Journal of Multivariate Analysis 145 (2016) 148–161

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Approximate uniform shrinkage prior for a multivariate generalized linear mixed model

Hsiang-Chun Chen [a,*], Thomas E. Wehrly [b]

[a] Department of Biostatistics, University of Texas MD Anderson Cancer Center, Houston, TX, USA

[b] Department of Statistics, Texas A&M University, College Station, TX, USA

ARTICLE INFO

Article history:
Received 5 August 2014
Available online 29 December 2015

Keywords:
Bayesian inference
Multivariate generalized linear model
Noninformative prior distribution
Uniform shrinkage prior

ABSTRACT

Multivariate generalized linear mixed models (MGLMM) are used for jointly modeling the clustered mixed outcomes obtained when there are two or more responses repeatedly measured on each individual in scientific studies. Bayesian methods are widely used techniques for analyzing MGLMM. The need for noninformative priors arises when there is insufficient prior information on the model parameters. The main aim of the present study is to propose an approximate uniform shrinkage prior for the random effect variance components in the Bayesian analysis for the MGLMM. This prior is an extension of the approximate uniform shrinkage prior proposed by Natarajan and Kass (2000). This prior is easy to apply and is shown to possess several nice properties. The use of the approximate uniform shrinkage prior is illustrated in terms of both a simulation study and osteoarthritis data.

1. Introduction

Clustered mixed outcomes arise in scientific studies such as longitudinal trials when more than one response is repeatedly measured on each individual. Methods have been proposed for modeling the clustered mixed outcomes. The multivariate generalized linear mixed model (MGLMM) is one of the most widely used models for accommodating these measurements when they are assumed to independently follow distributions in the exponential family, depending on fixed effects and subject-specific correlated random effects [3,10,9,15,1]. Approaches such as the adaptive Gauss–Hermite quadrature, Monte Carlo EM algorithm, generalized estimating equations approach, and penalized quasi-likelihood (PQL) have been developed for maximum likelihood estimation in MGLMM.

Since generalized linear mixed models can be viewed as hierarchical models containing two stages, a Bayesian approach [23,22,7] has been widely used in estimating the joint posterior distributions of the fixed-effect parameters and the variance components of the random effects. Several assumptions for the prior distribution on the fixed effect parameters and the variance components of random effects have been studied. The standard noninformative prior, or Jeffreys prior [21], is one of the most widely used assumption in Bayesian approach. The drawback of using the Jeffreys prior is that it may lead to an improper joint posterior distribution for fixed effects and variance components of the random effect [12,17].

For univariate generalized linear mixed model, Natarajan and Kass [16] introduced the approximate uniform shrinkage prior as an alternative prior for the variance component of the random effects. The main idea of the approximate uniform shrinkage prior is that the weight of the prior mean used in the approximate shrinkage estimate is assumed to be componentwise uniformly distributed. Using the transformation theorem, we can then find the distribution of the variance structure

* Corresponding author.
E-mail address: hchen8@mdanderson.org (H.-C. Chen).

http://dx.doi.org/10.1016/j.jmva.2015.12.004
0047-259X/

of random effect. This prior has several desirable properties. Kass and Natarajan [13] also proposed a default conjugate prior for variance component. This inverse Wishart prior uses 1 as its degree of freedom, and an estimate of variance component obtained from approximate shrinkage estimate and GLM model as its scale matrix. When the clustered mixed outcomes are considered, MGLMM is applied and the random effects are assumed to follow a multivariate normal distribution. Bayesian analysis can also be applied [5,6,18]. The inverse Wishart distribution is one of the most widely used prior for the covariance matrix of the random effects since the inverse Wishart distribution is the conjugate prior of the multivariate normal distribution. However, the estimation is very sensitive to the choice of scale matrix in the prior distribution. Thus, the need of noninformative priors arises when the prior information on the model parameters is insufficient.

The rest of this study is organized as follows. Section 2 briefly introduces Bayesian methods for univariate and multivariate generalized linear mixed model. In Section 3, approximate uniform shrinkage prior for multivariate generalized linear mixed model is derived. Model specification examples are also provided in this section. Section 4 presents properties of the approximate uniform shrinkage prior distribution and its corresponding posterior distributions. Section 5 explains how the posterior simulation is performed. In Section 6, evaluation of the performance of the approximate uniform shrinkage prior in simulation study is described. An example is provided to illustrate the application of the approximate uniform shrinkage prior in Section 7. Lastly, Section 8 concludes with implications for future research.

