Journal of Multivariate Analysis 135 (2015) 163–174

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Joint prior distributions for variance parameters in Bayesian analysis of normal hierarchical models

Haydar Demirhan [a], Zeynep Kalaylioglu [b,*]

[a] Department of Statistics, Hacettepe University, Beytepe, 06800 Ankara, Turkey

[b] Department of Statistics, Middle East Technical University, 06800 Ankara, Turkey

ARTICLE INFO

Article history:
Received 12 August 2013
Available online 12 January 2015

AMS subject classification:
62F15

Keywords:
Hierarchical models
Multi-level models
Multivariate log gamma
Random coefficient
Random effect
Variance components
Hyperprior
Hyperparameter
Directional derivative
Sensitivity analysis

ABSTRACT

In random effect models, error variance (stage 1 variance) and scalar random effect variance components (stage 2 variances) are a priori modeled independently. Considering the intrinsic link between the stages 1 and 2 variance components and their interactive effect on the parameter draws in Gibbs sampling, we propose modeling the variances of the two stages a priori jointly in a multivariate fashion. We use random effects linear growth model for illustration and consider multivariate distributions to model the variance components jointly including the recently developed generalized multivariate log gamma (G-MVLG) distribution. We discuss these variance priors as well as the independent variance priors exercised in the literature in different aspects including noninformativeness and propriety of the associated posterior density. We show through an extensive simulation experiment that modeling the variance components of different stages multivariately results in better estimation properties for the response and random effect model parameters compared to independent modeling. We scrutinize the sensitivity of response model coefficient estimates to the parameters of considered noninformative variance priors and find that their full conditional expectations are insensitive to noninformative G-MVLG prior parameters. We apply independent and joint models for analysis of a real dataset and find that multivariate priors for variance components lead to better fitted hierarchical model than the univariate variance priors.

1. Introduction

Hierarchical models are extensively used to model response data obtained from repeated measures designs, longitudinal studies, and multi-level randomized experiments designed in latin square, split plot, balanced/imbalanced block with random effects. Random effects models are currently very popular in a wide variety of fields such as medicine, pharmacology, psychology, regional sciences, agriculture, sports, modeling of traffic accidents, and energy economy [23,14,16,24,17,5,1,8,18,9].

In a hierarchical model, regression coefficients or treatment effects are viewed as random variables. The top stage (stage 1) of a hierarchical model consists of the response model whereas the next stage (stage 2) consists of models for the random coefficients (random effects). For responses obtained from a repeated measures design or a longitudinal study, the random coefficients of a linear hierarchical model account for the heterogeneity among the subjects as well as the correlation among the observations collected from the same subject at different time points. For data obtained from a randomized experiment in which the groups are viewed as a random selection from a population of groups, random effects encapture group specific effects as well as between group variation. For Bayesian analysis of hierarchical models, the hierarchical

* Corresponding author.
E-mail addresses: haydarde@hacettepe.edu.tr (H. Demirhan), kzeynep@metu.edu.tr (Z. Kalaylioglu).

http://dx.doi.org/10.1016/j.jmva.2014.12.013

structure is enlarged to include yet another stage at which the variances of the random coefficients (random effects) are given prior distributions. This stage is the focus of the current article.

As there is usually and unsurprisingly no sufficient prior knowledge regarding what could be the variance of the random coefficient, the user prefers noninformative hyperpriors and let the likelihood dominate the inference on the stage 2 variances. Therefore of interest are the diffuse priors and researchers in the area have been in quest for what could be regarded as default hyperprior for the stage 2 variance parameters or a one to one transformation of them. Of all the diffuse priors considered in the literature, gamma distribution with small shape and scale parameters (denoted thereof by $Ga(\epsilon, \epsilon)$) has been the most commonly used default prior for the inverse of stage 2 variance parameter (an equivalent representation being $Inv - Ga(\epsilon, \epsilon)$ for stage 2 variance) owing its common use to its conjugacy to Normality and resulting computational benefits in softwares such as BUGS that perform Gibbs sampling for posterior inference. A dangerous but often overlooked characterization of diffuse hyperprior distributions based on gamma distribution is that it may result in near or complete improper posteriors. For instance, Natarajan and McCulloch [21] discuss diffuse inverted gamma priors in probit-Normal hierarchical models resulting in improper posterior distributions and inaccurate posterior estimates. Motivated for developing proper hyperprior, Natarajan and Kass [20] proposed for generalized linear mixed models an approximate uniform shrinkage and Jeffreys priors for the unstructured second stage variance matrix and showed that their priors lead to proper posteriors and have better frequentist properties relative to inverse-gamma and Wishart hyperpriors.

