2528-9705

Örgütsel Davranış Araştırmaları Dergisi
Journal Of Organizational Behavior Research
Cilt / Vol.: 3, Sayı / Is.: 2, Yıl/Year: 2018, Sayfa/Pages:56-65

USING AUXILIARY VARIABLE METHOD IN HIERARCHICAL BAYESIAN SPATIOTEMPORAL MODELS, APPLICATION TO AUDITORY fMRI DATA

Nasrin BORUMANDNIA[1], Hamid ALAVI MAJD[2*], Seyed Mohammad TABATABAEE[3]

[1] Urology and Nephrology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran,
[2] Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Science, Tehran, Iran,
[3] Medical Informatics Department, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Science, Tehran, Iran,

* Corresponding Author

ABSTRACT

Sometimes the Bayesian models include a spatial prior which is computationally intractable, because normalizing constants are appeared in the posterior distributions. Computing of normalizing constants is a fundamental computational problem in many Spatiotemporal Bayesian inferences. Functional Magnetic Resonance Imaging (fMRI) data sets are a popular example of huge data sets and big data analytics that their spatial and temporal dependence structures are very complex. Therefore, Spatiotemporal Bayesian inferences for analyzing fMRI data are lionized, but normalizing constant problems often make these models be computationally problematic. In this paper we have focused on the computational schemes for practical Bayesian estimation in binary spatial Ising prior which is one of the problematic priors and is widely used in Spatiotemporal modeling of fMRI data. We investigate the new application of Auxiliary Variable method proposed by Møller for Bayesian estimation in a Hierarchical Spatiotemporal model including an Ising prior, where the posterior involves a normalizing constant. This method avoids approximations such as those in earlier works and also incorporates the normalizing constant and external field problems of Ising prior, simultaneously. We explore the performance of the method on simulated 3D correlated time series. Our approach does a good performance through the simulations. Also, we proceed with real fMRI data set, auditory data from SPM software.

Keywords: *Auxiliary Variable Method, Binary Spatial Ising, Intractable Normalizing Constant.*

INTRODUCTION

Intractable normalizing constants arise in a number of statistical problems, such as image analysis, neural networks, Markov point processes and Markov random fields priors (Møller et al., 2006). In Bayesian approach, sometimes the posterior distributions for the parameters of interest involve an intractable normalizing constant which is also a function of a parameter, so it makes that sampling from Bayesian posterior distributions be problematic. A simple example of a distribution with an intractable normalizing constant is the Ising prior (Melnykov, Maitra, 2010; Fang , Kim, 2012). This prior is used in fMRI Bayesian modeling to consider the spatial correlation among the observation.

Computational aspect of the Bayesian statistical models with intractable normalizing constants is a fundamental computational problem for many statistical models (Gelman, Meng, 1998). Three common approaches that were used for approximating intractable normalizing constants are analytic approximation, numerical integration and Monte Carlo simulation (Evans, Swartz,

Geliş tarihi/Recieved: 23.05.2018 – Kabul tarihi/Accepted: 30.09.2018 – Yayın tarihi/Published: 22.12.2018

1995; DiCiccio et al., 1997). Besag (1991) used a Bayesian hierarchical model which included Ising prior, and they considered an ad hoc procedure for solving the normalizing constant problem (Besag et al., 1991). Heikkinen and Hogmander (1994) used a pseudo likelihood function of easily derivable full conditional distributions for approximating the likelihood term including an unknown constant (Heikkinen and Hogmander, 1994). Heikkinen and Penttinen (1999) tried to find the maximum a posteriori estimate for the interaction function in a Bayesian model, where the normalizing constant in likelihood function is unknown (Heikkinen and Penttinen, 1999). Among the three common approaches, Monte Carlo methods is widely used in statistics (Gelman, Meng, 1998; Green et al., 2002; Dryden, Scarr, 2003; Berthelsen, Møller, 2002).

Monte Carlo techniques, such as importance sampling and some of its variations, such as bridge sampling and Umbrella sampling, can be used to estimate the normalizing constants in distributions. Another Monte Carlo technique that have been used in some papers is path sampling (Gelman, Meng, 1998). The Wang and Landau Algorithm, proposed by Fugao Wang (2001) is another Monte Carlo method designed to perform a non-markovian random walk to build the density of states by quickly visiting all the available energy spectrum (Wang, Landau, 2008). Zhang (2007) used the Wang-Landau scheme to demonstrate the efficiency of simulations via direct computation of the partition function under various macroscopic conditions, such as different temperatures or volumes (Cheng Zhang, 2007).

