RANDOM EFFECTS' DISTRIBUTION ASSUMPTION ON JOINT MIXED MODELLING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

CELAL OĞUZ ÖZDEMIR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
STATISTICS

SEPTEMBER 2018

Approval of the thesis:

RANDOM EFFECTS' DISTRIBUTION ASSUMPTION ON JOINT MIXED MODELLING

submitted by **CELAL OĞUZ ÖZDEMIR** in partial fulfillment of the requirements for the degree of **Master of Science in Statistics Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Ayşen Dener Akkaya _____
Head of Department, **Statistics**

Assoc. Prof. Dr. Zeynep Işıl Kalaylıoğlu _____
Supervisor, **Department of Statistics, METU**

Examining Committee Members:

Assoc. Prof. Dr. Özlem İlk Dağ
Department of Statistics, METU _____

Assoc. Prof. Dr. Zeynep Işıl Kalaylıoğlu
Department of Statistics, METU _____

Assoc. Prof. Dr. Beyza Doğanay Erdoğan
Department of Biostatistics, Ankara University _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: CELAL OĞUZ ÖZDEMIR

Signature :

ABSTRACT

RANDOM EFFECTS' DISTRIBUTION ASSUMPTION ON JOINT MIXED MODELLING

Özdemir, Celal Oğuz
M.S., Department of Statistics
Supervisor : Assoc. Prof. Dr. Zeynep Işıl Kalaylıoğlu

September 2018, 134 pages

Joint mixed model is an appealing approach in medical research where it is critical to estimate the odds of a fatal complication that occurs to a patient given the covariate profile such as a risk factor observed over time. For this kind of estimation, joint mixed model is used. In the standard Bayesian analysis of the model, the error variance and random effects' variance-covariance matrix are apriori modeled independently with Inverse-Gamma and Inverse-Wishart distributions respectively. Recently however, it is shown that joint apriori modeling via Generalized Multivariate Log-Gamma (G-MVLG) distribution is more efficient than the standard Bayesian analysis for these variance components. Our current aim is to inverstigate the robustness of G-MVLG based and standard analysis to random effects' distributions. Bivariate Gamma, Bivariate Skew-Normal, Normal distribution and their mixture distributions were considered for the true distribution of random effects. Results show that the G-MVLG approach is robust to the underlying true distribution of random effects when the sample size is sufficiently large. For small samples, a robust approach. Simula-

tions and real data study show that DPP for the random effects distributions is less biased and more efficient.

Keywords: Joint Mixed Model, Random Effects, Dirichlet Process Prior, Generalized Multivariate Log-Gamma Distribution, Bayesian Analysis

ÖZ

ORTAKLAŞA KARIŞIK MODELLEMEDE RASSAL ETKENLERİN DAĞILIM VARSAYIMLARI

Özdemir, Celal Oğuz
Yüksek Lisans, İstatistik Bölümü
Tez Yöneticisi : Doç. Dr. Zeynep Işıl Kalaylıoğlu

Eylül 2018, 134 sayfa

Sağlık alanı çalışmalarında rassal etki modellerinin kullanılması oldukça yaygındır ve genelde bir hastalığın gidişatında hayatı tehdit edici bir sonuçla karşılaşmanın olabilirliğini tahmin için kullanılırlar. Bu tür sonuçların tahmini için bağımlı değişkenin iki sonuçlu değişken ve bağımsız değişkenlerin zamana bağlı boylamsal değişken olduğu ortaklaşa karışık modeli kullanılmıştır. Standart Bayesci modelde hataların varyansı ve rassal etkilerin varyans matriksi bağımsız düşünülüp ayrı ayrı önsellerle modellenmiş ve hataların varyansı için ters-Gamma ve rassal etkilerin varyans matriksi için ters-Wishart dağılımları kullanılmıştır. Daha yeni çalışmalarda, hata terimlerinin bağımsız düşünülemediği ve birleşik önselle modellendikleri Genellenmiş Çok Değişenli log Gamma (G-MVLG) dağılımını kullanmış ve gerçek dağılımı normal dağılım olan rassal etkilerde yanlılığı ve etkinliği açısından standart modelden sağlam (robust) olduğu gösterilmiştir. Bu tez çalışmasında ise, yeni yaklaşımın rassal etkinin gerçek dağılımlarına örneğin iki değişkenli Gamma dağılımı, iki değişkenli eğik-Normal da-

ğılımı gibi sağlamlığı araştırılmıştır. Araştırma sonuçlarında örneklem büyüklüğü yeteri miktarda büyük olan durumlarda bu yöntemin denenen dağılımlara sağlam olduğu görülmüştür. Küçük örneklemlerde görece yetersiz kalan model için rassal etkilerin önsel dağılımlarına normal dağılım tabanlı Dirichlet Süreç Önseli (DSÖ) tanımlanmış ve gerçekleştirilen benzetim çalışmaları sonucunda DSÖ'nün rassal etkilerin dağılımları Normal dağılım önseli ile modellenmiş metoda göre daha az yanlı ve daha etkili olduğu görülmüştür.

Anahtar Kelimeler: Ortaklaşa Karışık Model, Rassal Etkiler, Dirichlet İşlem Önseli, Genellenmiş Çok Değişkenli Log-Gamma Dağılımı, Bayesci Analiz

to my family

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisor, Assoc. Prof. Dr. Zeynep Işıl Kalaylıoğlu, for her guidance, encouragement, understanding and continuous support through the learning and writting process of this master thesis.

I would like to thank my examining committee members; Assoc. Prof. Dr. Özlem İlk Dağ, Assoc. Prof. Dr. Beyza Doğanay Erdoğan for their suggestions and comments.

This thesis would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended with their valuable assistance. I sincerely thank to METU Statistics Department staff especially, Büşra Taş, Buket Coşkun, Petek Aydemir, Berna Tuncer, Ezgi Ayyıldız, Duygu Varol, Ozancan Özdemir and International Student Office staff and my friends Ecem Ünal and Tülay Akal for their help.

I also want to express my gratitude to firstly Assoc. Prof. Dr. Ahmet Yozgatlıgil and all my workmates, Ceren Aksoy Güleç, Sina Soylu, Ayşe Elif Kıratlı, Sara Banu Akkaş, Asuman Özgür Keysan and Cihan Tuğrul Çiçek from Office of Sponsored Projects for their understanding, kindness and friendly manner. I would like to thank also my colleagues from other offices Saadet Beygu, Anıl Şengezer, Aydın Tiryaki, Dilek Ariç, Işık Tekin, Gülsenem Emiralioğlu and the most important pillars of Research Coordination Office, Gülay Doğan and Adem Akyüz.

I would like to thank my beloved friends, Selim Can, İrem Ece Can, Ayşe Hülya Örnek, Gizem Alaçam, Kübra Yılmaz, who have been there for me all the time when I needed. My teacher and the best friend forever, Deniz Çelikel, has supported me throughout this master study with her moral speeches and laughter. I will be grateful forever for your encouragement and friendship.

Last but not the least, I would like to thank my lovely family my father, Cemal, and

my sister, Olcay who have provided all kinds of support in my life and understanding throughout my entire study. And my gorgeous cat, Ateş, thank you for not eating my papers. I know that was very hard. I am very grateful to have you. I would like to present my sincere thankfulness to my deceased mother, Dudu Ayşe, for her supportive and loving role in my life and her numerous help and sacrifices for me and my sister.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction and Motivating Example

Researchers in medical field use longitudinally collected data (i.e. observations made over number of time points) more commonly compared to other fields. They use longitudinal data to gain a better understanding on important life threatening health situations. In this thesis, our focus is the analysis of data sets consisting of cross sectional response variables and longitudinal covariate data. Our research motivation comes from a study named GDM shortly (Gestational Diabetes Mellitus recorded over gestation period), which was conducted at Zekai Tahir Burak Women Health, Care and Research Hospital in Ankara (Özgü-Erdinç et al. (2015)). The aim of their study was to investigate the association between glycated hemoglobin (HbA1C) levels of pregnant women with gestational diabetes mellitus and the risk of obstetrics complication. Diabetic mothers are prone to have more labor complications compared to women without diabetics, meaning that they may need special preparation at delivery room. It is therefore important to know the risk the mother may have because such preparations usually have considerable cost for the hospital. The ability of knowing possibility of the risk is critical and worth working on. The data set of current study is based on a record of a dichotomous indicator for occurrence of complication at primary end-point, longitudinal covariates collected at every visit of the mother and cross-sectional covariates collected at her visit.

This type of data are analyzed using joint mixed model (Wang et al. (2000)). There are both frequentist and Bayesian approaches to analyze the model. Such complex

data structure and analysis method can be challenging on the side of mathematics. Bayesian method for solving nonlinear and intractable equations in the analysis provides an efficient approach compared to classical way and it is the focus of this thesis.

The main concern about the analysis of joint mixed model is the specification of the distribution of random effects and what consequences misspecification has on the posterior inference. Our aim is to investigate the effects of misspecification in the Bayesian analysis of JMM and propose a method that relaxes the distributional assumption on the random effects.

1.2 Literature Review

In this section, we provide a full literature review on JMM and its estimation. Parametric Bayesian estimation methods for JMM are given in the literature and reviewed below. To the best of our knowledge, current thesis provides the first Non-parametric Bayesian estimation of JMM.

1.2.1 Joint Mixed Model

The estimation model for the primary endpoint response is joint modelling approach proposed by Wang et al. (2000) which consists of two models; (i) the response model and (ii) the longitudinal covariate model. These two models are joined together through a common set random effects. Mathematical representation of the joint model is

$$f(Y_i|Z_i, X_i, \beta, \gamma) = exp\left(\frac{Y_i\theta_i - b(\theta_i)}{a(\phi)} + c(Y_i, \phi)\right) \tag{1.1}$$

$$W_i = D_i X_i + U_i \tag{1.2}$$

where $\theta_i = Z_i^T\beta + X_i^T\gamma$. The one in equation 1.1 is the response model, which is in the form of generalized linear model and 1.2 is the longitudinal covariate's part

in the form of linear random effects model. Y_i denotes response variable which is collected at a primary endpoint for subject i, $\boldsymbol{W_i}$=$\{\boldsymbol{W_{i1}}, \boldsymbol{W_{i2}}, ..., \boldsymbol{W_{in_i}}\}$ is the set of longitudinal measurement with $\boldsymbol{n_i}$ being the number of measurements for subject $\boldsymbol{i}$, $\boldsymbol{Z_i}$ be the $\boldsymbol{1 \times p}$ vector of cross-sectional covariates of mixed type measured for subject $\boldsymbol{i}$ including 1 for intercept. $\boldsymbol{X_i}$ denotes $\boldsymbol{q \times 1}$ latent effects which characterizes the $\boldsymbol{w_i}$ profile over time. Let $\boldsymbol{\beta}$ be the $\boldsymbol{p \times 1}$ vector of unknown regression coefficients associated with the cross-sectional covariates and $\boldsymbol{\gamma}$ be the $\boldsymbol{q \times 1}$ vector of unknown regression coefficients associated with $\boldsymbol{X_i}$. $\boldsymbol{D_i}$ is the $\boldsymbol{n_i \times q}$ design matrix for the time variables. $\boldsymbol{U_i}$=$\{\boldsymbol{U_{i1}}, \boldsymbol{U_{i2}}, ..., \boldsymbol{U_{in_i}}\}$ be the vector of uncorrelated error terms specified under the conditional independency assumption for $\boldsymbol{W_i}$'s given $\boldsymbol{X_i}$. $\boldsymbol{\phi}$ is a precision parameter, a,b and c are known functions.

This model is used for estimating regression parameters, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. While estimating, there are several assumptions which are

- all variables are independent across $\boldsymbol{i}$,
- $\boldsymbol{W_i}$ and $\boldsymbol{Y_i}$ are conditionally independent given $\boldsymbol{X_i}$,
- $\boldsymbol{U_i}$ and $\boldsymbol{X_i}$ are independent and normally distributed.

In this model $\boldsymbol{X_i}$'s are latent factors in (1.1) and have the role of random effects in (1.2). Therefore the terms *random effects* and *latent factors* will be used interchangeably throughout the text.

1.2.2 Standard Bayesian Analysis of Joint Mixed Model

In the standard Bayesian analysis of the JMM (JMM-St), variance components are traditionally assumed independent as usual. Restating the model in Bayesian point of view starts with understanding the distribution of the observed data and afterwards defining prior distribution for the parameters related distributions of observed data.