2. Review of Bayesian methods for GLMM and MGLMM

2.1. Generalized linear mixed model

Given data composed of N subjects and T_i repeated measurements within the ith subject. Let y_{it} be the tth univariate measurement on the ith subject. Conditional on a subject-specific random effect b_i, $\{y_{i1}, \ldots, y_{iT_i}\}$ is assumed to independently follow a distribution with density in the exponential family

$$f(y_{it}|b_i) \propto \exp\left\{\frac{y_{it}\theta_{it} - a(\theta_{it})}{\phi}\right\}$$

where the dispersion parameter ϕ is assumed to be known, and θ_{it} is the canonical parameter. Assume the conditional mean is related to the linear form of predictors by the link function:

$$g(\mu_{it}) = x_{it}^T\beta + z_{it}^T b_i = \beta_1 x_{1,it} + \cdots + \beta_p x_{p,it} + b_{i1}z_{1,it} + \cdots + b_{iq}z_{q,it},$$

where $g(\cdot)$ is a monotonic differentiable link function, $x_{it} = (x_{1,it}, \ldots, x_{p,it})^T$ is a vector of covariates, $\beta = (\beta_1, \ldots, \beta_p)^T$ is a vector of fixed effect parameters, and $z_{it} = (z_{1,it}, \ldots, z_{q,it})^T$ is a vector of covariates corresponding to the random effect $b_i = (b_{i1}, \ldots, b_{iq})^T$. The random effect b_i is shared by repeated measurements within the same subject. Assume that b_i has a multivariate normal distribution with mean 0 and covariance matrix D. This model is known as the GLMM [23,2]. However, the maximum likelihood estimates for β and D cannot be simplified or evaluated in closed form. Because of the complexity of the likelihood in a GLMM, several numerical integration methods, such as Gauss–Hermite quadrature and the Bayesian approach, are proposed for analyzing data in GLMM.

The Bayesian approach is a very popular method used in the analysis of GLMM. A GLMM can be thought of as a two-stage hierarchical model. The measurements conditional on given subject-specific random effects are assumed to follow a particular distribution from the exponential family at the first stage, whereas the random effects are assumed to follow a multivariate normal distribution at the second stage. There is a need of the specification of the prior distribution for the fixed effect parameters β and the random effect variance components D. We assume that β and D are independent of each other in this study. When there is no subjective prior information about the fixed effect coefficient, the most widely used noninformative prior assumption for β is the improper uniform distribution, which we used throughout this study. However, various noninformative prior distributions for the variance components of the random effects, D, have been suggested in the previous literature, including Jeffreys prior, a proper conjugate prior and the approximate uniform shrinkage prior.

The standard noninformative prior, or a Jeffreys prior, $\pi(D) \propto |D|^{-\frac{q+1}{2}}$ [21,23] is one of the most widely used prior assumptions in the Bayesian approach. It is obtained by applying Jeffreys rule to the second-stage random effect distribution. The posterior distribution of D corresponding to a Jeffreys prior follows an inverse Wishart distribution with a scale matrix $S = \sum_{i=1}^N b_i b_i^T$ and degrees of freedom N, $IW(N, S)$. The advantage of Jeffreys prior is that the posterior distribution is specified and easy to implement, however the disadvantage is that it may lead to an improper joint posterior distribution for β and D [12,17].

Another popular choice for prior distributions is a proper conjugate prior. The inverse Wishart distribution with a scale matrix Ψ and degrees of freedom λ, $IW(\lambda, \Psi)$, is a conjugate prior for D. Since a univariate specialization of the inverse Wishart distribution is the inverse gamma distribution, the prior reduces to an inverse gamma distribution when the dimension of D is one. The most popular choice is to set $\lambda = q$ and $\Psi = qD^0$, where D^0 is the prior guess of D [19]. The advantage of this conjugate prior is that it is computationally easy to implement, while the disadvantage is that the estimation results can be very sensitive to the choices of D^0 [16].

Natarajan and Kass [16] introduced the approximate uniform shrinkage prior as an alternative prior for D. It is a generalization of the uniform shrinkage prior proposed by Strawderman [20]. The main idea in the approximate uniform