Thus far, all methods proposed in the literature but one entail approximations that do not vanish asymptotically. Auxiliary Variable method, the only method which was introduced by Moller (2004), avoids approximations such as those in earlier methods. Indeed, this method introduces an Auxiliary Variable x into the Metropolis-Hastings algorithm for the other parameters, so that ratios of normalizing constants will be omitted in Metropolis-Hastings ratio, while the posterior distributions for the other parameters are retained (Møller, 2006). The typical single-subject fMRI experiment runs in this way: A subject in a MRI scanner performs a task in response to a stimulus while three-dimensional images of the subject's brain are captured during the time. The signal measured in fMRI depends on local blood oxygen and is referred to as the blood oxygenation level dependent, BOLD signal (Poldrack et al., 2011). BOLD activity peaks 4 to 6 seconds after neuronal activity and experiences a marked undershoot after 10 to 12 seconds. It returns to baseline after 20 to 30 seconds (Penny et al., 2005). As a result, it is necessary to convolve the input functions by a Hemodynamic Response Function, HRF (Borumandnia et al., 2017). Due to use for pre surgical purposes and for meta-analyses investigations, single-subject scanning is popular in the fMRI experiments (Bowman et al., 2008). An Image is divided into a regular grid of volume elements, called voxels. The BOLD signal is observed at each voxel and at each time point, this leads to an enormous amount of data. These 4-D data sets have a complicated structure of correlations (Borumandnia et al., 2017). Due to complex spatial and temporal correlation structures of fMRI time series, statistical methods play a crucial role in the analysis of fMRI data (Poldrack et al., 2011; Lazar, 2008; Lindquist, 2009). Bayesian approach has received considerable attention for modeling of fMRI data (Bowman et al., 2008; Xia et al., 2009; Woolrich et al., 2004; Smith et al., 2003; Smith and Fahrmeir, 2007; Goldsmith et al., 2014; Genovese, 2000). The size and complexity of data makes computational feasibility be important as well as model efficiency (Zhang et al., 2015) In fact, we are faced with a situation where computational issues drive some of the modeling decisions.

In the fMRI literature, one of the statistical Bayesian models that incorporate spatial among brain responses using Ising prior, have been proposed by Lee and Et al. (2014). They modeled the BOLD response for a single subject with a linear regression model that setting prior distributions on different parameters of the model allows to account for the temporal and spatial dependence among observations. There is a normalizing constant in the posterior quantity given by the Ising which is analytically intractable that writer noted the path sampling (Gelman, 1998) or the Wang-Landau algorithm for solving it (Gelman, Meng, 1998; Landau et al., 2004). In addition, an external field in the Ising model was specified to incorporate anatomical prior information. They used a two-steps procedure proposed by Smith and Fahrmeir (2007) handle the external field in binary spatial Ising prior (Smith and Fahrmeir, 2007). If the writhers used the Auxiliary Variable method, they could find a solution for the problem of normalizing constant as well as the issue of external field. In this way, they could conquer both of those tasks simultaneously. This will leads to simplifying of model estimation that is a notable point in fMRI statistical modeling. Therefore, according to these advantages, we decide to use the Auxiliary Variable method in the process of model estimation. For presenting a new application of the Auxiliary Variable method, we estimate Lee's models by different method in the process of estimation. We used the Auxiliary Variable method in this special model as an example of Hierarchical Bayesian models including of Ising prior that the mentioned method can be used for them, similarly.

The rest of this paper is organized as follows. A brief introduction about the Bayesian model and the Auxiliary Variable method are presented in the statistical methods, section 2. How to generate simulated data and results obtained by applying the method to simulated data and real fMRI data, are listed in the section 3. Finally, section 4 involves additional remarks and discussion.

STATISTICAL METHODS

At first, a brief introduction about the modeling of fMRI data is explained. Following, the priors and posteriors quantities for the model of interest are rewired. At last the Auxiliary Variable Method will be proposed for using in this special model as an example of Hierarchical Bayesian models including of Ising prior which auxiliary variable method can be used for them.