JMM-St is presented in Horrocks and van Den Heuvel (2009) as:

$$f(Y_i|Z_i, X_i, \beta, \gamma) = exp\left(\frac{Y_i\theta_i - b(\theta_i)}{a(\phi)} + c(Y_i, \phi)\right)$$

$$W_i = D_i X_i + U_i$$

$$X_i \mid \mu_X, \Sigma_X \sim F_X$$

$$U_i \mid \sigma_U^2 \sim F_U$$

$$\beta \sim F_\beta$$

$$\gamma \sim F_\gamma$$

$$\mu_X \sim F_{\mu_X}$$

$$\phi \sim F_\phi$$

$$\Sigma_X \sim Inverse - Wishart(\delta, G)$$

$$\sigma_U^2 \sim Inverse - Gamma(\delta_U, a_u)$$

where $\theta_i = Z_i^T\beta + X_i^T\gamma$.

In JMM-St model, proper vague priors which have large variances are used for prior distributions. Normal distribution with large variance is assigned for regression parameters (β and γ). Inverse-Wishart prior with parameters (δ, G) is assigned for variance-covariance matrix of random effects where δ is $p \times p$ scale matrix and $G > p - 1$ is degrees of freedom and Inverse-Gamma prior with (δ_U, a_U), where δ_U is shape parameter and a_U is scale parameter, is assigned for uncorrelated error variance component.

1.2.3 Bayesian Analysis of JMM Using Joint Variance Prior Approach

Kalaylıoğlu and Demirhan (2017) developed a novel Bayesian analysis of JMM taking into the account of the fact that variance components in the random effets model are inherently composite. To jointly model them apriori they considered the G-MVLG prior introduced by Demirhan and Hamurkaroğlu (2011). The Joint Mixed Model us-

ing G-MVLG prior will be called hereafter JMM-KD. Their approach is given below.

$$f(Y_i|Z_i, X_i, \beta, \gamma) = exp\left(\frac{Y_i\theta_i - b(\theta_i)}{a(\phi)} + c(Y_i, \phi)\right)$$

$$W_j = D_i X_i + U_i$$

$$X_i \mid \mu_X, \Sigma_X \sim F_X$$

$$U_i \mid \sigma_U^2 \sim F_U$$

$$\beta \sim F_\beta$$

$$\gamma \sim F_\gamma$$

$$\mu_X \sim F_{\mu_X}$$

$$\phi \sim F_\phi$$

$$(\Sigma_X, \sigma_U^2) \sim G - MVLG(\delta, \nu, \lambda, \eta)$$

where $\theta_i = Z_i^T\beta + X_i^T\gamma$ and $(\delta, \nu, \lambda, \eta)$ is positive parameter vector of G-MVLG.

Appropriate distributions are considered for the random effects and random error terms, e.g. $X \sim N_m(\mu_X, \Sigma_X)$ and $U_i \sim N(0, \sigma_U^2)$ respectively.

In the work by Kalaylıoglu and Demirhan (2017), the advantages of joint prior approach was illustrated over JMM-St which resulted in less relative bias and more efficiency.

1.3 Objectives and Scope of the Study

In this thesis, our general aim is to investigate the robustness of JMM-KD model to underlying true random effects distribution.

Our specific aims are listed below:

- Investigate the effect of assumed latent factors' distribution on the analysis results,
- Investigate if this effect depends on the outcome being a rare event or common event,

- Investigate how the effect changes with respect to sample size,
- Relax the distributional assumption on latent factors and develop a robust method for JMM.

This thesis organized as follows. Chapter 2 covers theoretical application of a non-parametric JMM. DPM application to JMM-St and JMM-SKD is presented in this chapter. Chapter 3 gives simulation study in details. The section includes a detailed outline of the random number generation process for the distributions under study whose codes are unavailable in MATLAB. In addition to that, convergence diagnostics for all the unknown variates of the model are discussed. The results of the simulation study are covered in this chapter. In Chapter 4, summary statistics of GDM data and application of JMM-St with DPM model and JMM-KD with DPM model on GDM dataset are presented and discussed. Lastly, chapter 5 gives summary and conclusion of this thesis.

CHAPTER 2

A NON-PARAMETRIC BAYESIAN JOINT MIXED MODEL

We consider Dirichlet Process Prior for the random effects. The general information about Dirichlet Process Prior (DPP) and Dirichlet Process Mixture (DPM) as well as the application of this non-parametric approach DPM to JMM-Stmodel and JMM-KD are given in this chapter.

2.1 Dirichlet Process Prior

Recently, researchers have been working on latent factors models without any assumption on random effects' distributions. Models that require no distribution assumption are more robust than the ones that are based on assumptions. The problem with assuming a specific distribution about the random effects is that identification of random effects are impossible and assuming normal distribution for random effects is too restrictive (Heinzl et al. (2012)). In order to create robust approach without any distribution assumption on random effects, a non-parametric method Dirichlet Process Prior (DPP) is considered. Dirichlet Process is a stochastic process whose realizations are probability distribution. With this aspect, DPP is more robust approach than assuming a certain distribution in the beginning. DPP simply creates a mixture distribution from a pre-defined distribution $\boldsymbol{G_0}$. The DPP starts with a baseline distribution or a centring distribution $\boldsymbol{G_0}$, $\boldsymbol{G}$ about the mean $\boldsymbol{G_0}$ (Congdon, (2006)). For any proper set of A, $\boldsymbol{G(A) \sim Beta(\alpha G_0(A), \alpha G(A^c))}$ in general, and so that $\boldsymbol{G(A) \xrightarrow{p} G_0(A)}$ as concentration parameter $\boldsymbol{\alpha \to \infty}$. When $\boldsymbol{\alpha}$ increases, the mixture distribution $\boldsymbol{G}$ gets closer to baseline distribution $\boldsymbol{G_0}$ in other words flexibility is

getting smaller when α increases.

DP prior has a wide usage in statistics especially in Bayesian analysis. DPP is used for the random effects distribution prior in Li et al. (2010). In the article, random effects distribution prior was created with univariate Normal baseline distribution. The study concluded that DP prior of random effects gives better results compared to other models discussed.

DP prior model for latent factor $\boldsymbol{X_i}$ is

$$\boldsymbol{X_i} \mid \varphi_i \sim \boldsymbol{F}(\varphi)$$
$$\varphi_i \mid \boldsymbol{G} \sim \boldsymbol{G}$$
$$\boldsymbol{G} \sim \boldsymbol{DP}(\boldsymbol{\alpha_0}, \boldsymbol{G_0})$$
$$\varphi = (\mu, \Sigma_X) \sim \boldsymbol{G_0}$$

φ is parameter vector of baseline distribution.

2.2 Dirichlet Process Mixture (DPM)

In Bayesian analysis, Dirichlet Process Mixture was used in Daniels and Pourahmadi (2002) to ease obtaining full conditional distribution of variance components of their model. DPM construction is based on Hanson et al. (2005) and the formula is

$$\boldsymbol{f}(. \mid \boldsymbol{G}) = \int \boldsymbol{f}(. \mid \nu)\boldsymbol{G}(\boldsymbol{d\nu}) = \sum_{\boldsymbol{k=1}}^{\infty} \boldsymbol{w_k f}(. \mid \boldsymbol{\nu_k}) \tag{2.1}$$

where $\boldsymbol{\nu}$ is parameters of baseline distribution $\boldsymbol{G_0}$. $\boldsymbol{G}$ is the random mixture distribution which is mixture of random distributions distributed Dirichlet Process ($\boldsymbol{DP}(\boldsymbol{\alpha}, \boldsymbol{G_0})$). For practical application, Ishwaran and Zarepour (2000) and Ishwaran and James (2002) suggest that truncated maximum number of clusters like C $\leq$ n components. Therefore, finite mixture of distributions are created with larger C.

$$\sum_{\boldsymbol{k=1}}^{\boldsymbol{C}} \boldsymbol{w_k f}(. \mid \boldsymbol{\nu_k})$$

The random weight of mixtures calculated as $w_c = p_c \prod_{c=1}^{C-1} (1 - p_c)$ with $p_c \overset{iid}{\sim}$Beta(1,α). The sum of weights are arranged to be equal to 1. c is the number of clusters for considered population. The concentration (precision) parameter α has critical impact on number of clusters in prediction phase.

DPM approach to random effects distribution prior was used by De la Cruz-Mesia et al. (2007) in which the main goal was to achieve the classification between normal and abnormal pregnancy. The response variable was $\beta - HCG$ (continuous; longitudinal) for 173 young women. The sample size was 376 with all contribution of 173 women. With the help of DP, they avoided assuming the random effect distribution with a parametric method because it can critically change the predictive classification probabilities when assumed distribution has detailed features.

2.3 Dirichlet Process Mixture Application to Joint Mixture Models

The standard joint mixed model assumes that random effects have known multivariate distribution such as multivariate normal or multivariate t. Distribution assumption on random effects are generally considered weakness in modelling. These parametric assumptions may be relaxed by a more general non-parametric model for making robust inferences. Instead of using restrictive parametric assumption, we proposed non-parametric model development by non-parametric method Dirichlet Process Mixture for random effect distribution assumption. The aim of this method is to avoid bias resulting of the estimation from misleading random effect distribution assumption.

We consider mean parameters for each baseline distributions equal and zero. The variances of mixture distributions are taken as common variance which is obtained from model itself.

For DPM concentration parameter α_0, we consider Gamma distribution with a shape parameter and b scale parameter. Selecting the concentration parameter is an important issue for the Dirichlet Process because it regulates the number of clusters. In the paper by Murugiah and Sweeting (2012), the concentration parameter prior selection

was discussed and various specification methods were detailed. We considered one of these methods for the Gamma distribution parameters selection. This prior parameters selection method is based on the simulation study conducted for different maximum number of clusters. There are equations for selecting concentration parameter's prior listed in Murugiah and Sweeting (2012). Instead of solving previous equations for the parameter selection, they suggested to fit the approximating curve that is obtained from previous equations by simulations and following equation is found

$$\boldsymbol{a = b = e^{-0.033C}} \tag{2.2}$$

where C is the considered maximum number of cluster. Here, Equation 2.3 is only investigated for $\mathbf{2 \leq C < 6}$. An important note that the previous equations that suggested for selecting the best gamma distribution parameters have feature when maximum number of clusters $\boldsymbol{C \to \infty}$, $\boldsymbol{a}$ and $\boldsymbol{b}$ both approach to zero as in equation 2.3.

2.3.1 JMM-St with DPM Model

The implementation of DPM to JMM-St Model is given below:

$$\boldsymbol{f(Y_i|Z_i, X_i, \beta, \gamma) = exp\left(\frac{Y_i\theta_i - b(\theta_i)}{a(\phi)}\right) + c(Y_i, \phi)}$$

$$\boldsymbol{\theta_i = Z_i^T\beta + X_i^T\gamma}$$

$$\mathrm{W}_{ij} = \boldsymbol{D_{ij}X_i + U_{ij}}$$

$$\boldsymbol{X_i \mid K_i, \varphi \sim N_m(\mu_{K_i}, \Sigma_{X,K_i}) \quad i = 1, ..., n}$$

$$\mathrm{K}_i \mid \boldsymbol{p \sim Discrete(p_1, ..., p_C)}$$

$$\boldsymbol{\varphi_k = (\mu_k, \Sigma_{X,k}) \sim G_0 \quad k = 1, ..., C}$$

$$\mathrm{G} \sim \boldsymbol{DP(G_0, \alpha_0)}$$

$$\boldsymbol{\alpha_0 \sim Gamma(a, b)}$$

$$\mathrm{U}_i \mid \boldsymbol{\sigma_U^2 \sim F_U}$$

$$\boldsymbol{\beta} \sim \boldsymbol{N_p}(\boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta})$$

$$\boldsymbol{\gamma} \sim \boldsymbol{N_q}(\boldsymbol{\mu_\gamma}, \boldsymbol{\Sigma_\gamma})$$

$$\boldsymbol{\phi} \sim \boldsymbol{F_\phi}$$

$$\boldsymbol{\Sigma_X} \sim \boldsymbol{IW}(\boldsymbol{\delta}, \boldsymbol{G})$$

$$\boldsymbol{\sigma_U^2} \sim \boldsymbol{IG}(\boldsymbol{\delta_U}, \boldsymbol{a_u})$$

Letting $\boldsymbol{P} = (\boldsymbol{p_1}, \boldsymbol{p_2}, ..., \boldsymbol{p_C})$, $\boldsymbol{\varphi} = (\boldsymbol{\varphi_1}, ..., \boldsymbol{\varphi_C})$, $\mathbb{X} = \{\boldsymbol{X_1}, ..., \boldsymbol{X_n}\}$, $\boldsymbol{K} = (\boldsymbol{K_1}, ..., \boldsymbol{K_n})$ The joint posterior of the model JMM-St is

$$\boldsymbol{f}(\boldsymbol{\varphi}, \boldsymbol{p}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \mathbb{X}, \boldsymbol{\beta}, \boldsymbol{\gamma} \mid \boldsymbol{Y}, \boldsymbol{W}) \propto \prod_{\boldsymbol{i=1}}^{\boldsymbol{n}} \prod_{\boldsymbol{j=1}}^{\boldsymbol{n_i}} \boldsymbol{f}(\boldsymbol{Y_i}, \boldsymbol{W_{ij}} \mid \boldsymbol{\varphi}, \boldsymbol{P}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \boldsymbol{X_i}, \boldsymbol{\beta}, \boldsymbol{\gamma})$$
$$\times \boldsymbol{f}(\boldsymbol{\mu_k}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \mathbb{X}) \times \boldsymbol{f}(\boldsymbol{P}) \times \boldsymbol{f}(\boldsymbol{\gamma}) \times \boldsymbol{f}(\boldsymbol{\beta}) \times \boldsymbol{f}(\boldsymbol{\sigma_U^2}) \times \boldsymbol{f}(\boldsymbol{\Sigma_X})$$

