Journal of Multivariate Analysis 140 (2015) 325–342

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Geometric ergodicity of random scan Gibbs samplers for hierarchical one-way random effects models

Alicia A. Johnson [a,*], Galin L. Jones [b]

[a] Department of Mathematics, Statistics, and Computer Science, Macalester College, United States

[b] School of Statistics, University of Minnesota, United States

ARTICLE INFO

Article history:
Received 14 May 2014
Available online 11 June 2015

AMS subject classifications:
60J05
60J22
65C40
62F15

Keywords:
Markov chain Monte Carlo
Convergence
Gibbs sampling
Geometric ergodicity
One-way random effects

ABSTRACT

We consider two Bayesian hierarchical one-way random effects models and establish geometric ergodicity of the corresponding random scan Gibbs samplers. Geometric ergodicity, along with a moment condition, guarantees a central limit theorem for sample means and quantiles. In addition, it ensures the consistency of various methods for estimating the variance in the asymptotic normal distribution. Thus our results make available the tools for practitioners to be as confident in inferences based on the observations from the random scan Gibbs sampler as they would be with inferences based on random samples from the posterior.

1. Introduction

Suppose that for $i = 1, \ldots, K$

$$\begin{aligned} Y_i|\theta_i, \gamma_i &\overset{ind}{\sim} N(\theta_i, \gamma_i^{-1}) \\ \theta_i|\mu, \lambda_\theta, \lambda_i &\overset{ind}{\sim} \mathrm{N}(\mu, \lambda_\theta^{-1}\lambda_i^{-1}) \\ (\mu, \lambda_\theta, \gamma_1, \ldots, \gamma_K, \lambda_1, \ldots, \lambda_K) &\sim p(\mu, \lambda_\theta, \gamma_1, \ldots, \gamma_K, \lambda_1, \ldots, \lambda_K) \end{aligned} \tag{1}$$

where p is a generic proper prior. Eventually we will consider two distinct choices for the prior p, but we leave the description until Section 2. In both cases, the hierarchy in (1) yields a proper posterior which is intractable in the sense that the posterior quantities, such as expectations or quantiles, required for Bayesian inference are not available in closed form. Thus we will consider the use of Markov chain Monte Carlo (MCMC) methods.

Let y denote all of the data, $\lambda = (\lambda_1, \ldots, \lambda_K)^T$, $\xi = (\theta_1, \ldots, \theta_K, \mu)^T$, and $\gamma = (\gamma_1, \ldots, \gamma_K)^T$. In Section 2 we will see that the specific forms of the posterior full conditional densities $f(\xi|\lambda_\theta, \lambda, \gamma, y)$, $f(\lambda_\theta|\xi, \lambda, \gamma, y)$, $f(\lambda|\xi, \lambda_\theta, \gamma, y)$ and $f(\gamma|\xi, \lambda_\theta, \lambda, y)$ are available and hence it is easy to construct Gibbs samplers to help perform posterior inference. Gibbs samplers can be implemented in either a deterministic scan or a random scan, among other variants [22,13]. Although deterministic

* Corresponding author.
E-mail addresses: ajohns24@macalester.edu (A.A. Johnson), galin@umn.edu (G.L. Jones).

http://dx.doi.org/10.1016/j.jmva.2015.06.002

scan MCMC algorithms are currently popular in the statistics literature, random scan algorithms were some of the first used in MCMC settings [25,9] and remain useful in applications [20,28]. Random scan Gibbs samplers can also be implemented adaptively while the deterministic scan version cannot. In addition, there has been recent interest in the theoretical properties of random scan algorithms [2,13,18,21,33,36].

We will study the random scan Gibbs sampler which is now described. Let $r = (r_1, r_2, r_3, r_4)$ with $r_1 + r_2 + r_3 + r_4 = 1$ and each $r_i > 0$ where we call r the *selection probabilities*. If $(\xi^{(n)}, \lambda_\theta^{(n)}, \lambda^{(n)}, \gamma^{(n)})$ is the current state of the Gibbs sampler, then the $(n+1)$st state is obtained as follows.

Draw $U \sim \text{Uniform}(0, 1)$ and call the realized value u.

If $u \le r_1$, draw $\xi' \sim f(\xi|\lambda_\theta^{(n)}, \lambda^{(n)}, \gamma^{(n)}, y)$ and set

$$(\xi^{(n+1)}, \lambda_\theta^{(n+1)}, \lambda^{(n+1)}, \gamma^{(n+1)}) = (\xi', \lambda_\theta^{(n)}, \lambda^{(n)}, \gamma^{(n)})$$

else if $r_1 < u \le r_1 + r_2$, draw $\lambda_\theta' \sim f(\lambda_\theta|\xi^{(n)}, \lambda^{(n)}, \gamma^{(n)}, y)$ and set

$$(\xi^{(n+1)}, \lambda_\theta^{(n+1)}, \lambda^{(n+1)}, \gamma^{(n+1)}) = (\xi^{(n)}, \lambda_\theta', \lambda^{(n)}, \gamma^{(n)})$$