Regression model

A linear regression model for the BOLD response of a given voxel is $y_v = X_v\beta_v + \varepsilon_v$, where X is the T×p covariate matrix and $\beta_v=(\beta_{v1},\ldots,\beta_{vp})$ is a p×1 vector of regression coefficients. The goal of this analysis is detecting neuronal activation in a voxel which corresponds to identifying nonzero β_v. Let γ_v be binary random variable that indicates whether the voxel is activated by a task. That is, the coefficient β_v is equal to zero if $\gamma_v = 0$ and β_v is nonzero if $\gamma_v = 1$. So the model can be rewritten as $y_v = X_v(\gamma_v)\beta_v(\gamma_v) + \varepsilon_v$.

Priors

Referring again that the priors mentioned here have been proposed by Lee et al. (2014). The temporal dependence between observations for a given voxel was taken into account by considering an autoregressive processes on the structure of the error terms in model (3): $\varepsilon_v \sim N_{T_v}(0,\sigma_v^2\Lambda_v)$. The prior for σ^2 and ρ were respectively $\pi(\sigma_v^2)\alpha\frac{1}{\sigma_v^2}$, $\pi(\rho) = \prod_{v=1}^{N}\pi(\rho_v)\ \alpha\ \prod_{v=1}^{N}U(-1<\rho_v<1)$. A binary spatial Ising prior for γ parameter allows to incorporate anatomical prior information with a constant external field, parameter θ_0, as well as spatial interaction between voxels with parameter θ_1: $\pi(\gamma|\theta)\alpha\{\sum_{v=1}^{v}\theta_0\gamma_v + \theta_1 \sum_{v\sim k} w_{v,k} I(\gamma_v = \gamma_k)\}$. Parameter θ_1 is the positive parameter to represent the strength of the interaction between any two neighbor voxels. The term $v\sim k$ means that two voxels v and k are neighbors. It can be assumed a uniform prior on θ_1 and θ_0, that is [min θ_0, max θ_0]×[0, max θ_1]. Also a Zellner's g-prior have been used on regression coefficients (Zellner, 1986).

Posterior Inference

The posterior distribution for parameters of interest, obtained by combining the prior information and the likelihood function via the Bayes theorem, was provided as follow:

$$q(\gamma,\rho,\theta|y)\ \alpha\ \pi(\gamma|\theta)\,\pi(\theta)\,\pi(\rho)\ \prod_{v=1}^{N} \frac{1}{(1+T_v)^{\frac{q_v}{2}}}\frac{1}{\Lambda_v^{\frac{1}{2}}}\Big(\big(y_v - X_v'\hat{\beta}_v\big)'\Lambda_v^{-1}\big(y_v - X_v'\hat{\beta}_v\big)\Big)^{-\frac{T_v}{2}} \qquad (1)$$

According to (1), conditional posteriors can be reached. Details can be find in Lee's paper (Lee et al., 2011). For inference, component-wise Markov Chain Monte Carlo (MCMC) sampling techniques can be used to sample the individual parameters conditional upon the others. The metropolis Hastings ratio for parameters θ is

$$\frac{Z(\theta_{1j},\theta_{0j})\, exp\left\{\theta_{1j}^{*}\sum_{v\sim k} w_{v,k}\, I(\gamma_{v,j}=\gamma_{k,j})\right\} I(0<\theta_{j1}^{*}<\theta_{1\ max})\, I(\theta_{0\ min}<\theta_{j0}^{*}<\theta_{0\ max})}{Z_j\left(\theta_{1j}^{*},\theta_{0j}\right) exp\left\{\theta_{1j}\sum_{v\sim k} w_{v,k}\, I(\gamma_{v,j}=\gamma_{k,j})\right\} I(0<\theta_{j1}<\theta_{max})\, I(\theta_{0\ min}<\theta_{j0}<\theta_{0\ max})}\ \frac{p_{\theta_j}(\theta_j|\gamma,y)}{p_{\theta_j}(\theta_j^{*}|\gamma,y)} \qquad (2)$$

The ratio $\frac{Z_j(\theta_j,\theta_{0j})}{Z_j\left(\theta_j^{*},\theta_{0j}\right)}$ is analytically intractable and should be solved for estimating proses. In the following, we go over the Auxiliary Variable method, then we will use it to find a solution for the intractable normalizing constant.