Then,

$$= \prod_{\boldsymbol{i=1}}^{\boldsymbol{n}} \prod_{\boldsymbol{j=1}}^{\boldsymbol{n_i}} \boldsymbol{f}(\boldsymbol{Y_i} \mid \boldsymbol{W_{ij}}, \boldsymbol{\varphi}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{P}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \boldsymbol{X_i}) \times \boldsymbol{f}(\boldsymbol{W_{ij}} \mid \boldsymbol{\varphi}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \boldsymbol{X_i}) \times$$
$$\boldsymbol{f}(\mathbb{X} \mid \boldsymbol{K}, \boldsymbol{\varphi}) \times \boldsymbol{f}(\boldsymbol{K} \mid \boldsymbol{P}) \times \boldsymbol{f}(\boldsymbol{\gamma}) \times \boldsymbol{f}(\boldsymbol{\beta}) \times \boldsymbol{f}(\boldsymbol{\sigma_U^2}) \times \boldsymbol{f}(\boldsymbol{\Sigma_X})$$

Here, $\boldsymbol{f}(\boldsymbol{p})$ is equivalent to following stick-breaking approach,

$$\boldsymbol{p_h} \overset{\boldsymbol{iid}}{\sim} \boldsymbol{Beta}(\boldsymbol{1}, \boldsymbol{\alpha_0}) \quad \boldsymbol{h = 1, 2, .., C - 1,}$$

where C is maximum cluster number (Heinzl et al. (2012)),

$$\boldsymbol{p_h} \mid \boldsymbol{\alpha_0} \sim \boldsymbol{Be}\left(\boldsymbol{1 + n_h}, \boldsymbol{\alpha_0} + \sum_{\boldsymbol{l=h+1}}^{\boldsymbol{C}} \boldsymbol{n_l}\right)$$

$$\boldsymbol{\alpha_0} \mid \boldsymbol{p_1}, \boldsymbol{p_2}, ...\boldsymbol{p_{C-1}} \sim \boldsymbol{Ga}\left(\boldsymbol{C - 1 + a}, \boldsymbol{b} - \sum_{\boldsymbol{h=1}}^{\boldsymbol{C-1}} \boldsymbol{log}(\boldsymbol{1 - p_h})\right)$$

the full conditional of φ is

$$f(\varphi) = f(\mu_k) \times f(\Sigma_X)$$

since μ_k is selected fixed as 0 and Σ_X is distributed IW independently. Conjugate Wishart Updater is used for posterior density of Σ_X.

Full conditional of X is

$$f(\mathbb{X} \mid K, \varphi, p, \alpha_0) \propto \prod_{i=1}^{n} \prod_{j=1}^{n_i} f(Y_i \mid W_{ij}, \varphi, \beta, \gamma, p, K, \alpha_0, X_i) \times f(W_{i,j} \mid \varphi, K, \alpha_0, X_i) \times f(\mathbb{X} \mid K, \varphi)$$

$$f(\mathbb{X} \mid K, \varphi) = \prod_{i=1}^{n} N_m(\mu_{K_i}, \Sigma_X)$$

For β and γ, normal priors are assigned and MCMC methods (i.e. Gibbs Sampling) are used for obtaining random samples from posterior density.

$$f(\beta \mid .) \propto \prod_{i=1}^{n} f(Y_i \mid W_{ij}, \varphi, \gamma, p, K, \alpha_0, X_i) \times f(\beta) \tag{2.3}$$

$$f(\beta) = N_m(\mu_\beta, \Sigma_\beta)$$

$$f(\gamma \mid .) \propto \prod_{i=1}^{n} f(Y_i \mid W_{ij}, \varphi, \beta, p, K, \alpha_0, X_i) \times f(\gamma) \tag{2.4}$$

$$f(\gamma) = N_m(\mu_\gamma, \Sigma_\gamma)$$

2.3.2 JMM-KD with DPM model

The implementation of DPM to JMM-KD model is given below:

$$f(Y_i|Z_i, X_i, \beta, \gamma) = exp\left(\frac{Y_i\theta_i - b(\theta_i)}{a(\phi)}\right) + c(Y_i, \phi)$$

$$\boldsymbol{\theta}_i = \boldsymbol{Z}_i^T \boldsymbol{\beta} + \boldsymbol{X}_i^T \boldsymbol{\gamma}$$

$$\mathrm{W}_{ij} = \boldsymbol{D}_i \boldsymbol{X}_i + \boldsymbol{U}_{ij}$$

$$\mathrm{X}_i \mid \boldsymbol{K}_i, \boldsymbol{\varphi} \sim \boldsymbol{N}_m(\boldsymbol{\mu}_{\boldsymbol{K}_i}, \boldsymbol{\Sigma}_{\boldsymbol{X},\boldsymbol{K}_i}) \quad \boldsymbol{i = 1, ..., n}$$

$$\mathrm{K}_i \mid \boldsymbol{p} \sim \boldsymbol{Discrete}(\boldsymbol{p_1}, ..., \boldsymbol{p_C})$$

$$\boldsymbol{\varphi_k} = (\boldsymbol{\mu_k}, \boldsymbol{\Sigma_{X,k}}) \sim \boldsymbol{G} \quad \boldsymbol{k = 1, ..., C}$$

$$\mathrm{G} \sim \boldsymbol{DP}(\boldsymbol{G_0}, \boldsymbol{\alpha_0})$$

$$\boldsymbol{\alpha_0} \sim \boldsymbol{Gamma}(\boldsymbol{a}, \boldsymbol{b})$$

$$\mathrm{U}_i \mid \boldsymbol{\sigma_U^2} \sim \boldsymbol{F_U}$$

$$\boldsymbol{\beta} \sim \boldsymbol{N_p}(\boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta})$$

$$\boldsymbol{\gamma} \sim \boldsymbol{N_q}(\boldsymbol{\mu_\gamma}, \boldsymbol{\Sigma_\gamma})$$

$$\boldsymbol{\phi} \sim \boldsymbol{F_\phi}$$

$$(\boldsymbol{\Sigma_X}, \boldsymbol{\sigma_U^2}) \sim \boldsymbol{G - MVLG}(\boldsymbol{\delta}, \boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\eta})$$

Letting $\boldsymbol{P} = (\boldsymbol{p_1}, \boldsymbol{p_2}, ..., \boldsymbol{p_C})$, $\boldsymbol{\varphi} = (\boldsymbol{\varphi_1}, ..., \boldsymbol{\varphi_C})$, $\mathbb{X} = \{\boldsymbol{X_1}, ..., \boldsymbol{X_n}\}$, $\boldsymbol{K} = (\boldsymbol{K_1}, ..., \boldsymbol{K_n})$, the joint posterior distribution of the model parameters and latent variables with DPM is

$$\boldsymbol{f}(\boldsymbol{\varphi}, \boldsymbol{p}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \mathbb{X}, \boldsymbol{\beta}, \boldsymbol{\gamma} \mid \boldsymbol{Y}, \boldsymbol{W}) \propto \prod_{\boldsymbol{i=1}}^{\boldsymbol{n}} \prod_{\boldsymbol{j=1}}^{\boldsymbol{n_i}} \boldsymbol{f}(\boldsymbol{Y_i}, \boldsymbol{W_{ij}} \mid \boldsymbol{\varphi}, \boldsymbol{P}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \boldsymbol{X_i}, \boldsymbol{\beta}, \boldsymbol{\gamma})$$
$$\times \boldsymbol{f}(\boldsymbol{\mu_k}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \mathbb{X}) \times \boldsymbol{f}(\boldsymbol{P}) \times \boldsymbol{f}(\boldsymbol{\gamma}) \times \boldsymbol{f}(\boldsymbol{\beta}) \times \boldsymbol{f}(\boldsymbol{\sigma_U^2}, \boldsymbol{\Sigma_X})$$

Then,

$$= \prod_{i=1}^{n} \prod_{j=1}^{n_i} f(\boldsymbol{Y_i} \mid \boldsymbol{W_{ij}}, \boldsymbol{\varphi}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{p}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \boldsymbol{X_i}) \times f(\boldsymbol{W_{i,j}} \mid \boldsymbol{\varphi}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \boldsymbol{X_i}) \times$$

$$f(\boldsymbol{\mu_k}, \boldsymbol{K}, \boldsymbol{\alpha_0}, \boldsymbol{X}) \times f(\boldsymbol{p}) \times f(\boldsymbol{\gamma}) \times f(\boldsymbol{\beta}) \times f(\boldsymbol{\sigma_U^2}, \boldsymbol{\Sigma_X})$$

where $f(\boldsymbol{\beta})$, $f(\boldsymbol{\gamma})$, $f(\boldsymbol{\sigma_U^2}, \boldsymbol{\Sigma_X})$ and $f(\boldsymbol{\alpha_0})$ are priors.

CHAPTER 3

SIMULATION STUDY

In this chapter, a Monte Carlo (MC) simulation study is presented. The primary aim of this MC simulation study is investigating robustness of JMM-KD under the possible true distributions for random effects. Process of generating random numbers from specified distributions is detailed and Mixture Distribution method also is explained. Base model for the simulation study is the model JMM-KD and JMM-St. For Markov chain Monte Carlo application, OpenBUGS is used (Spiegelhalter et al., (2003)). MATLAB environment is used to generate random numbers from specified distributions. MATLAB and OpenBUGS in simulations are used together with the help of open sourced package "mat2bugs". Convergence diagnostics of both JMM-St and JMM-KD are placed in Appendix A and Appendix B respectively and results are evaluated in this chapter.

We consider Bivariate Normal Distribution, Bivariate Normal Mixture Distribution, Bivariate Gamma Distribution, Bivariate Gamma Mixture Distribution, Bivariate Skew-Normal Distribution and Bivariate Skew-Normal Mixture Distribution as true random effects generating mechanisms. In section 3.1, we provide details of data generation from Bivariate Gamma and Bivariate Skew-Normal.

3.1 Random Effects Generating Distributions

In this section, Bivariate Gamma Distribution and Skew-Normal Distribution and their random number generation method are given.

3.1.1 Bivariate Gamma Distribution

The probability density function of bivariate Gamma distribution is based on five parameters which are two shape, two location and one correlation parameters of Bivariate Gamma distribution function (Smith et al. (1982)). Letting the random vector (X,Y) has a Bivariate gamma distribution, its probability density function (pdf) is given by

$$f_{X,Y}(t_1,t_2;\gamma_1,\gamma_2,\eta)=\frac{t_1^{(\gamma_1-1)}t_2^{(\gamma_2-1)}exp(-\frac{(t_1+t_2)}{(1-\eta)})}{(1-\eta)^{\gamma_1}\Gamma(\gamma_1)\Gamma(\gamma_2-\gamma_1)}\times\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{n^{j+k}}{(1-\eta)^{2j+k}}\frac{\Gamma(\gamma_2-\gamma_1+k)}{\Gamma(\gamma_2+j+k)}\frac{(t_1t_2)^j t_2^k}{j!k!} \tag{3.1}$$

where $t_1=\beta_1 x$, $t_2=\beta_2 y$, β_1 and β_2 are known scale parameters, γ_1 and γ_2 are shape parameters. $\eta=\rho\sqrt{\gamma_2/\gamma_1}$ where ρ is the correlation coefficient between the variables X and Y. The random number generation algorithm is based on Schmeiser and Lal (1982). The algorithm is

- Generate Y_1 from $Ga(\gamma_1-\rho(\gamma_1\gamma_2)^{(1/2)},1)$
- Generate Y_2 from $Ga(\gamma_2-\rho(\gamma_1\gamma_2)^{(1/2)},1)$
- Generate $Y_3\sim Ga(\rho(\gamma_1\gamma_2)^{(1/2)},1)$
- $X_1=(Y_1+Y_3)\beta_1$
- $X_1=(Y_2+Y_3)\beta_2$
- Return (X_1,X_2)

3.1.2 Multivariate Skew-Normal Distribution

Skew-Normal (SN) probability density function is proposed by Azzalini and Dalla Valle (1996). MATLAB code was given in web page of Azzalini. Multivariate skew-normal distribution random number generation MATLAB codes are available as a

library in the web page. The library has univariate and multivariate software of Azzalini's Skew-Normal algorithm. The pdf of p-variate SN distribution and random number generation algorithm is given as:

$$\boldsymbol{Y} \sim \boldsymbol{SN_p(\mu, \Sigma, \lambda)}$$
$$\boldsymbol{f(y) = 2\phi_p(y \mid \mu, \Sigma)\Phi_1(\lambda^T \Sigma^{-(1/2)}(y - \mu)), \quad y \in \mathbb{R}}$$

where $\phi_p(. \mid \mu, \Sigma)$ is the pdf of p-variate normal distribution. Φ_1 denotes the cumulative distribution function of the standard normal distribution.