More recently Lambert et al. [15] compare effects of 13 different prior settings induced on stage 2 scale parameters of a random effects hierarchical model via a simulation study using WinBUGS. They consider various gamma, Pareto and logistic distributions as prior for stage 2 precision, various uniform distributions as prior for stage 2 variance, its square root and natural logarithm, and various half-normal distributions as prior for square root of stage 2 variance. Not a particular prior setting is identified as best in all scenarios and they note that uniform prior is not a good alternative if a vague prior is intended for stage 2 variance.

Browne and Draper [4] for Bayesian analysis of mixed linear and random effects logistic regression models consider $Inv - Ga(\epsilon, \epsilon)$ and uniform prior on $(0, 1/\epsilon)$ for the stage 2 variance. Their simulation study demonstrates that Bayesian interval inference with these priors face undercoverage problems in mixed linear models when the number of level 2 units of the experimental design is small. Gelman [11] considers traditional $Inv - Ga(\epsilon, \epsilon)$ and Uniform(0,A) hyperpriors and constructs a folded-noncentral-t family of priors as hyperpriors for variance parameters in hierarchical models. Unlike Lambert et al. [15], Gelman [11] suggests the use of uniform prior for a noninformative prior setting. He recommends half-Cauchy (denoted thereof by $HC(0, 1)$) distribution, which is included in folded-noncentral-t family, as weakly informative prior for stage 2 standard deviation and advises not to use the inverse-gamma setting. Of these prior distributions, as indicated in the article $\lim_{A\to\infty}$ Uniform(0,A) yields proper posterior whereas $\lim_{\epsilon\to 0} Ga(\epsilon, \epsilon)$ does not and the posterior inference is sensitive to the choice of ϵ.

Polson and Scott [22] propose to induce half-Cauchy distribution on stage 2 standard deviation and obtained inverted-beta priors for stage 2 variance which ultimately led to the class of hypergeometric inverted-beta distributions resulting in a generalization of the half-Cauchy prior. They qualify the half-Cauchy prior as a sensible default prior for scale parameters in hierarchical models.

One should note, however, that there are two main aspects with these priors that need attention. First, with these priors, variance components of different stages are a priori modeled independently although they are linked as they are the components of the total variation in a response. Second, as presented in Section 3.2, the drawback of these prior structures is that the posterior inference on the response model coefficients in a hierarchical model is highly sensitive to the choice of the parameters of these prior distributions. In this article we a priori model the variance components of different stages jointly by specifying a multivariate prior distribution. Desirable properties of such a joint variance prior density are 1. noninformative, 2. leads to proper posteriors, and 3. change in the parameters of the variance priors do not effect the posterior inference on response model coefficients.

For joint prior modeling, we stack stage 1 and stage 2 variances and induce a multivariate hyperprior distribution. We consider multivariate normal, multivariate skew normal, and generalized multivariate log-gamma distribution as the multivariate hyperprior distribution on natural logarithms of the variance components and investigate their properties based on our prototype hierarchical model.

The rest of the article is organized as follows. In Section 2, we discuss certain modeling aspects concerning the variance components including the informativeness issue, present the proposed joint variance prior setting, and discuss its propriety. Section 3 presents an extensive simulation study in which we investigate and compare sensitivity of the posterior estimators of the proposed joint prior to those in the literature where variances of different stages are a priori modeled independently. In this section, the notion of noninformativeness for a multivariate prior density is furnished and subspace of variance hyperparameters to which the posterior inference is rather insensitive is sought through the directional derivative concept. A data application is presented in Section 4. Finally a discussion on the evaluation of the results and generalization of the proposed approach for further modeling extensions is given in Section 5.

2. Modeling the variance components

We will consider the basic random coefficient model given in (1). Such basic models are also considered in Bayesian literature to study variance components in normal hierarchical models [11,22]. The model is