Auxiliary Variable Method

The Auxiliary Variable method is an efficient Markov chain Monte Carlo approach for distributions with intractable normalizing constant. The idea behind this approach is based on two point: it adds an Auxiliary variable in the Metropolis-Hastings algorithm and chooses the proposal distribution so that the algorithm does not depend upon the unknown normalizing constant (Møller et al., 2006).

To sample from the posterior $\pi(\theta|y)\ \alpha\ \pi(\theta)\pi(y|\theta)$ that the likelihood is $\pi(y|\theta) = \frac{q_\theta(y)}{Z_\theta}$, The Metropolis-Hastings ratio is $H(\theta'|\theta) = \frac{\pi(\theta')q_{\theta'}(y)P(\theta|\theta')}{\pi(\theta)q_\theta(y)P(\theta'|\theta)} / \frac{Z_{\theta'}}{Z_\theta}$ where $P(\theta'|\theta)$ is the proposal density for θ. The normalizing constant Z_θ is not available analytically and an exact computation is not feasible. For solving this problem, Møller introduced an Auxiliary variable x with conditional distribution $g(x|\theta,y)$ and constructed a Metropolis-Hastings chain with target distribution $\pi(\theta,x|y)\alpha\pi_0(\theta)g(x|\theta,y)\frac{q_\theta(y)}{Z_\theta}$. This chain proposes a new state (x',θ') jointly by drawing θ' from $P(\theta'\,|\theta,x)$ and then x' from $f(\,x'|\theta) = \frac{q_{\theta'}(x')}{Z_{\theta'}}$. An appropriate choice for Auxiliary variable density $f(x|\theta,y)$ and proposal density $P(\theta'|\theta)$ cause the algorithm to have

Cilt / Vol.: 3, Sayı / Is.: 2, Yıl / Year: 2018, Sayfa / Pages: 56–65

a good mixing and convergence properties (Møller et al., 2006). One approach is to make $f(x|\theta, y) = \frac{q_{\tilde{\theta}}(x)}{Z_{\tilde{\theta}}}$. Then the problematic normalizing constant cancels in the Metropolis-Hastings ratio $H(\theta', x'|\theta, x) = \frac{g(x'|\theta', y)\, q_{\theta'}(y) q_\theta(x) \pi_0(\theta') P(\theta|\theta', x')}{g(x|\theta, y) q_\theta(y) q_{\theta'}(x')\, \pi_0(\theta) P(\theta'|\theta, x)}$ As mentioned before, normalizing constants problem sometimes occurs for Bayesian hierarchical models, such as Spatio-Temporal model with Ising prior for fMRI time series data. In case of Ising prior, independent normal distributions can be used for proposal densities of θ_0 and θ_1. This choice leads to $\frac{P(\theta|\theta')}{P(\theta'|\theta)} = 1$. Also uniform priors can be assumed on θ_0 and θ_1, where $\theta \in \Theta = [\min \theta_0, \max \theta_0] \times [0, \max \theta_1]$. So Metropolis-Hastings ratio reduces to $H(\theta', x'|\theta, x) = 1[\theta' \in \Theta] \frac{q_{\tilde{\theta}}(x')\, q_{\theta'}(y) q_\theta(x)}{q_{\tilde{\theta}}(x) q_\theta(y) q_{\theta'}(x')}$, That $1[\theta' \in \Theta]$ is an indicator function. In practice, the exact values of min $\theta_0 < 0$, max$\theta_0 > 0$ and max θ_1 have very little influence on the chain, so ranges (-1, +1) for θ_0 and and [0; 1) for θ_1 are quite adequate (Møller et al., 2006; Lee et al., 2014).

RESULTS

Simulation Study

In this section we report the results of a simulation study undertaken to validate the model and estimation procedure based on Auxiliary Variable method. For different values of θ_0 and θ_1 we generated 5 data sets based on 10*10*10 activated-inactivated 3-dimentional images. We performed a posteriori inference on these data sets using the defined model with the Axillary Variable method in the estimation process.