3.1.3 Bivariate Mixture Distribution

Bernoulli distribution baseline approach is used for generating random numbers from mixture distributions. When Bernoulli distribution has the success probability p, mixture pdf can be found by the formula

$$\boldsymbol{g(y) = p \times f_1(y) + (1 - p) \times f_2(y)}$$

By using this approach, random number generation algorithm is

- Step 1 : Generate $\boldsymbol{U}$ from $\boldsymbol{Ber(p)}$
- Step 2 : Generate $\boldsymbol{X_1}$ from $\boldsymbol{F_{X_1}}$ and $\boldsymbol{X_2}$ from $\boldsymbol{F_{X_2}}$
- Step 3 : $\boldsymbol{M = U \times X_1 + (1 - U) \times X_2}$, where $\boldsymbol{M}$ is distributed from $\boldsymbol{g}$ mixture pdf.

3.2 Simulation

We conducted to sets of simulations studies. In the first study, the simulation scenarios basically focused on to explore the robustness of JMM-St and JMM-KD to true distributions of latent factors. These true distributions are selected as normal, normal mixture, skew-normal, skew-normal mixture, gamma and gamma mixture distributions. Selected sample sizes are 100 for small sample size case, and 500 for

moderately large sample size case. Two different response distributions are considered. These are $\boldsymbol{P(Y \mid x) = 0.1}$ representing rare disease and $\boldsymbol{P(Y \mid x) = 0.5}$ representing common disease. Number of MCMC samples to get efficient estimates is decided based on the rule that Monte Carlo (MC) Error should be less than 5% of the sample standard deviation. We observed this rule at $\mathbf{50.000}$ MCMC samples. Scenarios are repeated 100 times for all.

In the second study, we tested Dirichlet Process Mixture modification on JMM-St and JMM-KD via simulation study. In order to compare efficiency, we used Gaussian Normal distribution for true random effect distribution. We consider only small sample size case as 100 observations. As in the previous scenario, two different disease types are considered which are rare and common disease. The rule of choosing number of samples is applied. Number of samples is $\mathbf{50.000}$. Scenarios are repeated 100 times. Convergence diagnostics and posterior distributions of parameters are given in Appendices D and E. In addition, in Appendix H, OpenBUGS codes of DPM models can be found.

Convergence diagnostics are used to ensure that the Gibbs Sampling random draws converge in distribution to the posterior distribution. We used the following OpenBUGS convergence diagnostics and posterior statistics such as:

- Brooks-Gelman-Rubin (BGR) (1998) diagnostics is calculated with two chain of initial values and analysis concluded that all scenarios are converged (BGR statistics < 1.1)
- Auto-correlation graphics of OpenBUGS,
- Posterior density.

Convergence diagnostics of distribution sensitivity analysis and convergence diagnostics of DPM versions of the models can be found in Appendix A, B, D and E. As a result of these convergence diagnostic indicators, burn-in (warm-up) iteration number was decided 5000 for all scenarios. In addition to convergence diagnostics results, posterior densities of estimated parameters are given in appendices.

Results of these simulation study are tabulated for summary statistics which are Relative Bias (RB) and Interquartile Range (IQR). These indicators were calculated for posterior medians which is more robust estimate than posterior mean because of asymmetricity in posterior densities.

- **Relative Bias**: Relative bias is a comparable statistic between different distributions of random effects. The calculation formula is

$$\boldsymbol{RB} = \frac{\boldsymbol{E}[\hat{\theta}] - \boldsymbol{\theta}}{\boldsymbol{\theta}}$$

  where $\boldsymbol{\theta}$ is the parameter and $\boldsymbol{E}[\hat{\theta}]$ is estimated using the MC averages of posterior medians.

- **Interquartile Range**: Interquartile range of posterior median is used to evaluate efficiency. Narrower IQR means more efficient estimate. The calculation is

$$\boldsymbol{IQR} = \boldsymbol{Q_3} - \boldsymbol{Q_1}$$

  where $\boldsymbol{Q_1}$ is the first quartile and $\boldsymbol{Q_3}$ is the third quartile of median in the posterior distribution.

- **MC Standard Error**: MC Standard Error is standard error of estimated median values of all MCMC chains.

$$\boldsymbol{s} = \sqrt{\frac{\sum_{k=1}^{100}(\hat{\theta}_k - \bar{\hat{\theta}})^2}{100 - 1}}$$

$$\boldsymbol{MCSE} = \frac{\boldsymbol{s}}{\sqrt{100}}$$

$\hat{\theta}_k$ are the posterior median values of each simulated data set and $\bar{\hat{\theta}}$ is their average.

3.3 Distribution Sensitivity Analysis

In this section, results of the simulation study is presented in four tables for JMM-St and JMM-KD. MC Standard Error tables can be found in Appendix C.

Table 3.1: Relative Bias of JMM-St

		Relative Bias					
n	Parameter	$X \sim N$	$X \sim N$ Mixture	$X \sim SN$ (Skewness=0.2)	$X \sim SN$ Mixture (Skewness=0.1)	$X \sim Ga$	$X \sim Ga$ Mixture
$P(Y=1 \mid x)=0.1$							
100	β	0.1910	0.2319	0.0603	0.1957	0.1462	0.3552
	γ_1	0.2081	0.3217	-0.0484	0.3283	0.2357	0.4250
	γ_2	0.1946	0.2302	-0.0100	0.2655	0.1815	0.1763
	σ_U^2	0.0151	0.0059	0.0418	0.0296	0.0177	0.0162
	$Var(X_{i0})$	-0.0358	-0.0524	-0.1613	-0.0942	-0.0929	-0.0199
	$Var(X_{i1})$	-0.0156	0.0041	-0.0172	-0.0110	-0.0339	-0.0047
	$Cov(X_{i0}, X_{i1})$	0.0307	0.0140	-0.0351	-0.0455	-0.0347	0.0067
500	β	0.0283	0.0784	0.1046	0.0437	-0.1395	0.0451
	γ_1	0.0842	0.1095	0.1825	0.2269	-0.1010	-0.0004
	γ_2	-0.0146	0.0719	0.1496	0.0941	0.0184	0.0699
	σ_U^2	0.0031	0.0014	0.0018	0.0006	0.0021	0.0052
	$Var(X_{i0})$	-0.0199	-0.0129	-0.0157	-0.0222	-0.0169	-0.0122
	$Var(X_{i1})$	-0.0040	0.0126	-0.0010	0.0051	-0.0464	0.0031
	$Cov(X_{i0}, X_{i1})$	0.0004	-0.0388	0.0081	0.0261	0.0358	-0.0186
$P(Y=1 \mid x)=0.5$							
100	β	0.2268	0.2351	0.1801	0.2551	0.3765	0.1266
	γ_1	0.2188	0.3741	0.2208	0.2495	0.7451	0.2007
	γ_2	0.1571	0.1946	0.0763	0.2401	-0.0567	0.0699
	σ_U^2	0.0208	0.0247	-0.0027	0.0148	0.0255	0.0397
	$Var(X_{i0})$	-0.0307	-0.0743	-0.0244	-0.0253	-0.0983	-0.0587
	$Var(X_{i1})$	-0.0225	0.0010	-0.0023	0.0100	-0.0471	0.0137
	$Cov(X_{i0}, X_{i1})$	-0.0421	-0.1028	-0.0389	0.0926	-0.0270	0.0289
500	β	0.0823	0.0446	0.0594	0.2160	0.1142	0.0531
	γ_1	0.0845	0.0526	0.0521	0.1761	0.1447	0.0705
	γ_2	0.0665	0.0125	0.0367	0.2641	0.0775	0.0199
	σ_U^2	0.0077	-0.0005	0.0034	0.0057	0.0072	0.0068
	$Var(X_{i0})$	-0.0140	-0.0316	-0.0254	-0.0112	-0.0253	-0.0219
	$Var(X_{i1})$	-0.0015	-0.0031	0.0101	0.0165	-0.0315	0.0165
	$Cov(X_{i0}, X_{i1})$	-0.0248	-0.0421	0.0130	0.1299	0.0585	-0.0042

Table 3.2: Relative Bias of JMM-KD

		Relative Bias					
n	Parameter	$X \sim N$	$X \sim N$ Mixture	$X \sim SN$ (Skewness=0.2)	$X \sim SN$ Mixture (Skewness=0.1)	$X \sim Ga$	$X \sim Ga$ Mixture
$P(Y=1 \mid x)=0.1$							
100	β	0.2250	0.2287	0.1680	0.1854	0.5354	0.2082
	γ_1	0.4817	0.3482	0.2027	0.0207	0.4903	0.2022
	γ_2	0.3192	0.1384	0.1431	0.3489	0.2468	0.2033
	σ_U^2	0.0054	0.0166	0.0233	0.0191	0.0166	0.0102
	$Var(X_{i0})$	0.0053	-0.0137	-0.0287	-0.0347	-0.0354	0.0080
	$Var(X_{i1})$	0.0453	0.0321	0.0051	0.0321	-0.0324	0.0233
	$Cov(X_{i0}, X_{i1})$	0.0154	-0.0909	0.0078	-0.0237	0.0462	0.0535
500	β	0.0256	0.0857	0.0815	0.0454	-0.0611	-0.0061
	γ_1	0.0721	0.1288	0.1334	0.2272	-0.0793	-0.0337
	γ_2	-0.0188	0.0668	0.1217	0.0997	0.0649	0.0459
	σ_U^2	0.0060	0.0041	0.0065	0.0045	0.0097	-0.0054
	$Var(X_{i0})$	-0.0044	0.0145	-0.0020	0.0072	-0.0120	-0.0051
	$Var(X_{i1})$	0.0032	0.0119	0.0033	0.0119	-0.0373	0.0231
	$Cov(X_{i0}, X_{i1})$	-0.0270	0.0010	0.0206	0.1196	0.0650	0.0164
$P(Y=1 \mid x)=0.5$							
100	β	0.1509	0.2177	0.0704	0.1752	0.4114	0.1763
	γ_1	0.2067	0.3436	0.0927	0.1569	0.6826	0.2322
	γ_2	0.1406	0.1869	-0.0121	0.2305	0.3055	0.1232
	σ_U^2	0.0050	0.0210	-0.0040	0.0064	0.0176	0.0195
	$Var(X_{i0})$	0.0057	-0.0238	0.0071	0.0418	-0.0280	0.0327
	$Var(X_{i1})$	0.0351	0.0321	0.0105	0.0241	0.0082	0.0492
	$Cov(X_{i0}, X_{i1})$	0.0299	-0.0920	-0.0188	-0.0956	0.0664	0.0133
500	β	0.0902	0.0824	0.1022	0.1920	0.0936	0.0307
	γ_1	0.1002	0.0756	0.0942	0.1615	0.0499	0.0358
	γ_2	0.0649	0.0618	0.0588	0.2256	0.1372	0.0385
	σ_U^2	0.0036	0.0058	0.0064	0.0019	0.0074	0.0010
	$Var(X_{i0})$	0.0065	0.0024	-0.0026	0.0073	-0.0133	0.0042
	$Var(X_{i1})$	0.0066	0.0111	-0.0022	0.0103	-0.0482	0.0079
	$Cov(X_{i0}, X_{i1})$	-0.0184	-0.0575	-0.0019	-0.0063	-0.0256	0.0306