else if $r_1 + r_2 < u \le r_1 + r_2 + r_3$, draw $\lambda' \sim f(\lambda|\xi^{(n)}, \lambda_\theta^{(n)}, \gamma^{(n)}, y)$ and set

$$(\xi^{(n+1)}, \lambda_\theta^{(n+1)}, \lambda^{(n+1)}, \gamma^{(n+1)}) = (\xi^{(n)}, \lambda_\theta^{(n)}, \lambda', \gamma^{(n)})$$

else if $r_1 + r_2 + r_3 < u \le 1$, draw $\gamma' \sim f(\gamma|\xi^{(n)}, \lambda_\theta^{(n)}, \lambda^{(n)}, y)$ and set

$$(\xi^{(n+1)}, \lambda_\theta^{(n+1)}, \lambda^{(n+1)}, \gamma^{(n+1)}) = (\xi^{(n)}, \lambda_\theta^{(n)}, \lambda^{(n)}, \gamma').$$

Our goal is to investigate the conditions required for the random scan Gibbs sampler to produce reliable simulation results. Specifically, we will investigate conditions under which the Markov chain is geometrically ergodic, which we now define. Let $\mathsf{X} = \mathbb{R}^K \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+^K \times \mathbb{R}_+^K$ and $\mathcal{B}(\mathsf{X})$ denote the Borel sets. Let $P^n : \mathsf{X} \times \mathcal{B}(\mathsf{X}) \to [0, 1]$ denote the n-step Markov kernel for the random scan Gibbs sampler so that if $A \in \mathcal{B}(\mathsf{X})$ and $n \ge 1$

$$P^n((\xi^{(1)}, \lambda_\theta^{(1)}, \lambda^{(1)}, \gamma^{(1)}), A) = \Pr((\xi^{(n+1)}, \lambda_\theta^{(n+1)}, \lambda^{(n+1)}, \gamma^{(n+1)}) \in A \,|\, (\xi^{(1)}, \lambda_\theta^{(1)}, \lambda^{(1)}, \gamma^{(1)})).$$

Let F denote the posterior distribution associated with (1) and $\|\cdot\|$ denote the total variation norm. Then the random scan Gibbs sampler is geometrically ergodic if there exists $M : \mathsf{X} \to [0, \infty)$ and $t \in [0, 1)$ such that for all $\xi, \lambda_\theta, \lambda, \gamma$ and $n = 1, 2, \ldots$

$$\|P^n((\xi, \lambda_\theta, \lambda, \gamma), \cdot) - F(\cdot)\| \le M(\xi, \lambda_\theta, \lambda, \gamma)\, t^n. \tag{2}$$

Geometric ergodicity is a useful stability property for MCMC samplers [16,32] in that it ensures rapid convergence of the Markov chain since $t < 1$, the existence of a central limit theorem (CLT) [1,12,14,30], and consistency of various methods to estimate asymptotically valid Monte Carlo standard errors [5,7,8,15]. To see the connection between (2) and the CLT let $g : \mathsf{X} \to \mathbb{R}$ and f be the posterior density and suppose we want to calculate

$$\mu_g := \int_{\mathsf{X}} g(\xi, \lambda_\theta, \lambda, \gamma) f(\xi, \lambda_\theta, \lambda, \gamma|y)\, d\xi\, d\lambda_\theta\, d\lambda\, d\gamma.$$

Assuming μ_g exists and the Markov chain is irreducible, aperiodic and Harris recurrent (see Meyn and Tweedie [26] for definitions and Section 2 for discussion of these conditions for our two random scan Gibbs samplers), then, as $n \to \infty$,

$$\mu_n := \frac{1}{n}\sum_{i=1}^{n} g(\xi^{(i)}, \lambda_\theta^{(i)}, \lambda^{(i)}, \gamma^{(i)}) \to \mu_g \quad \text{with probability 1.}$$

Of course, μ_n will be much more valuable if we can equip it with an asymptotically valid standard error. If the random scan Gibbs sampler is geometrically ergodic and

$$\int_{\mathsf{X}} g^2(\xi, \lambda_\theta, \lambda, \gamma) f(\xi, \lambda_\theta, \lambda, \gamma|y)\, d\xi\, d\lambda_\theta\, d\lambda\, d\gamma < \infty, \tag{3}$$

then there exists $\delta_g^2 < \infty$ such that, as $n \to \infty$, and for any initial distribution

$$\sqrt{n}(\mu_n - \mu_g) \xrightarrow{d} \mathrm{N}(0, \delta_g^2). \tag{4}$$

The quantity δ_g^2 is complicated [10], but if the Markov chain is geometrically ergodic there are several methods for estimating it consistently [7,12,15]. This then allows construction of asymptotically valid interval estimates of μ_g to describe the precision of μ_n [5] and hence the reliability of the simulation. A similar approach is available for estimating posterior quantiles which, of course, are often useful for constructing posterior credible intervals; see the recent work of Doss et al. [3].