Data generation process is as follow. The auto-regression coefficient ρ_v, is generated from Uniform (-1; 1) for each voxel. Also parameter σ set fix, σ=3 for all of voxels. We consider different scenarios for θ_0 and θ_1 that have been shown in table 1. Given θ_0 and θ_1, we generate a 3-D activated-inactivated lattice cubic of size 10*10*10 from (8) formula. We used a 3-dimensional neighborhood which contains the 6 directly adjacent voxels in 3-dimensional space. Commonly used neighborhood 3-dimentional structures for voxels are 6-neighbors (if voxels share a face), 18-neighbors (if voxels share face and edge) and 26-neighbors (if voxels share face, edge and corner). The weights $w_{v,k}$ were taken to be the same for all voxels.

Simulating from Ising models can be done using a perfect sampling technique or MCMC approach (Novotny, 1999). A Monte Carlo algorithm for a 2-dimensional Ising model is proposed by Gunnar Ingelman (Ingelman, 2009). We have extended their algorithm for 3-dimensional simulation of Ising model. Given parameter γ_v, we simulate a time-series y_v in each voxel v of length 54 from the model. We built a design matrix by convolving a stimulus function for a block design with a Poisson HRF. The baseline level in a human brain and amplitude of activation in response to a stimulus at each voxel v are assume 300 and 10, respectively. Parameter γ_v indicates if β_v is equal to 0 or not. If γ_v is 1, we simulate data from the model with $\beta_v = (300,10)^T$, otherwise data was simulated with $\beta_v = \beta_{v,0} = 300$.

For each simulated data set, the model with the Auxiliary Variable approach was applied to detect the activated-inactivated voxels. We classified a voxel as active one if $\hat{p}(\gamma_v = 1) > 0.8722$

(Raftery, 1996). The estimation results are based on 10000 draws (with 2000 burned) from the posterior using our MCMC algorithm with Auxiliary Variable method.

Sensitivity analysis

We performed a sensitivity analysis to study how different values of θ_0 and θ_1 affect the performance of the model with the Auxiliary Variable approach. In table 1, accuracy and FPR rates are reported for simulated data sets. Accuracy, the percentage of voxels that are correctly identified, and False Positive Rate (FPR), the proportion of active voxels falsely identified relative to all the inactive voxels, averaged over the 5 replicates.

Table 1. Simulated data: accuracy and FPR for different choices of θ_1 and θ_0 parameters in Ising prior. Results are in percentages.

θ_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7
Accuracy	95.92	95.92	97.64	98.38	98.87	99.16	98.16
FPR	1.11	1.46	0.66	0	2.58	10	4.36
θ_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7
Accuracy	95.67	96.20	96.42	96.12	96.13	96.70	97.10
FPR	1.35	1.08	0.53	0.94	0.00	0.00	0.00

The result shows that our approach have had a good performance on simulated data sets. Result in table 1 imply that setting larger values of the parameter θ_1 in the Ising prior distribution for the β's, may moderately impact on the performance and lead to a higher accuracy. The best performance is when the parameter θ_1 is nearly 0.6 in a 3-Dimentional data set. The FPR rates are relatively low for different values of θ_1. In addition, larger values of θ_0 lead to reducing of accuracy. In this case, FPR ratios are almost close to zero. In general, we can say the approach has an acceptable performance; at least 96 percent of accuracy in different situations.

