Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 52 (2008) 1347 – 1361
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Flexible modelling of random effects in linear mixed models—A Bayesian approach

Remus K.W. Ho[a,*], Inchi Hu[b]

[a]*Department of Statistics, Chinese University of Hong Kong, Shatin, Hong Kong*
[b]*Department of Information and Systems Management, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong*

Received 5 February 2006; received in revised form 21 June 2007; accepted 11 September 2007
Available online 18 September 2007

Abstract

Flexible modelling of random effects in linear mixed models has attracted some attention recently. In this paper, we propose the use of finite Gaussian mixtures as in Verbeke and Lesaffre [A linear mixed model with heterogeneity in the random-effects population, J. Amu. Statist. Assoc. **91**, 217–221]. We adopt a fully Bayesian hierarchical framework that allows simultaneous estimation of the number of mixture components together with other model parameters. The technique employed is the Reversible Jump MCMC algorithm (Richardson and Green [On Bayesian Analysis of Mixtures with an Unknown Number of Components (with discussion). J. Roy. Statist. Soc. Ser. B **59**, 731–792]). This approach has the advantage of producing a direct comparison of different mixture models through posterior probabilities from a single run of the MCMC algorithm. Moreover, the Bayesian setting allows us to integrate over different mixture models to obtain a more robust density estimate of the random effects. We focus on linear mixed models with a random intercept and a random slope. Numerical results on simulated data sets and a real data set are provided to demonstrate the usefulness of the proposed method.

Keywords: Bayesian; Hierarchical model; Linear mixed model; Mixture models; Random effects; RJMCMC algorithm

1. Introduction

In longitudinal studies, measurements are obtained repeatedly from individual subjects. We are interested in effects that, we believe, are common to all individuals and also effects that are different among individuals. A commonly used model to capture these among-subject variation is mixed effects model (Laird and Ware, 1982). Specifically, we assume that the data y_{ij}, for $i = 1, \dots, c$ and $j = 1, \dots, n_i$, where c is the number of subjects and n_i is the number of measurements of the ith subject, are independently drawn from densities $f(y_{ij}; \boldsymbol{\alpha}_i, \boldsymbol{\beta})$. The unknown random parameters $\boldsymbol{\alpha}_i$ represent the effects of some covariates which vary among subjects and $\boldsymbol{\beta}$ the effects of some covariates that are the same across subjects.

Traditionally, the $\boldsymbol{\alpha}_i$'s are assumed to follow a multivariate Gaussian distribution for mathematical convenience. Although some studies suggest that inference on fixed effects may be robust to non-normality of random effects (Butler and Louis, 1992; Verbeke and Lesaffre, 1997), there are also findings of inconsistencies in fixed and random effects

* Corresponding author.
E-mail addresses: kwho@sta.cuhk.edu.hk (R.K.W. Ho), imichu@ust.hk (I. Hu).

0167-9473/$ - see front matter
doi:10.1016/j.csda.2007.09.005

estimations under misspecification of random effects distributions (Neuhaus et al., 1992). Moreover, modelling the random effects with less restrictive distributional assumptions may provide important insights. A skewed or even multimodal random effect may indicate exclusion of important factors and suggest improvements in model settings. Therefore, substantial efforts have been made to introduce nonparametric or semi-parametric estimations of the random effects. Examples include the discrete nonparametric MLE (Laird, 1978; Lindsay, 1983), the smooth nonparametric MLE (Magder and Zeger, 1996), predictive recursive estimation (Tao et al., 1999), mixture of normals via EM algorithm (Verbeke and Lesaffre, 1996), etc. More recently, Zhang and Davidian (2001) proposed semi-parametric estimation using a class of densities introduced in Gallant and Nychka (1987). Ghidey et al. (2004) used a penalised Gaussian mixture to model the random effects. In their approach, they did not estimate the number of mixture components.

For most of the existing methods, the complexity of the model for the random effects is controlled by some tuning parameters, for example, the penalty coefficient λ in Ghidey et al. (2004), the tuning parameter K in Zhang and Davidian (2001) and the number of mixture components in Verbeke and Lesaffre (1996). Usually, criteria like Akaike Information Criterion (AIC) or Schwartz Information Criterion (BIC) are used for such decisions. Although an "optimal" model under specific criteria can be chosen from several candidates, the comparisons often lack intuitive meanings. Sole use of the "optimal" model may also lead to the concern of model uncertainties.

In this paper, we propose an approach that has advantage over existing methods in these aspects. We focus on linear mixed models under the fully Bayesian hierarchical framework. Similar to Verbeke and Lesaffre (1996), we model the random effects through finite Gaussian mixtures. However, we do not fix the number of mixture components but estimate that along with other model parameters. This is made possible by adopting the Reversible Jump MCMC algorithm developed by Richardson and Green (1997) and Green (1995). The same idea has been developed in many areas such as the analysis of times of coal mining disasters (Green, 1995), disease mapping (Green and Richardson, 2002), hidden Markov models (Robert et al., 2000) and measurement error models (Richardson et al., 2002).

This approach allows us to perform estimations for mixture models with different number of components in a single MCMC run. Moreover, the fully Bayesian setting provides us posterior probability estimates of the number of mixture components that enable us to make coherent and intuitive comparisons of different mixture models. Besides picking the mixture model with highest posterior probability, we can also integrate over all mixture models to obtain a density estimate of the random effects that is more robust to model uncertainties.

Watier et al. (1998) apply the algorithm of Richardson and Green (1997) to random effects models. Unfortunately, the one-dimensional algorithm can only be applied to models with a random intercept only. In many applications such as repeated measurement problems, it is natural to consider models with random intercepts and random slopes. Therefore an extension of the RJMCMC algorithm to multivariate setting is needed and it turns out to be quite challenging. A general multivariate extentsion can be found in Ho (2005). We remark that another possible approach to fully Bayesian analysis of the problem is the nonparametric Bayesian approach (e.g. Van Der Merwe and Pretorius, 2003) making use of Dirichlet process priors (e.g. Escobar, 1994; MacEachern, 1994; West, 1992).

In this paper, we provide a self-contained treatment of the RJMCMC algorithm for linear mixed models with a random intercept and a random slope. After a brief introduction of our method, we describe the Bayesian hierarchical model setting in Section 2. In Section 3, we give some details of the RJMCMC algorithm for linear mixed models with a random intercept and a random slope. In Sections 4 and 5, we perform studies on simulated data and real data, respectively. We end by giving some concluding remarks in Section 6.

2. Bayesian hierarchical model

2.1. Gaussian mixture random effects

We consider the following linear mixed effects model:

$$\mathbf{y}_i = \mathbf{Z}_i \boldsymbol{\alpha}_i + \mathbf{X}_i \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i \quad (i = 1, \ldots, c), \tag{1}$$

where $\mathbf{y}_i$ is an $n_i \times 1$ vector of responses from the ith subject , $\boldsymbol{\beta}$ is a $q \times 1$ vector of fixed effects, $\mathbf{Z}_i$ and $\mathbf{X}_i$ are covariate matrices and $\boldsymbol{\varepsilon}_i$ is an $n_i \times 1$ vector of errors following $\mathrm{N}(0, \sigma^2\mathbf{I})$. Random effects of the ith subject are modelled by a