - P_{(I-P_0)X}$$

$$= P_{P_2'} - P_{P_2'X} = P_2'(I - P_{P_2X})P_2,$$

$$\hat{\sigma}^2(2) = \frac{[P_2y - P_2X\hat{\beta}_2]'[P_2y - P_2X\hat{\beta}_2]}{(n-r) - r(P_2X)}$$

$$= \frac{y'(I - P_{X \vdots U})y}{n - [r + r(P_2X)]}$$

$$= \frac{y'(I - P_{X \vdots U})y}{n - r(X \vdots U)}$$

$$= \hat{\sigma}^2(A).$$

By the former theorems, we obtain ANOVAE in common use is LSE of error variance in the model of (3).

Following that we discuss the estimation of variance component σ_1^2 when UU' has only one characteristic value. When $UU' = \lambda^2 P_U$, (2) is a ordinary linear model. Let $e = \Lambda_r\tilde{\xi_1} + \tilde{\varepsilon_1}$, where $\alpha = \lambda^2\sigma_1^2 + \sigma^2$. By Least Square Method, we have

$$\hat{\alpha} = \frac{(P_1y)'(I - P_{P_1X})(P_1y)}{r - (I - P_{P_1X})}$$

$$= \frac{y'(P_U - P_{P_UX})y}{r(P_U - P_{P_UX})},$$

thus analyzing $\hat{\sigma}^2(2)$ about σ^2 in the model of (3), we get an estimation of σ_1^2,

$$\hat{\sigma_1}^2(1) = \frac{\hat{\alpha} - \hat{\sigma}^2(2)}{\lambda^2}.$$

Theorem 4. *In the model of (1), new estimation is better than ANOVAE under Stein function $L(\hat{\Sigma}, \Sigma) = tr(\hat{\Sigma} - \Sigma)^2$ when the design matrix U has only one non-zero singular value.*

Proof. In the model of (1), the covariance matrix of y is

$$\Sigma = UU'\sigma_1^2 + I\sigma^2$$

$$= \lambda^2\sigma_1^2P_U + (P_U + Q_U)\sigma^2$$

$$= \alpha P_U + \sigma^2 Q_U,$$

then Stein function [6] is

$$L(\hat{\Sigma}, \Sigma) = tr((\hat{\alpha} - \alpha)P_U + (\hat{\sigma}^2 - \sigma^2)Q_U)^2$$

$$(\hat{\alpha} - \alpha)^2 tr(P_U) + (\hat{\sigma}^2 - \sigma^2)^2 tr(Q_U).$$

Its relative risk function [7] is

$$R = tr(P_U)E(\hat{\alpha} - \alpha)^2 + tr(Q_U)E(\hat{\sigma}^2 - \sigma^2)^2,$$

it is readily to get the risk function of new estimation is

$$R_1 = tr(P_U)E(\hat{\alpha} - \alpha)^2 + tr(Q_U)E(\hat{\sigma}^2(2) - \sigma^2)^2$$

while the risk function of ANOVAE in original model is

$$R_A = tr(P_U)E(\hat{\alpha} - \alpha)^2 + tr(Q_U)E(\hat{\sigma}^2(A) - \sigma^2)^2$$

where $\tilde{\alpha} = \lambda^2\hat{\sigma_1}^2(A) + \hat{\sigma}^2(A)$. Since $\hat{\alpha}$ is UMVUE of α and $\tilde{\alpha}$ is UE of α, then

$$E(\hat{\alpha} - \alpha)^2 \le E(\tilde{\alpha} - \alpha)^2.$$

Hence this new estimation is better than ANOVAE.

IV. Conclusions

In statistics parameter theory [8], people pay more and more attention to linear mixed models. Parameter estimation is front-burner issue at present. An estimation was offered and its properties were proved in this paper. But for parameter estimation of linear model,we can make a further research from the following several aspects:

(1)People can conduct a further study about the contacts between fixed effects and random effects;

(2)The connections among several defined estimations should be further investigated in the future;

(3)Try to find another better estimation such that it can overcome more defects;

(4)An example should be given to illustrate the theoretical results of new estimation by computer calculations.

References

[1] W.L. Xu. *A estimation of variance component in linear mixed model (J)*. Applied probability and statistics, 2009, 25(3), pp.301-308.

[2] S.G. Wang, J.H. Shi, S.J. Yin. *Linear model introduction (M)*.Beijing science press. 2004.

[3] Y.H. Fan, S.G. Wang. *The improvement about ANOVAE of variance component in linear mixed model (J)*. Applied mathematics A journal of Chinese universities, 2007, 22(1), pp.67-73.

[4] M.X. Wu, S.G. Wang. *The optimal estimation about fixed effect and variance component simultaneously (J)*. Chinese science ser.A, 2004, 15(3):3732384.

[5] K. Tatsuga. *Estimation of variance components in mixed linear models (J)*. Journal of multivatiate analysis, 1995, 53:2102236.

[6] L.R. Lamotte. *One non-negative quadratic unbiased estimation of variance components (J)*. Journal of the american statistical association, 1973, 68, pp.728-730.

[7] J.H. Shi, S.G. Wang. *A non-negative estimation of variance component (J)*. Chinese journal of engineering mathematics, 2004, 21(4):6232627.

[8] X.R. Chen. *Statistics introduction*. Beijing Science Press, 1981, pp.104-108.