A Review on Mixed Models

Zaixing Li

1China University of Mining and Technology (Beijing), Beijing, PR China
2State Key Laboratory of Coal Resource and Safe Mining (CUMT), Beijing, PR China

Mixed model/mixed modeling [1,2] is an important area/tool in statistics. It includes fixed effects and random effects. In fact, random effects (mixed) models were introduced by Fisher [3] where the correlations of trait values between relatives were studied. One-way analysis of variance (ANOVA) model and two way ANOVA are two ordinary and widely-used mixed models. Now two kinds of mixed models are mainly mentioned in literatures. One is to model clustered data/repeated data/longitudinal data [4] where the response may be divided into independent sub-vectors and the covariance matrix of random effects is very general, the other is similar to the two-way ANOVA model and some is to act as a representation tool for the nonparametric function [5,6], where the response may not be divided into independent sub-vectors and the covariance matrix of random effects is usually with special structures.

For different mean structures, mixed models usually include the following: linear mixed models [7,8]. Nonlinear mixed models NLMM [9,10]. Semi parametric mixed models SMM [10], varying coefficient mixed models VCMM [11]. Generalized linear mixed models GLMM [12], generalized additive mixed models GAMM [6], generalized varying coefficient mixed models GVCMM [13].

Statistical inference (estimation/prediction and hypothesis testing) of mixed models is the main topic in this area.

As for the estimation/prediction of mixed models, Henderson et al. [7] is the earliest literature to the best of my knowledge where LMM is considered with the fixed effects estimated and the random effects predicted. Laird and Ware [8] developed the EM algorithm to estimate the fixed effect and the covariance matrix of random effects in the framework of LMM for longitudinal data with normality assumptions. There are many other literatures about estimation of LMM, for instance [14-17]. Besides, NLMM is also estimated by many authors such as Lin [12], Nguyen and Mentr [18], Li [19], Lin and Zhang [6] considered GAMM and Zhang [13] studied GVCMM.

Most literatures about mixed models focus on the hypothesis testing especially for the existence of random effects or their sub-vectors. The testing problem is equivalent to testing whether the corresponding (co)variances of random effects are zero or not since the mean of random effects is zero. Since the true values are on the boundary of the parametric space, it is a nonstandard testing problem and no Wilks phenomenon holds [20]. It is of interest and challenge. There are two kinds of literatures: one is under parametric distributional assumptions and Monte Carlo (MC) method is usually used, the other is distribution-free and some are tractable in the sense that the critical values do not resort to MC method.

Under the normality distributions about random effects and random errors, LMM is studied by many authors. For instance, Stram and Lee [21] and Giampoli and Singer [22] considered likelihood ratio tests (LRTs) according to Self and Liang [20] and Vu and Zhou [23] respectively; Crainiceanui and Ruppert [24] and Greven et al. [25] developed some algorithms for this nonstandard testing problem; Saville and Herring [26] applied Bayes factors in LMM. For other mixed models, Russo et al. [27] considered variance components testing in NLMM with elliptical distributions by score-type test SST [28]. Besides, Zhang and Lin [29] examined GLMM with normally distributed random effects by the adapted LRT based on the theory of Self and Liang [20] and SST based on Silvapulle and Silvapulle [28]. Sinha [30] also considered the existence of random effects in GLMM where the responses are in the exponential family by a one-sided score test based on SST.

For the distribution-free tests for random effects, some main publications are as follows: Drikvandi et al. [31] and Li and Zhu [32] proposed the trace-based tests TDV and Tmd for LMM respectively; Nobre et al. [33] developed a tractable U-test. Li and Zhu [10] proposed a difference-based test for the existence of random effects TmD in SMM. Li et al. [34] developed two distribution-free and easily tractable tests based on the quasi-likelihood for VCMM. Li et al. [35] studied ANOVA-type LMM and Li [36] investigated the existence of any sub-vector of random effects in NLMM.

Moreover, the topics in mixed models include variable selection and high dimensional problems, which is of interest, too. For examples, Fan and Li [34] and Chen et al. [37].

References

*Corresponding author: Li Z, China University of Mining and Technology (Beijing), Beijing, PR China. Tel: 0516-83592326; E-mail: lxgcaes@126.com

Received May 19, 2017; Accepted May 27, 2017; Published May 31, 2017


Copyright: © 2017 Li Z. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.