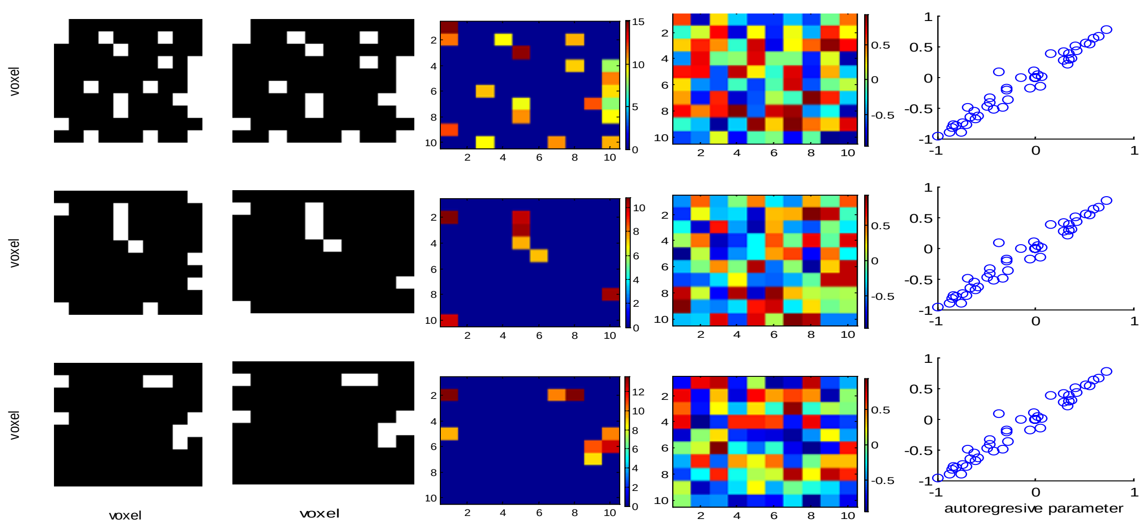

Figure 1: Simulated data with block design, first 3 slices of randomly selected 3-D simulated lattice cubic: True map of the activation indicators γ (first column); Predicted map of the activation indicators γ (second column); The posterior mean map of β (third column); The posterior mean map of ρ (forth column); Scatter plot of posterior mean estimates vs. true values for ρ parameter in the slices.

In figure 1, we have shown first 3 slices of 10 slices of a randomly selected 3-D simulated lattice cubic of our simulated data set with θ_0=0 and θ_1=0.3. In other words, it is a true map of the activation indicators γ for the first 3 slice of 10 slices of this data. The posterior activation map for these slices of dataset have been displayed in second column of figure 1. By comparing the first two columns with each other, it is obviously that the approach does a good performance at detecting the active voxels. In this simulated data set, a small number of active voxels are falsely identified as inactive (0.88 %). Posterior mean map of β and ρ parameters for the slices of dataset have been shown in third and fourth columns, respectively. In the last columns scatter plot of posterior mean estimates vs. true values for ρ parameter in the slices. This plot shows that our approach produces good estimates, close to the true values of ρ parameters.

Real Data Example

- ***Auditory data***

Here we apply our technique to the data set collected by Geraint Rees in Functional Imaging Laboratory (FIL) which is known as the mother of all experiments, available at http://www.fil.ion.ucl.ac.uk/spm/data/auditory/. The experiment was performed on a single subject, under 2 different conditions, rest and auditory stimulation. Auditory stimulation was bi-syllabic words presented binaurally at a rate of 60 per minute. The subject was scanned during 6 blocks, with each block lasting 42 s. 96 acquisitions were made (TR=7s). We discarded the first 12 scans, and we did our analysis on leaving 84 scans. Data set comprises whole brain BOLD/EPI images acquired on a modified 2T Siemens MAGNETOM Vision system. Each acquisition consisted of 64 contiguous slices (64×64×64 3×3×3 mm^3 voxels). The posterior activation maps overlay on the structural MRI images for slices 23 up to 28 of brain, containing some parts of the Temporal lobe, are shown in figure 2. As images show activations appear mainly in the area, which is known to be involved in the perception of auditory cortex.

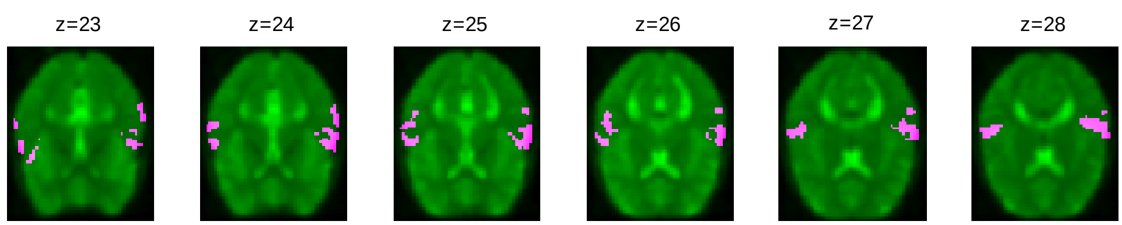

Figure 2: Real fMRI data: A transverse plane of the slice number 23 up to 28 containing the temporal lobe. posterior activation maps for the slices, obtained by assigning value 1 to those voxels with $P(\gamma_v = 1|y) > 0.8722$, and value 0 otherwise.

The results for the other slices of whole brain and the code implementing our methodology are available upon request. Our Matlab code performed 10000 MCMC iterations in 7 hours for real data, on a computer with CPU 3.30 GHz, and 4 GB of RAM.