Table 3.3: Interquartile Range of JMM-St

		Interquartile Range					
n	Parameter	X $\sim N$	X $\sim N$ Mixture	X $\sim SN$ (Skewness=0.2)	X $\sim SN$ Mixture (Skewness=0.1)	X $\sim Ga$	X $\sim Ga$ Mixture
$P(Y=1 \mid x)=0.1$							
100	β	1.0555	2.5505	3.2311	0.8670	2.1803	2.2791
	γ_1	0.8399	1.6380	1.6145	0.7528	1.3237	1.1213
	γ_2	0.7139	1.3254	1.5591	0.6644	1.4225	1.3916
	σ_U^2	0.0603	0.0579	0.0679	0.0614	0.0608	0.0591
	$Var(X_{i0})$	0.3175	0.3275	0.3457	0.3456	0.2733	0.2735
	$var(X_{i1})$	0.1212	0.1353	0.0823	0.1255	0.1117	0.1055
	$Cov(X_{i0}, X_{i1})$	0.1824	0.1489	0.1352	0.1782	0.2552	0.2684
500	β	0.3581	1.1266	1.1194	0.3016	0.7557	0.8259
	γ_1	0.2868	0.6909	0.4617	0.2478	0.3991	0.3554
	γ_2	0.2751	0.5755	0.5568	0.2534	0.5501	0.5452
	σ_U^2	0.0256	0.0256	0.0255	0.0256	0.0256	0.0257
	$Var(X_{i0})$	0.1554	0.1345	0.2989	0.1227	0.1555	0.1028
	$Var(X_{i1})$	0.0544	0.0608	0.0350	0.0568	0.0490	0.0472
	$Cov(X_{i0}, X_{i1})$	0.1124	0.0933	0.3057	0.0908	0.1180	0.1007
$P(Y=1 \mid x)=0.5$							
100	β	1.6820	0.7530	2.6856	1.3078	1.0355	2.1486
	γ_1	1.7907	1.3891	1.7751	1.9559	0.5279	0.6654
	γ_2	1.1733	1.0441	1.3714	1.2029	0.4506	0.7816
	σ_U^2	0.0576	0.0587	0.0553	0.0568	0.0606	0.0604
	$Var(X_{i0})$	0.4042	0.3703	0.7153	0.2885	0.2032	0.3016
	$var(X_{i1})$	0.1202	0.1346	0.0789	0.1283	0.1109	0.1075
	$Cov(X_{i0}, X_{i1})$	0.2803	0.2366	0.7003	0.2135	0.2418	0.2785
500	β	0.8315	0.2850	1.4354	0.7301	0.3664	0.8483
	γ_1	0.9095	0.4824	0.9050	1.1485	0.1686	0.2454
	γ_2	0.5837	0.3902	0.7309	0.6935	0.1727	0.3187
	σ_U^2	0.0256	0.0255	0.0253	0.0255	0.0257	0.0257
	$Var(X_{i0})$	0.1527	0.1376	0.2990	0.1194	0.1576	0.1035
	$Var(X_{i1})$	0.0548	0.0598	0.0355	0.0574	0.0499	0.0478
	$Cov(X_{i0}, X_{i1})$	0.1100	0.0950	0.3032	0.0892	0.1187	0.1020

Table 3.4: Interquartile Range of JMM-KD

		Interquartile Range					
n	Parameter	X $\sim \boldsymbol{N}$	X $\sim \boldsymbol{N}$ Mixture	X $\sim \boldsymbol{SN}$ (Skewness=0.2)	X $\sim \boldsymbol{SN}$ Mixture (Skewness=0.1)	X $\sim \boldsymbol{Ga}$	X $\sim \boldsymbol{Ga}$ Mixture
$\boldsymbol{P(Y=1 \mid x)=0.1}$							
100	$\boldsymbol{\beta}$	1.0305	2.5519	2.8023	0.8308	2.2753	2.1292
	$\boldsymbol{\gamma_1}$	0.8009	1.6111	1.2350	0.7008	1.3577	1.0294
	$\boldsymbol{\gamma_2}$	0.7065	1.2801	1.3675	0.6461	1.4363	1.3496
	$\boldsymbol{\sigma_U^2}$	0.0580	0.0580	0.0591	0.0588	0.0587	0.0586
	$\boldsymbol{Var(X_{i0})}$	0.3168	0.3153	0.2834	0.3333	0.3048	0.3922
	$\boldsymbol{var(X_{i1})}$	0.1267	0.1407	0.0802	0.1331	0.1134	0.1100
	$\boldsymbol{Cov(X_{i0}, X_{i1})}$	0.1434	0.1493	0.1179	0.1470	0.1303	0.1481
500	$\boldsymbol{\beta}$	0.3577	1.1580	1.1078	0.3018	0.7728	0.8044
	$\boldsymbol{\gamma_1}$	0.2841	0.7038	0.4563	0.2440	0.4061	0.3452
	$\boldsymbol{\gamma_2}$	0.2723	0.5813	0.5555	0.2534	0.5656	0.5357
	$\boldsymbol{\sigma_U^2}$	0.0257	0.0257	0.0256	0.0257	0.0257	0.0254
	$\boldsymbol{Var(X_{i0})}$	0.1361	0.1422	0.1252	0.1495	0.1353	0.1708
	$\boldsymbol{Var(X_{i1})}$	0.0550	0.0608	0.0353	0.0574	0.0495	0.0483
	$\boldsymbol{Cov(X_{i0}, X_{i1})}$	0.0627	0.0668	0.0523	0.0662	0.0585	0.0658
$\boldsymbol{P(Y=1 \mid x)=0.5}$							
100	$\boldsymbol{\beta}$	1.4656	0.7439	2.5508	1.2282	0.9454	2.1876
	$\boldsymbol{\gamma_1}$	1.7711	1.3552	1.6465	1.8120	0.4585	0.6436
	$\boldsymbol{\gamma_2}$	1.1615	1.0337	1.3193	1.1658	0.4256	0.7965
	$\boldsymbol{\sigma_U^2}$	0.0565	0.0584	0.0552	0.0564	0.0591	0.0584
	$\boldsymbol{Var(X_{i0})}$	0.3014	0.3125	0.2748	0.3401	0.3082	0.3999
	$\boldsymbol{Var(X_{i1})}$	0.1293	0.1409	0.0806	0.1320	0.1180	0.1127
	$\boldsymbol{Cov(X_{i0}, X_{i1})}$	0.1418	0.1481	0.1172	0.1486	0.1328	0.1509
500	$\boldsymbol{\beta}$	0.7600	0.2933	1.4925	0.7185	0.3688	0.8238
	$\boldsymbol{\gamma_1}$	0.9158	0.4930	0.9424	1.1367	0.1664	0.2381
	$\boldsymbol{\gamma_2}$	0.5822	0.4109	0.7505	0.6714	0.1739	0.3230
	$\boldsymbol{\sigma_U^2}$	0.0254	0.0256	0.0252	0.0253	0.0257	0.0256
	$\boldsymbol{Var(X_{i0})}$	0.1347	0.1412	0.1227	0.1473	0.1352	0.1719
	$\boldsymbol{Var(X_{i1})}$	0.0553	0.0607	0.0351	0.0572	0.0490	0.0478
	$\boldsymbol{Cov(X_{i0}, X_{i1})}$	0.0624	0.0663	0.0517	0.0650	0.0579	0.0654

Tables 3.1 to 3.4 show that there is distribution sensitivity in the JMM-St and JMM-KD models. In terms of JMM-St results from Tables 3.1 and 3.3 show that in general when the true random effect distribution is Normal, estimation becomes less biased and more efficient as expected in JMM-St model. However, there are some unexpected results as well. For instance, when we consider rare response and small sample size, Skew Normal distribution gives less bias in estimation of $\boldsymbol{\beta}$, $\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$ for true random effects distribution. Although it results in small biases values, more efficient estimations; the narrowest IQR for these parameters are observed in Skew Normal Mixture distribution. In large sample size rare response case, JMM-St performs better for true Normal distribution compared to other distributions in terms of bias and efficiency for all parameters. JMM-St model, while estimating $\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$, results in better bias for true Gamma and Gamma Mixture Distributions in common response. The estimation of variance parameters does not generally change for true distribution. In other words, JMM-St is not sensitive to random effects' distribution when the estimating parameters are variance parameters, except Skew Normal distribution in common response. Further analysis may be done for to investigate this situation.

In terms of JMM-KD, from Tables 3.2 and 3.4 it is evident that JMM-KD model performs similar to JMM-St when we compare the estimation of parameters. There are cases where different reactions to the model are observed. To illustrate further, in both rare and common response cases with small sample size, the estimations of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$s result are less biased but not efficient in Skew Normal case. The most efficient outcomes for parameter $\boldsymbol{\beta}$ are observed in Skew Normal Mixture and Normal Distribution cases as expected. For the parameters $\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$, Gamma Mixture distribution shows less bias and is more efficient in all cases. For estimation of variance parameters, all scenarios result in almost the same results for JMM-KD except Skew Normal Distribution. JMM-KD model performs better for estimation of variance parameters in Skew Normal Distribution case.

3.4 DPM Simulation Study

In this section, results of JMM-St with DPM and JMM-KD with DPM simulation studies are presented in two tables for each. In tables, there are comparison of two different concentration parameters (α_0) in which one is Murugiah and Sweeting's equation 2.1 and other one is fixed α_0 value. The considered maximum cluster numbers are C=3 and C=5 and sample size is arranged as 100. The comparison is made by summary statistics introduced in section 3.2.

This simulation study aims to show that how DPM versions models result in terms of relative bias and IQR. The simulation scenarios are same as sensitivity analysis but the true distribution of latent factors is Normal distribution. Outcomes of JMM-St with DPM and JMM-KD with DPM simulation studies are given in Tables 3.5 and 3.6. When we compared 3.1-3.4 with 3.5 and 3.6, results show that considered models are less biased and more efficient than standard versions. In addition, this simulation study revealed maximum number of initial cluster does not have any effect.

Table 3.5: JMM-St with DPM Results

$P(Y=1 \mid x)=0.1$		$\alpha_0 \sim Ga(1,1)$			$\alpha_0=1$		
	Parameter	Relative Bias	IQR	MCSE	Relative Bias	IQR	MCSE
C=3	β	0.1923	1.0419	0.8935	0.1707	1.0083	0.9651
	γ_1	0.2152	0.8354	0.7206	0.1671	0.8586	0.8248
	γ_2	0.1854	0.7069	0.5835	0.0048	0.7031	0.5995
	σ_U^2	0.0163	0.0592	0.0421	0.0171	0.0594	0.0458
	$var(X_{i0})$	-0.0667	0.4309	0.2402	-0.0173	0.4757	0.2110
	$var(X_{i1})$	-0.0297	0.1181	0.0819	-0.0115	0.1201	0.0898
	$cov(X_{i0}, X_{i1})$	-0.0704	0.2820	0.1027	-0.1053	0.2911	0.1003
C=5	β	0.1712	1.0082	0.9595	0.1712	1.0135	0.9632
	γ_1	0.1760	0.8607	0.8228	0.1827	0.8748	0.8293
	γ_2	0.0038	0.6999	0.5955	-0.0045	0.7014	0.5678
	σ_U^2	0.0182	0.0596	0.0458	0.0209	0.0608	0.0466
	$var(X_{i0})$	-0.1266	0.4609	0.2068	-0.1557	0.4437	0.2207
	$var(X_{i1})$	-0.0132	0.1191	0.0889	-0.0244	0.1174	0.0842
	$cov(X_{i0}, X_{i1})$	-0.1613	0.2822	0.0992	-0.1857	0.2517	0.1011
$P(Y=1 \mid x)=0.5$		$\alpha_0 \sim Ga(1,1)$			$\alpha_0=1$		
C=3	β	0.2072	1.5285	0.7458	0.1796	1.4847	0.8733
	γ_1	0.2688	1.8758	0.7843	0.2323	1.7935	0.9326
	γ_2	0.1258	1.1597	0.5836	0.1160	1.1390	0.6354
	σ_U^2	0.0154	0.0572	0.0374	0.0081	0.0569	0.0427
	$var(X_{i0})$	-0.0923	0.3994	0.1937	-0.0814	0.4025	0.1955
	$var(X_{i1})$	-0.0052	0.1216	0.0829	-0.0116	0.1198	0.0898
	$cov(X_{i0}, X_{i1})$	-0.1099	0.2683	0.1193	-0.1193	0.2688	0.0908
C=5	β	0.1762	1.4790	0.8631	0.2098	1.5107	0.7931
	γ_1	0.2360	1.8040	0.9262	0.2098	1.8303	0.9356
	γ_2	0.1052	1.1276	0.6352	0.1092	1.1359	0.5833
	σ_U^2	0.0092	0.0570	0.0428	0.0213	0.0577	0.0427
	$var(X_{i0})$	-0.0908	0.3961	0.1931	-0.1113	0.4099	0.2115
	$var(X_{i1})$	-0.0128	0.1193	0.0888	-0.0303	0.1163	0.0844
	$cov(X_{i0}, X_{i1})$	-0.1738	0.2640	0.0975	-0.1706	0.2713	0.1012

