Journal of Multivariate Analysis 122 (2013) 377–392

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Constrained empirical Bayes estimator and its uncertainty in normal linear mixed models

Tatsuya Kubokawa

Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

ARTICLE INFO

Article history:
Received 19 June 2012
Available online 23 August 2013

AMS subject classifications:
62F12
62F15
62F40
62J07

Keywords:
Benchmarking
Best linear unbiased predictor
Constrained Bayes
Empirical Bayes
Linear mixed model
Mean squared error
Parametric bootstrap
Second-order approximation
Small area estimation

ABSTRACT

The empirical Bayes (EB) estimator or empirical best linear unbiased predictor (EBLUP) in the linear mixed model (LMM) is useful for the small area estimation in the sense of increasing the precision of estimation of small area means. However, one potential difficulty of EB is that when aggregated, the overall estimate for a larger geographical area may be quite different from the corresponding direct estimate like the overall sample mean. One way to solve this problem is the benchmarking approach, and the constrained EB (CEB) is a feasible solution which satisfies the constraints that the aggregated mean and variance are identical to the requested values of mean and variance. An interesting query is whether CEB may have a larger estimation error than EB. In this paper, we address this issue by deriving asymptotic approximations of MSE of CEB. Also, we provide asymptotic unbiased estimators for MSE of CEB based on the parametric bootstrap method, and establish their second-order justification. Finally, the performance of the suggested MSE estimators is numerically investigated.

1. Introduction

The linear mixed models (LMMs) and the model-based estimates including empirical Bayes estimator (EB) or empirical best linear unbiased predictor (EBLUP) have been recognized useful in small area estimation. The typical models used for the small area estimation are the Fay–Herriot model and the nested error regression model (NERM), and the usefulness of EB is illustrated by Fay and Herriot [7] and Battese, Harter and Fuller [1]. For a good review and account on this topic, see [9,14].

One potential difficulty of EB is that when aggregated, the overall estimate for a larger geographical area may be quite different from the corresponding direct estimate like the overall sample mean. One way to solve this problem is the benchmarking approach, which modifies EB so that one gets the same aggregate mean and/or variance for the larger geographical area. Ghosh [8] suggested the constrained Bayes estimator or the constrained EB (CEB) which satisfy the constraints that the aggregated mean and variance are identical to the mean and variance of the posterior distribution, and Datta, Ghosh, Steorts and Maples [4] gave some extensions. Since the sample variance of EB is smaller than the posterior variance, CEB modifies EB so that its sample variance is identical to the posterior variance. However, the usefulness and purpose of EB is that EB gives stable estimates with higher precision of estimation. We then have a concern whether the constrained EB may be against this purpose. Thus, it is quite interesting and important to assess the mean squared error (MSE) of CEB. In this paper, we address this issue for CEB in the general normal linear mixed model.

E-mail address: tatsuya@e.u-tokyo.ac.jp.

0047-259X/$ – see front matter
http://dx.doi.org/10.1016/j.jmva.2013.08.012

In Section 2, we consider the general constraint on the mean and variance of estimators and derive the general constrained estimators including CEB, which is an extension of Ghosh [8] and Datta et al. [4]. In Section 3, we derive the asymptotic approximations of MSE of CEB. When the variance constraint is the posterior variance, it is shown that MSE of CEB is larger than MSE of EB in the first order approximation. To modify this property, we suggest some modification of the variance constraint. We also provide an asymptotically unbiased estimator of MSE of CEB based on the parametric bootstrap method, and establish the second-order justification. In Section 4, we investigate the performance of MSE of CEB and the MSE estimators. Section 5 gives a concluding remark. The proofs of the asymptotic approximations are given in the Appendix.

2. Constrained Bayes estimators

2.1. Linear mixed model and mean–variance constraints

Consider the general linear mixed model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{v} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\boldsymbol{y}$ is an $N \times 1$ observation vector of the response variable, $\boldsymbol{X}$ and $\boldsymbol{Z}$ are $N \times p$ and $N \times M$ matrices, respectively, of the explanatory variables, $\boldsymbol{\beta}$ is a $p \times 1$ unknown vector of the regression coefficients, $\boldsymbol{v}$ is an $M \times 1$ vector of the random effects, and $\boldsymbol{\epsilon}$ is an $N \times 1$ vector of the random errors. Here, $\boldsymbol{v}$ and $\boldsymbol{\epsilon}$ are mutually independently distributed as $\boldsymbol{v} \sim \mathcal{N}_M(\boldsymbol{0}, \boldsymbol{Q})$ and $\boldsymbol{\epsilon} \sim \mathcal{N}_N(\boldsymbol{0}, \boldsymbol{R})$ where $\boldsymbol{Q}$ and $\boldsymbol{R}$ are positive definite matrices. Then, $\boldsymbol{y}$ has a marginal distribution $\mathcal{N}_N(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$ for $\boldsymbol{\Sigma} = \boldsymbol{R}+\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'$. It is assumed that $\boldsymbol{Q}$ and $\boldsymbol{R}$ are functions of unknown parameters $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_q)'$, namely, $\boldsymbol{Q} = \boldsymbol{Q}(\boldsymbol{\psi})$, $\boldsymbol{R} = \boldsymbol{R}(\boldsymbol{\psi})$ and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\psi})$. Denote unknown parameters by $\boldsymbol{\omega} = (\boldsymbol{\psi}, \boldsymbol{\beta})$.