DISCUSSION

In this paper we used the hierarchical Bayesian Spatio-Temporal model for fMRI data introduced by Lee and et al. (2014). Also we merged the Auxiliary Variable method in estimation proses of the model. They imposed a hyper-prior on the parameter of the Ising model that led to appearing of normalizing constant in posterior inference. A variety of approximate methods and computational schemes have been proposed for spatial Ising priors, but it seems unclear which one is preferred for practical use. The usual approach is to replace normalizing constant by an estimate using Markov Chain Monte Carlo methods, in order to achieve an equilibrium distribution close to real distribution. In addition, the other methods for Monte Carlo approximation of normalizing constants is computationally very demanding, requiring many samples and needed to be repeated for each iteration of the Metropolis-Hastings algorithm. In our work we also reviewed the Auxiliary Variable method for normalizing constant that was introduced by Moller and Et al. (2006). Application of the Auxiliary Variable method to the hierarchical models, with the normalizing constant, had been suggested by Moller et al., (2006). So we demonstrated how to use the Auxiliary Variable method in the Spatio-Temporal model consist of Ising prior, where the normalizing constant is intractable. A considerable advantage of our work is that we did simulation study on 3-dimentional data sets. In real situations, the voxels overlay a three-dimensional lattice, so slice-by-slice analyses are known to have some limitations. In previous works that we looked into, simulation was performed on 2-dimentional data sets. The simulation from 3-dimentional Ising model is not easy to do and it is often ignored. The applications we have presented here show that the method has a good performance in different values of the 3-dimentional Ising's parameters.

Thus far, usual methods for solving the problem of normalizing constant, are based on finding an approximation for it, indeed the Auxiliary Variable method does not. Estimation of normalizing constant ratios requires extensive MCMC runs to estimate before the analysis can start. By using the Auxiliary Variable method, we can remove this estimation, so this method will be somewhat easier to setup among the others. Another motivation for doing present study was that Auxiliary Variable method combines two aspects into a single framework, the problem of normalizing constant and the external field. Our results on real fMRI data have confirmed the result of the previous studies on real data.

ACKNOWLEDGEMENTS

The authors thank the Institute for OpenfMRI and also for My Connectome project for providing the data. We also express our sincere thanks and appreciation to professor Poldrack for his kind help in using of the data.

References

Berthelsen KK, Møller J. A primer on perfect simulation for spatial point processes. Bull Braz Math Soc. 2002;33(3):351–67.

Besag, J., York, J., and Mollie A. Bayesian Image Restoration, with two Applications in Spatial Statistics,". Ann Inst Stat Math. 1991; 43:1{59}.

Borumandnia N, Majd HA, Zayeri F, Baghestani AR, Gohari MR, Tabatabaei SM, Faeghi F. Bayesian 3-Dimensional Spatial Variable Selection Modeling of Voxel- Specific HRFs for Localization in fMRI Time Series Data. Advances and Applications in Statistics. 2017 Dec 1;51(6):397-426.

Borumandnia N, Majd HA, Zayeri F, Baghestani AR, Tabatabaei SM, Faeghi F. A 3- Dimensional Non-Parametric Bayesian Spatiotemporal Model for Brain Activation and Functional Connectivity in fMRI Data. Jp Journal of Biostatistics. 2017 Jun 1;14(1):31-60.

Bowman, F., Ca o, B., Bassett, S., and Kilts C. a Bayesian hierarchical framework for spatial modeling of fMRI data.". Neuroimage. 2008; 39:146–156.

Cheng Zhang JM. Simulation via direct computation of partition functions. NIH Public Access. 2007;76(3 Pt 2).

DiCiccio, T. J., Kass, R. E., Raftery, A. and Wasserman L. Computing Bayes factors by combining simulation and asymptotic approximations. J Amer Stat Assoc. 1997; 92:903–915.

Dryden IL, Scarr MR. Bayesian texture segmentation of weed and crop images using reversible jump Markov chain Monte Carlo methods. Appl Stat. 2003;31–50.

Evans, M. and Swartz T. Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. Stat Sci. 1995; 10:254–72.

Fang Z, Kim I. A Graphical View of Bayesian Variable Selection. 2012;(540):1–80.

Gelman A, Meng X. Simulating Normalizing Constants: From Importance Sampling to Bridge Sampling to Path Sampling. Stat Sci. 1998;13(2):163–85.