Table 3.6: JMM-KD with DPM Results

$P(Y=1 \mid x)=0.1$		$\alpha_0 \sim Ga(1,1)$			$\alpha_0=1$		
	Parameter	Relative Bias	IQR	MCSE	Relative Bias	IQR	MCSE
C=3	β	0.2275	1.0868	0.9453	0.2279	1.0940	0.9428
	γ_1	0.5102	0.8592	0.7093	0.5157	0.8742	0.7144
	γ_2	0.3282	0.7257	0.6158	0.3279	0.7274	0.6162
	σ_U^2	0.0098	0.0588	0.0426	0.0118	0.0591	0.0431
	$var(X_{i0})$	-0.0357	0.3027	0.2322	0.0564	0.2985	0.2345
	$var(X_{i1})$	0.0278	0.1274	0.1044	0.0181	0.1259	0.1029
	$cov(X_{i0}, X_{i1})$	-0.0087	0.1401	0.1093	-0.0224	0.1378	0.1079
C=5	β	0.2361	1.1030	0.9357	0.2319	1.1372	0.9447
	γ_1	0.6270	0.8866	0.7294	0.5714	0.9543	0.7123
	γ_2	0.5284	0.7289	0.6390	0.3566	0.7367	0.6162
	σ_U^2	0.0152	0.0594	0.0430	0.0251	0.0603	0.0507
	$var(X_{i0})$	-0.0835	0.2913	0.2333	-0.1182	0.2894	0.2564
	$var(X_{i1})$	0.0020	0.1232	0.0992	-0.0038	0.1226	0.1010
	$cov(X_{i0}, X_{i1})$	-0.0559	0.1348	0.1016	-0.0642	0.1339	0.1060
$P(Y=1 \mid x)=0.5$		$\alpha_0 \sim Ga(1,1)$			$\alpha_0=1$		
	Parameter	Relative Bias	IQR	MCSE	Relative Bias	IQR	MCSE
C=3	β	0.1321	1.4547	0.7291	0.1857	1.4819	0.9786
	γ_1	0.1950	1.7755	0.8289	0.2121	1.7787	1.0640
	γ_2	0.1127	1.1385	0.6794	0.0988	1.1225	0.6969
	σ_U^2	0.0141	0.0574	0.0446	0.0110	0.0571	0.0414
	$var(X_{i0})$	-0.0324	0.2931	0.2153	-0.0871	0.2842	0.1706
	$var(X_{i1})$	0.0008	0.1242	0.1020	-0.0155	0.1217	0.0926
	$cov(X_{i0}, X_{i1})$	-0.0890	0.1368	0.0905	-0.1589	0.1332	0.1040
C=5	β	0.1727	1.0411	0.9659	0.2327	1.5144	0.9052
	γ_1	0.1957	0.9070	0.8236	0.2441	1.7674	0.9841
	γ_2	0.0337	0.7061	0.6018	0.2076	1.1942	0.8107
	σ_U^2	0.0231	0.0605	0.0500	-0.0076	0.0561	0.0412
	$var(X_{i0})$	-0.1389	0.2844	0.2372	-0.0277	0.2892	0.1958
	$var(X_{i1})$	-0.0050	0.1219	0.0911	-0.0437	0.1171	0.0878
	$cov(X_{i0}, X_{i1})$	-0.0511	0.1313	0.1031	-0.0613	0.1323	0.1162

CHAPTER 4

APPLICATION

Our aim in this chapter is to illustrate our approach on a real data set and compare the results with the previous analysis of the data using JMM-KD.

4.1 Data Description

JMM-St with DPM and JMM-KD with DPM models, which are detailed in chapter 2, is applied to data set, which is from a study conducted at Zekai Tahir Women Health, Care and Research Hospital in Ankara/TURKEY (Özgü-Erdinç et al. (2015)) and previously has been analyzed by Ozgu-Erdinc (2015) and Kalaylıoğlu and Demirhan (2017). The data set consists of 259 women with Gestational Diabetes Mellitus (GDM) followed over their gestation period. There are two types of treatment that women have. One is diet treatment and other one is insulin treatment. In this application, we only consider insulin treatment mothers as a subset of GDM data set. The data have as a longitudinally observed measurement (log) serum HbA1C (HbA1C; continuous) which is a measurement for glycated (convert bounding with glucose molecule) hemoglobin. HbA1C level is an important indicator for diabetic people. High level of HbA1C increases the odds of having complication related with diabetes increases. Complications at labor risk mother and baby's life. The other measurements of the data set are age (Age; Continuous), Body Mass Index (BMI; Continuous) of a patient before 8 weeks of gestation, number of pregnancy mother had before current one, parity (P;Continuous), Macrosomia History (MSH; Binary), GDM History (GDMH; Binary), GDM Family History (GDMFH; Binary), Hemoglobin Concentra-

tion > 13 g/dl (Hb13; Binary) and the status of Obstetric Labor Complication (OLC). The outcome variable is OLC. There are binary responses in data set and the reference level of binary variables is set as absence. The aim is to investigate the effect of HbA1C profile over time on the complication at birth.

4.2 Exploratory Analysis

In this section, the data set is numerically presented and visualized. The descriptive statistics and plots are given apart from the main analysis in order to understand the GDM Data.

Table 4.1: GDM Data Average Values Comparison

	Women without OLC (min.,max.)	**Women with OLC (min.,max.)**
Age	33.35 (20, 45)	32.13 (20, 45)
BMI	32.25 (21, 52)	33.72 (21, 46)
HbA1C Level	(4.2, 7.8)	(4.3, 8.8)

Summary of the statistics of continuous variables of GDM data such as age, BMI and HbA1C level are given in Table 4.1. Table also contains minimum and maximum observed values of variables. As can be seen from Table 4.1, women in the data set have same age interval and almost same average age regardless of their OLC condition. The other inference about the data set is that in general, women with OLC have higher BMI and HbA1C level compared to women without OLC.

The relation between binary observed variables such as Macrosomia History, GDM History, GDM Family History and Hb13, and OLC are shown in contingency tables. In order to see independence between factors and having OLC, χ^2 test is applied each cross table and results are given. Also, Odds Ratios (OR) and confidence intervals (CI) of OR are given.

Table 4.2: χ^2 and Odds Ratio test for association of having OLC and factors

Factor	**Factor Levels**	**Women without OLC**	**Women with OLC**	**Total**	**P-Value**	**OR 95% CI**
		count (N)	*count (N)*			
MSH	*without (0)*	91	23	114	**0.135**	**0.23 (0.030, 1.80)**
	with (1)	17	1	18		
	Total	108	24	132		
GDMH	*without (0)*	80	13	93	**0.053**	**2.44 (0.98, 6.09)**
	with (1)	28	11	39		
	Total	108	24	132		
GDMFH	*without (0)*	88	18	106	**0.47**	**1.47 (0.618, 3.47)**
	with (1)	20	6	26		
	Total	108	24	132		
Hb13	*no (0)*	75	14	89	**0.293**	**1.62 (0.65,4.03)**
	yes (1)	33	10	44		
	Total	108	24	132		

Frequencies of the levels of categorical variables are given in Table 4.2. χ^2 test is applied each cross tables for testing the association between OLC and factors. Results revealed that GDM History has relation with OLC slightly (p=0.053>α=0.05) for GDM data set. The other variables tested have no association with OLC since calculated p-values do not show significant value for χ^2 test. Moreover, when we look to CIs of OR we can conclude that ORs are not significantly different than 1 since all CIs cover 1. Therefore, there is no significant association between factors and having OLC.

Although some variables are not related with odds of having OLC, these variables were thought to be in model and analysis (Özgü-Erdinç et al. (2015)). According to the literature, all the variables have an impact on having OLC.

Figure 4.1 shows (log) serum HbA1C levels of mother over gestation period without OLC and with OLC, respectively. The visual comparison is for primary investigation of the relation between HbA1C and the odds of having OLC. The main analysis is done for predicting the odds of having OLC with using the longitudinal covariate (log) serum HbA1C.

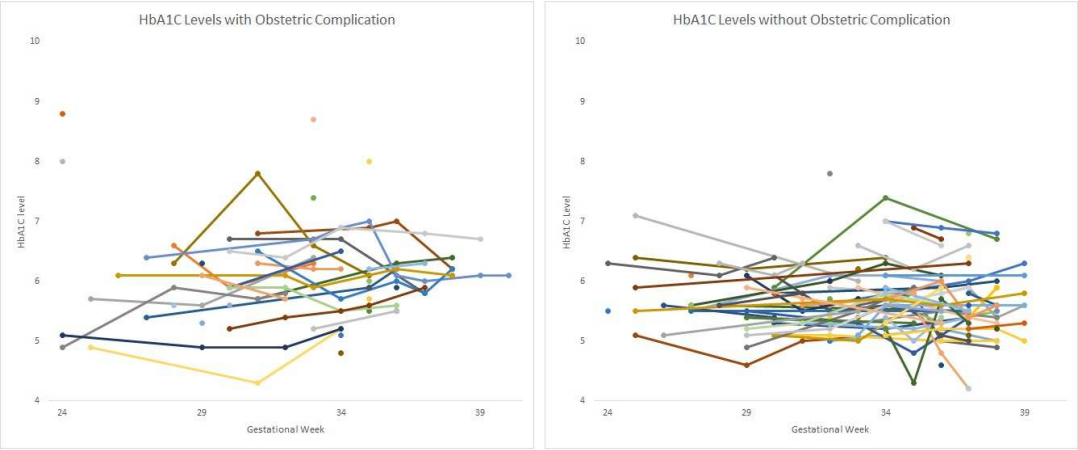

Figure 4.1: HbA1C Level vs Gestational Week without and with OLC

4.3 Data Analysis

In this section, GDM data analysis is covered. For analysis, four different models are considered for efficiency comparison. The models are JMM-KD, JMM-St, JMM-KD with DPM and JMM-St with DPM. The standard versions of the model are used unchanged as in the paper Kalaylıoğlu and Demirhan (2017). JMM-St with DPM and JMM-KD with DPM versions of models are selected in the light of simulation study 3.4.

The GDM data set is analyzed and posterior medians are given. Number of longitudinal measurements of HbA1C changing between only 1 and 8 per subject. As can be seen in Figure 4.1, this measurement does not have a functional form over week. Joint mixed model is applied and simple regression part of the model equation 1.2. For the estimation of the probability of having OLC the first part of the model equation 1.1 is used in which longitudinal log serum HbA1C are accounted for by including covariates that are the random coefficients from the longitudinal model. The considered mixed effects response models are,

$$\boldsymbol{probit}(\boldsymbol{P}(\boldsymbol{OLC}_i = 1 \mid \boldsymbol{covariates}_i)) = \boldsymbol{\beta_1} + \boldsymbol{\beta_2}\boldsymbol{Age}_i + \boldsymbol{\beta_3}\boldsymbol{BMI}_i + \boldsymbol{\beta_4}\boldsymbol{P}_i$$
$$+ \boldsymbol{\beta_5}\boldsymbol{GDMH}_i + \boldsymbol{\beta_6}\boldsymbol{GDMFH}_i + \boldsymbol{\beta_7}\boldsymbol{Hb13}_i + \boldsymbol{\beta_8}\boldsymbol{MSH}_i + \boldsymbol{\gamma_1}\boldsymbol{X}_{i1} + \boldsymbol{\gamma_2}\boldsymbol{X}_{i2}$$

$$\boldsymbol{log}(\boldsymbol{HbA1C}_{ij}) = \boldsymbol{X}_{i1} + \boldsymbol{X}_{i2}\left(\frac{\boldsymbol{t}_{ij} - \bar{\boldsymbol{t}}}{\boldsymbol{S}_t}\right) + \boldsymbol{U}_{ij}$$

where $\bar{\boldsymbol{t}}$ and $\boldsymbol{S}_t$ are mean and standard deviation of of all $\boldsymbol{t}_{ij}$ $\boldsymbol{i} = \boldsymbol{1, ..., 256}$ and $\boldsymbol{j} = \boldsymbol{1, ..., n_i}$. Probit link function is considered because of computational reasons (Albert and Chib, (1993)). Standardized version of time indicator is used since there are different observation numbers taken at different time for each study individual (irregular longitudinal data).

$\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$ are relation coefficients which indicates the association between the odds of *OLC* and *HbA1C* profile over gestation period of mother. Especially, if $\boldsymbol{\gamma_2}$ is equal to zero, it means that *OLC* status is not related with the slope of mother's *log serum HbA1C* across the gestation. $\boldsymbol{\gamma_1}$ gives mean relation between *OLC* and *log serum*

HbA1C. The main focus in this analysis is $\boldsymbol{\gamma_2}$ because it leads us whether to use *log serum HbA1C* on predicting probability of having OLC or not.

The priors that were used in Kalaylıoğlu and Demirhan (2017) are used in this application for making easy comparison between estimations. The convergence diagnostics are given in Appendix F and G and convergence ensured. The first 100,000 iterations are discarded and every 100th of next 1,400,000 iterations resulting in 14,000 iterations hired for posterior inferences. Table 4.3 shows posterior medians and 95% of posterior intervals. To find the best in models, $\boldsymbol{-2 * log(likelihood)}$ is calculated as a model indicator and given for each model.