Let $\boldsymbol{\theta} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{v}$ and let $\widehat{\boldsymbol{\theta}} = \widehat{\boldsymbol{\theta}}(\boldsymbol{y})$ be a predictor of $\boldsymbol{\theta}$. Suppose that $\widehat{\boldsymbol{\theta}}$ is evaluated relative to the quadratic loss function $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|^2_{\boldsymbol{\Omega}} = (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ for known positive definite matrix $\boldsymbol{\Omega}$. In the Bayesian framework, $\boldsymbol{\theta}$ has a prior distribution $\mathcal{N}_N(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}')$, and the posterior distribution of $\boldsymbol{\theta}$ given $\boldsymbol{y}$ is

$$\boldsymbol{\theta}|\boldsymbol{y} \sim \mathcal{N}_N\big(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}), \boldsymbol{Z}(\boldsymbol{Q}^{-1} + \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\big). \tag{2.2}$$

Since the Bayes estimator is the posterior mean $E[\boldsymbol{\theta}|\boldsymbol{y}]$ under the squared error loss, the Bayes estimator is

$$\widehat{\boldsymbol{\theta}}^B = \widehat{\boldsymbol{\theta}}^B(\boldsymbol{\omega}) = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}). \tag{2.3}$$

The covariance matrix of $\boldsymbol{\theta}|\boldsymbol{y}$ given in (2.2) is rewritten as $\boldsymbol{Z}(\boldsymbol{Q}^{-1} + \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}' = \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}' - \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'$, which is further expressed as $(\boldsymbol{R}^{-1} + (\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}')^{-1})^{-1}$ when $\boldsymbol{Z}$ is of full rank.

In this paper, we treat general constraints on a linear combination and residual variance of estimator $\widehat{\boldsymbol{\theta}}$. Let $\boldsymbol{W}$ be an $N \times L$ known matrix. The constraints are described by

$$(\mathbf{C}) \begin{cases} (\mathrm{C1})\ (\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}} = \boldsymbol{t}_1(\boldsymbol{y}) \quad \text{for } L\text{-variate function } \boldsymbol{t}_1(\boldsymbol{y}), \\ (\mathrm{C2})\ \|\widehat{\boldsymbol{\theta}} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}\|^2_{\boldsymbol{\Omega}} = t_2(\boldsymbol{y}) \quad \text{for function } t_2(\boldsymbol{y}). \end{cases} \tag{2.4}$$

The estimator $(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}$ is the weighted least squares estimator of $\boldsymbol{\xi}$ under the loss $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{W}\boldsymbol{\xi}\|^2_{\boldsymbol{\Omega}}$, and $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}\|^2_{\boldsymbol{\Omega}}$ is the residual variance. Since $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}\|^2_{\boldsymbol{\Omega}} = \widehat{\boldsymbol{\theta}}'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}$ for

$$\boldsymbol{P}_{\boldsymbol{\Omega}} = \boldsymbol{\Omega} - \boldsymbol{\Omega}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega},$$

the constraint (C2) is expressed as $\widehat{\boldsymbol{\theta}}'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}} = t_2(\boldsymbol{y})$. Thus, we call (C1) and (C2) the mean and variance constraints, respectively, in this paper. We give examples of the constraints (C1) and (C2) below.

The constraint (C1) was dealt by Datta et al. [4] who derived a constrained Bayes estimator $\widehat{\boldsymbol{\theta}}^{CB}$ satisfying $(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^{CB} = \boldsymbol{t}_1(\boldsymbol{y})$. A typical example of $\boldsymbol{t}_1(\boldsymbol{y})$ is

$$\boldsymbol{t}_1(\boldsymbol{y}) = (\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{y}. \tag{2.5}$$

As explained in examples given below, this corresponds to the constraint that the sampling mean of estimators is equal to the overall sample mean based on $\boldsymbol{y}$ in specific situations. For the Bayes estimator $\widehat{\boldsymbol{\theta}}^B$ given in (2.3), it is seen that

$$(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^B = (\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{y} - (\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}),$$

which is not equal to $(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{y}$. Thus, the Bayes estimator does not satisfy the mean constraint (C1). Another example is

$$\boldsymbol{t}_1(\boldsymbol{y}) = (\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^B, \tag{2.6}$$