Genovese C. a Bayesian time-course model for functional magnetic resonance imaging data.". J Am Stat Assoc. 2000; 95:691–703.

Goldsmith, J., Huang, L., and Crainiceanu CM. Smooth scalar-on-image regression via spatial Bayesian variable selection." J Comput Graph Stat. 2014; 23:46–64.

Green PJ, Richardson S. Hidden Markov Models and Disease Mapping. J Am Stat Assoc. 2002;97(460).

Heikkinen, J. and H ogmander H. Fully Bayesian Approach to Image Restoration with an Application in Biogeography. Appl Stat. 1994; 43:569–82.

Heikkinen, J. and Penttinen A. Bayesian Smoothing in the Estimation of the Pair Potential Function of Gibbs Point Processes. Bernoulli. 1999; 5:1119–36.

Ingelman G. Monte Carlo method and the Ising model. Course: Statistical methods in Physics. 2009.

Landau DP, Tsai S, Exler M, Introduction I. A new approach to Monte Carlo simulations in statistical physics: Wang-Landau sampling. 2004;72(10):1294–302.

Lazar N. A. The statistical analysis of functional MRI data. Stat Biol Heal. 2008.

Lee K. J., Jones G. L., Caffo B. S., & Bassett S. S. Spatial Bayesian Variable Selection Models on Functional Magnetic Resonance Imaging Time-Series Data. Bayesian Anal. 2014; 9(3):699–732.

Lee K. J., Jones G. L., Caffo B. S., & Bassett S. S. Spatial Bayesian Variable Selection Models on Functional Magnetic Resonance Imaging Time-Series Data. Bayesian Anal. 2011;(3):1–32.

Lindquist M a. The Statistical Analysis of fMRI Data. 2009;23(4):439–64.

Melnykov V, Maitra R. High-dimensional Ising model selection. Stat Surv [Internet]. 2010; 4:80–116.

Møller, A. Pettitt, K. Berthelsen and RR. An efficient Markov chain Monte Carlo method for distributions with intractable normalising constants. Tech Rep R-2004-02, Dept Math Sci Aalborg Univ 2004 To Appear Biometrica, 93, 2006.

Novotny MA. Introduction to the Propp-Wilson Method of Exact Sampling for the Ising Model. Comput Simulations Condens Matter Phys XII [Internet]. 1999;0(1):179–84.

Penny WD, Trujillo-Barreto NJ, Friston KJ. Bayesian fMRI time series analysis with spatial priors. Neuroimage. 2005;24(2):350–62.

Poldrack R.A., Mumford J.A., and Nichols T.E. Handbook of functional MRI data analysis. Cambridge University Press; 2011.

Raftery, A. Hypothesis testing and model selection. In Gilks, W., Richardson, S., and Spiegelhalter, D., editors, Markov Chain Monte Carlo in Practice. Chapman & Hall, London. 1996.

Smith, M. and Fahrmeir L. Spatial Bayesian variable selection with application to functional magnetic resonance imaging.", J Am Stat Assoc. 2007;102:417–31.

Smith, M. and Fahrmeir, L. (2007). Spatial Bayesian variable selection with application to functional magnetic resonance imaging." Journal of the American Statistical Association, 102: 417-431.

Smith, M., P utz, B., Auer, D., and Fahrmeirc L. Assessing brain activity through spatial Bayesian variable selection. Neuroimage. 2003; 20:802–15.

Wang F, Landau DP. An efficient, multiple range random walk algorithm to calculate the density of states. 2008; 30602:1–4.

Woolrich, M. W., Jenkinson, M., Brady, J. M., and Smith SM. Fully Bayesian spatio-temporal modeling for fMRI data. IEEE Trans Med Imaging. 2004; 23:213–31.

Xia, J., Liang, F., and Wang YM. FMRI analysis through Bayesian variable selection with a spatial prior. IEEE Int Symp Biomed Imaging. 2009;714–7.

Zellner A. On assessing prior distributions and Bayesian regression analysis with g-prior distributions. Bayesian inference and decision techniques: Essays in Honor of Bruno De Finetti. 1986; 6:233-43.

Zhang L, Guindani M, Vannucci M. Bayesian models for functional magnetic resonance imaging data analysis. Wiley Interdiscip Rev Comput Stat. 2015;7(February):21–41.