Table 4.3: GDM Data Anaysis Comparison

Covariates	**JMM-KD**	**JMM-KD with DPM**	**JMM-St**	**JMM-St with DPM**
Intercept	-7.81 (-13.50, -2.51)	-6.58 (-11.35, -1.956)	-9.93 (-17.28, -2.84)	-7.60 (-13.23, -2.06)
Age	-0.002 (-0.06, 0.06)	0.003 (-0.05, 0.06)	-0.003 (-0.07, 0.06)	0.002 (-0.06, 0.06)
BMI	0.003 (-0.06, 0.06)	0.005 (-0.05, 0.06)	0.006 (-0.07, 0.08)	0.004 (-0.06, 0.06)
Parity	0.23 (-0.02, 0.50)	0.23 (-0.02, 0.5)	0.26 (-0.02, 0.55)	0.23 (-0.02, 0.49)
GDMH	0.43 (-0.20, 1.05)	0.41 (-0.21, 1.03)	0.54 (-0.14, 1.24)	0.43 (-0.20, 1.05)
GDMFH	0.08 (-0.66, 0.78)	0.10 (-0.62, 0.80)	0.05 (-0.76, 0.82)	0.08 (-0.65, 0.78)
Hb13	0.52 (-0.09, 1.15)	0.50 (-0.11, 1.11)	0.56 (-0.11, 1.28)	0.51 (-0.1, 1.13)
MSH	-1.35 (-2.66, -0.26)	-1.35 (-2.62, -0.26)	-1.46 (-2.89, -0.28)	-1.35 (-2.63,-0.27)
$\boldsymbol{X_1}$	3.44 (0.73, 6.39)	2.66 (0.43, 5.02)	4.60 (0.77, 8.50)	3.31 (0.43, 6.22)
$\boldsymbol{X_2}$	0.37 (-5.83, 6.50)	-0.01 (-6.12, 6.19)	0.85 (-5.37, 7.14)	0.14 (-6.07,6.34)
$\boldsymbol{-2 * logL}$	-639.7	-618.3	-660	-660

*Note: The interval stands for 95% posterior interval of the posterior median of parameter.

It can be seen from Table 4.3 the posterior inferences have narrower posterior intervals in DPM models. Both JMM-St with DPM and JMM-KD with DPM have more efficient estimates than their standard versions. The best models are JMM-St and JMM-St with DPM since the lowest $\boldsymbol{-2 * logL}$ value resulted.

$\boldsymbol{MSH}$ and $\boldsymbol{X_1}$ variables are resulted significant in the model. In other words, Macrosomia History has negative affect on having obstetric labor complication. The mean-

ing of significance of $\boldsymbol{X_1}$ is that the average HbA1C level during gestation period is positively related with probablity of having OLC in all models.

CHAPTER 5

CONCLUSION

Joint mixed model is an appealing approach in medical research where it is critical to estimate the odds of a fatal complication will occur to a patient given the covariate profile observed over time. In this thesis, the standard Bayesian analysis of the model in which the error variance and random effects' variance-covariance matrix were apriori modeled independently with inverse-gamma and inverse-Wishart distributions respectively and joint apriori modeling of variance components via Generalized Multivariate Log-Gamma (G-MVLG) distribution are considered.

The conducted simulation studies have some limitations in distribution types and sample sizes. We have considered only six different distributions which are bivariate distributions.In addition, there are two different sample size, 100 as the small sample size and 500 as the large sample size.

Firstly, the sensitivity of models on the true distribution of latent factors are considered. Simulation study is then conducted for true distributions Normal, Normal Mixture, Skew-Normal, Skew-Normal Mixture, Gamma and Gamma Mixture. Results showed that the quality of posterior inference changes for different true distributions. Therefore, it can be said that the models have sensitivity on true distributions. Moreover, the study showed that both models give better results when applied to large sample sizes rather than small ones. As mentioned in Chapter 3, estimation is oddly either good or bad for some of the true distributions of random effects, which require further analysis on distributions and models. In addition, in some common response cases, the estimations are not as accurate as rare response cases. This situation also

need further investigation.

Secondly, Dirichlet Process Mixture is used for latent factors' assumption for robust inferences on parameters. It is known that assuming a distribution on latent factors affects the posterior estimates when the assumption has more information. In order to avoid this consequence, DPM model was proposed in this study for latent factors' distribution assumption. We conducted simulation study on both two models and showed that DPM makes model more efficient and robust compared to their standard forms.

The new models were applied on GDM data set. Joint Mixed model application of this data set had been done in Kalaylıoğlu and Demirhan (2017). Our new models and previous standard models were compared and it was concluded that models with DPM results in almost the same posterior median inferences and a narrower 95% posterior intervals for the considered parameters. In the aspect of not assuming a distribution for random effects, DPM models are more robust than models, which have assumption for random effects' distribution.

For future work, we are planning to develop a DPM prior for error variance term $\boldsymbol{\sigma_U^2}$. We will reduce the number of assumption by one on models with DPM on $\boldsymbol{\sigma_U^2}$. Secondly, further analysis will be done for the specific situation mentioned previously and report the findings.

REFERENCES

[1] J. H. Albert and S. Chib, "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, vol. 88, no. 422, pp. 669–679, 1993.

[2] C. Archambeau and M. Verleysen, "Robust Bayesian clustering," *Neural Networks*, vol. 20, no. 1, pp. 129–138, 2007.

[3] A. Azzalini and A. Dalla Valle, "The multivariate skew-normal distribution," *Biometrika*, vol. 83, no. 4, pp. 715–726, 1996.

[4] P. Congdon, *Bayesian Statistical Modelling*. Wiley Series in Probability and Statistics, 2nd editio ed., 2007.

[5] M. J. Daniels and M. Pourahmadi, "Bayesian Analysis of Covariance Matrices and Dynamic Models for Longitudinal Data," *Biometrika Trust*, vol. 89, no. 3, pp. 553–566, 2002.

[6] R. De la Cruz-Mesia, F. A. Quintana, and P. Müller, "Semiparametric Bayesian Classification with Longitudinal Markers," *Journal of the Royal Statistical Society*, vol. 56, no. 2, 2007.

[7] H. Demirhan and C. Hamurkaroglu, "On a multivariate log-gamma distribution and the use of the distribution in the Bayesian analysis," *Journal of Statistical Planning and Inference*, vol. 141, no. 3, pp. 1141–1152, 2011.

[8] Erning Li, Daowen Zhang and M. Davidian, "Likelihood and Pseudo-likelihood Methods for Semiparametric Joint Models for a Primary Endpoint and Longitudinal Data," *Computational Statistics and Data Analysis*, vol. 51, no. 12, pp. 5776–5790, 2007.

[9] T. E. Hanson, A. J. Branscum, and W. O. Johnson, "Bayesian Nonparametric Modeling and Data Analysis: An Introduction," *Handbook of Statistics*, vol. 25, pp. 245–278, jan 2005.

[10] F. Heinzl, L. Fahrmeir, and T. Kneib, "Additive mixed models with Dirichlet process mixture and P-spline priors," *AStA Advances in Statistical Analysis*, vol. 96, no. 1, pp. 47–68, 2012.

[11] J. Horrocks and M. J. van Den Heuvel, "Prediction of pregnancy: A joint model for longitudinal and binary data," *Bayesian Analysis*, vol. 4, no. 3, pp. 523–538, 2009.

[12] Y. Huang, X. J. Hu, and G. A. Dagne, "Jointly modeling time-to-event and longitudinal data: A Bayesian approach," *Statistical Methods and Applications*, vol. 23, no. 1, pp. 95–121, 2014.

[13] H. Ishwaran and M. Zarepour, "Exact and approximate sum representations for the Dirichlet process," *Can. J. Stat.*, vol. 30, no. 2, pp. 269–283, 2002.

[14] H. Ishwaran and L. F. James, "Approximate Dirichlet process computing in finite normal mixtures: Smoothing and prior information," *Journal of Computational and Graphical Statistics*, vol. 11, no. 3, pp. 508–532, 2002.

[15] A. Jara, T. E. Hanson, and E. Lesaffre, "Robustifying generalized linear mixed models using a new class of mixtures of multivariate Polya trees," *Journal of Computational and Graphical Statistics*, vol. 18, no. 4, pp. 838–860, 2009.

[16] Z. Kalaylioglu and H. Demirhan, "A joint Bayesian approach for the analysis of response measured at a primary endpoint and longitudinal measurements," *Statistical Methods in Medical Research*, vol. 26, no. 6, pp. 2885–2896, 2017.

[17] E. Li, N. Wang, and N.-Y. Wang, "Joint Models for a Primary Endpoint and Multiple Longitudinal Covariate Processes," *Biometrics*, vol. 63, no. 4, pp. 1068–1078, 2007.

[18] E. Li, D. Zhang, and M. Davidian, "Conditional estimation for generalized lin-

ear models when covariates are subject-specific parameters in a mixed model for longitudinal parameters," *Biometrics*, vol. 60, no. March, pp. 1–7, 2004.

[19] Y. Li, X. Lin, and P. Müller, "Bayesian Inferences in Semiparametric Meixed Models for Longitudinal data," *Biometrics*, vol. 66, no. 1, pp. 70–78, 2010.

[20] S. Litière, A. Alonso, and G. Molenberghs, "The impact of a misspecified random-effects distribution on the estimation and the performance of inferential procedures in generalized linear mixed models," *Statistics in Medicine*, vol. 27, no. 16, pp. 3125–3144, 2008.

[21] C. E. McCulloch and J. M. Neuhaus, "Misspecifying the Shape of a Random Effects Distribution: Why Getting It Wrong May Not Matter," *Statistical Science*, vol. 26, no. 3, pp. 388–402, 2011.

[22] S. Murugiah and T. Sweeting, "Selecting the precision parameter prior in Dirichlet process mixture models," *Journal of Statistical Planning and Inference*, vol. 142, no. 7, pp. 1947–1959, 2012.

[23] I. Ntzoufras, *Bayesian Modeling Using WinBUGS*. JOHN WILEY & SONS, INC., 2009.

[24] A. S. Ozgu-Erdinc, C. Iskender, D. Uygur, A. Oksuzoglu, K. D. Seckin, M. I. Yeral, Z. I. Kalaylioglu, A. Yucel, and A. N. Danisman, "One-hour versus two-hour postprandial blood glucose measurement in women with gestational diabetes mellitus: which is more predictive?," *Endocrine*, vol. 52, no. 3, pp. 561–570, 2016.

[25] Peter Congdon, *Applied Bayesian Hierarchical Methods*, vol. 39. TAYLOR & FRANCIS GROUP, 2010.

[26] F. A. Quintana, W. O. Johnson, E. Waetjen, and E. Gold, "Bayesian nonparametric Longitudinal Data Analysis," vol. 111, no. 515, pp. 1168–1181, 2016.

[27] B. W. Schmeiser and R. Lal, "Bivariate Gamma Random Vectors," vol. 30, no. 2, pp. 355–374, 1982.

[28] O. E. Smith, S. I. Adelfang, and J. D. Tubbs, "A Bivariate Gamma Probability Distribution with Application to Gust Modeling," *NASA*, vol. 1982, 1982.

[29] D. Spiegelhalter, A. Thomas, N. Best, and R. Way, "WinBUGS User Manual," *Components*, vol. 2, no. January, pp. 1–60, 2003.

[30] S. Walker and J. Wakefield, "Population Models with a Nonparametric Random Coefficient Distribution," *The Indian Journal of Statistics*, vol. 60, no. 1, pp. 196–214, 1998.

[31] C. Y. Wang, N. Wang, and S. Wang, "Regression Analysis When Covariates Are Regression Parameters of a Random Effects Model for Observed Longitudinal Measurements," *Biometrics*, vol. 56, no. 2, pp. 487–495, 2000.

APPENDIX A

CONVERGENCE DIAGNOSTICS OF JMM-ST MODEL

In this chapter, convergence diagnostics are presented for each interested parameters. Thinning, burn-in (warm up) iteration number and MC chain size in chapter 3 Simulation Study were determined in the light of these results. There are three images to observe the convergence in the posterior density. They are Auto-correlation between iterations, BGR (Brooks-Gelman-Rubin) (Gelman and Rubin (1992))diagnostic and Posterior density. These images are created with the help of OpenBUGS. The simulation procedure is same in simulation study part. There are two chains of initials for BGR Diagnostics, sample size for each chain is 50,000 which means 100,000 total samples. In each sample, sample size is $\boldsymbol{n = 100}$. Visual outputs of thinning=10 and burn-up=5,000 are given. A special note that in Bivariate Normal Mixture distribution thinning is set equal 50 and burn-up is set 10,000 and sample size of each chain is 100,000 for convergence.

A.1 Bivariate Gamma Distribution

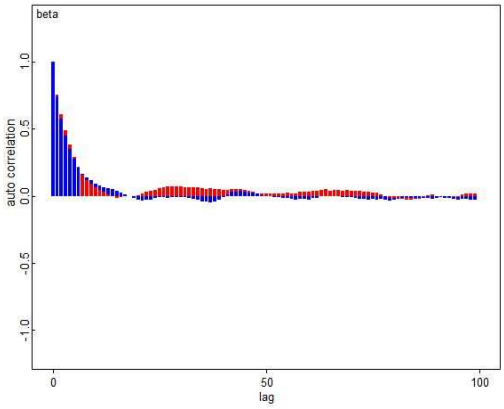

Figure A.1: Auto-Correlation of Parameter β

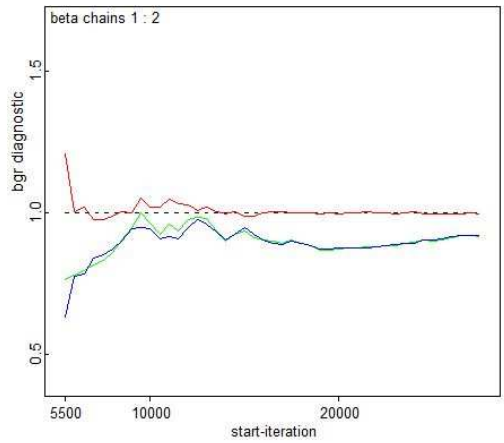

Figure A.2: BGR Diagnostic of Parameter β

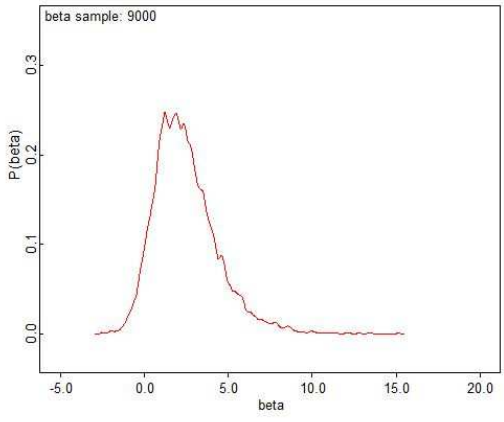

Figure A.3: Posterior Density of Parameter β

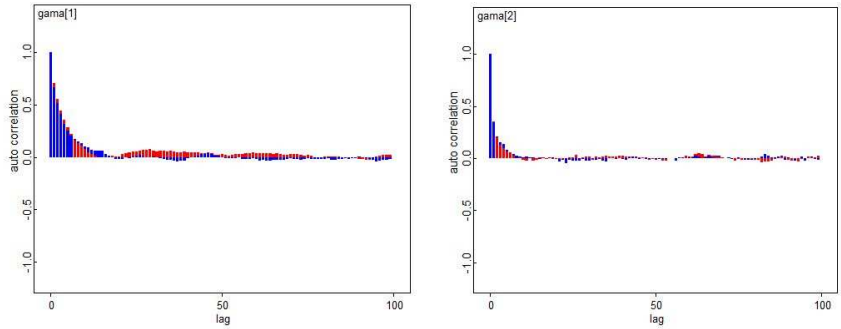

Figure A.4: Auto-Correlation of Parameter γ

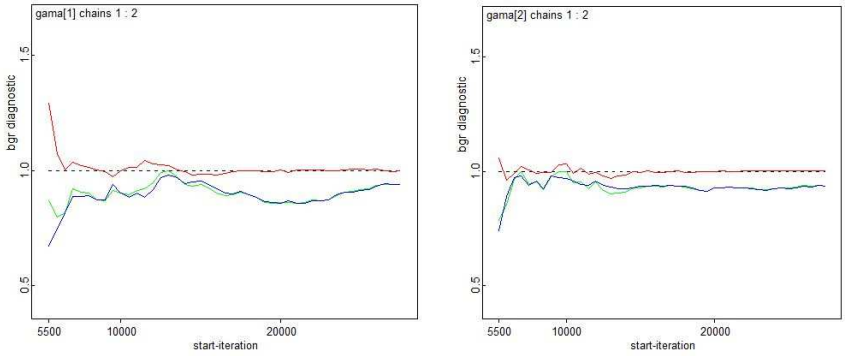

Figure A.5: BGR Diagnostic of Parameter γ

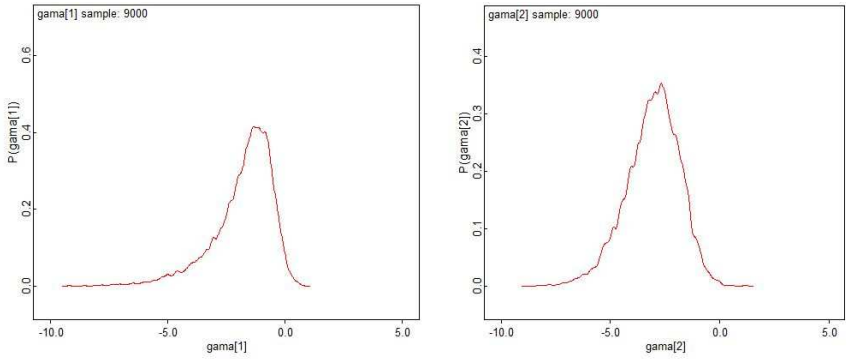

Figure A.6: Posterior Density of Parameter γ

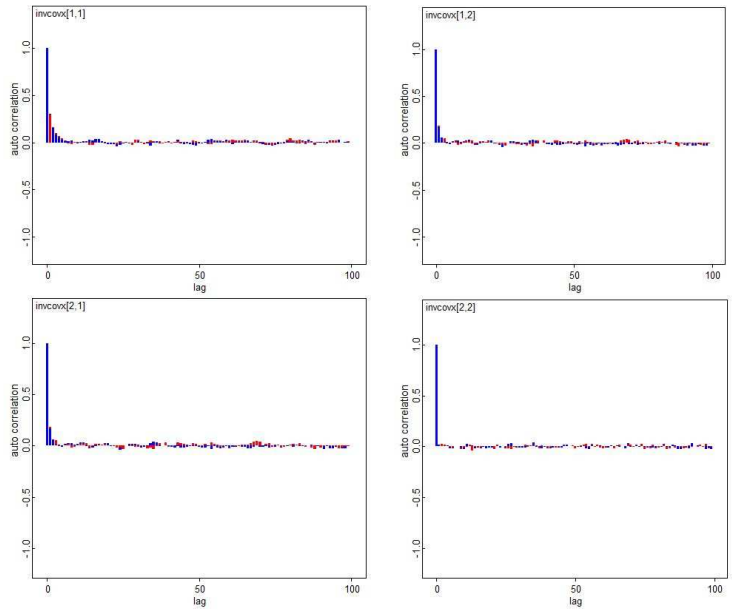

Figure A.7: Auto-Correlation of Parameter inverse $\boldsymbol{\Sigma_X}$

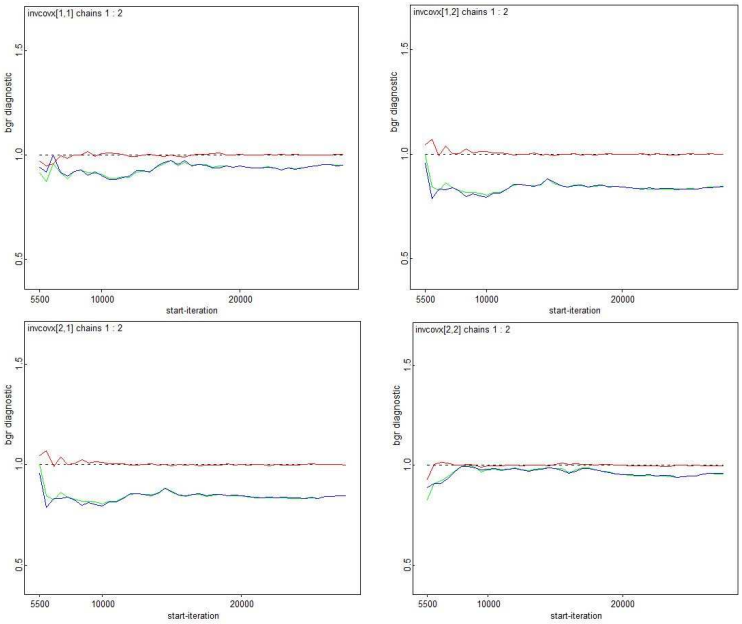

Figure A.8: BGR Diagnostic of Parameter inverse $\boldsymbol{\Sigma_X}$

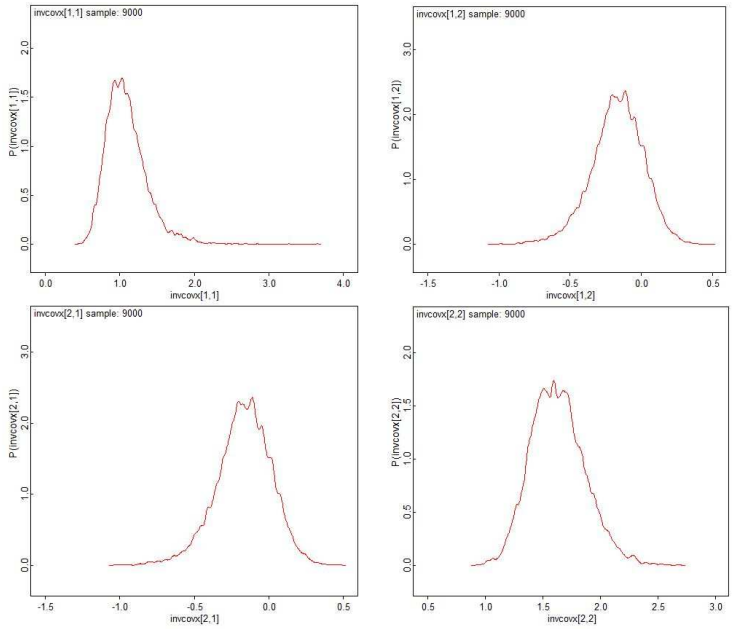

Figure A.9: Posterior Density of Parameter inverse $\boldsymbol{\Sigma_X}$

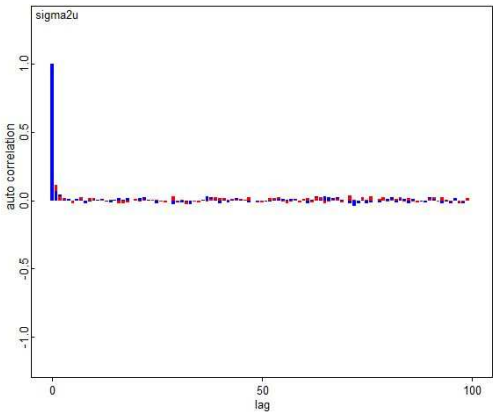

Figure A.10: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

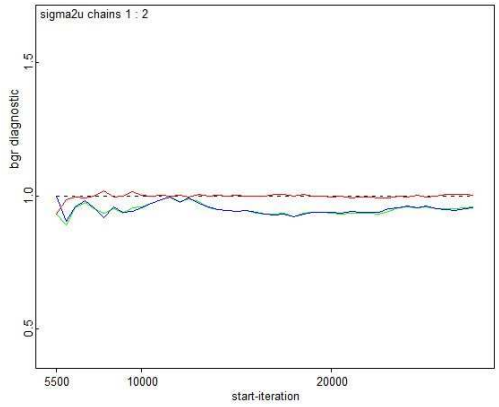

Figure A.11: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

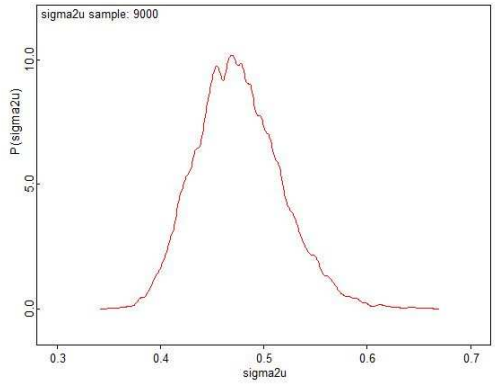

Figure A.12: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

A.2 Bivariate Gamma Mixture Distribution

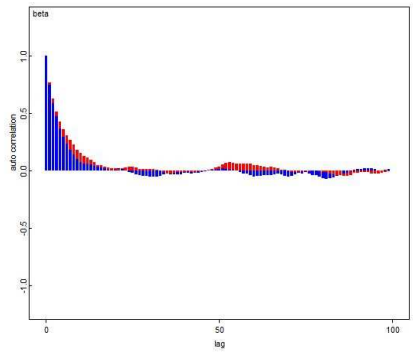

Figure A.13: Auto-Correlation of Parameter β

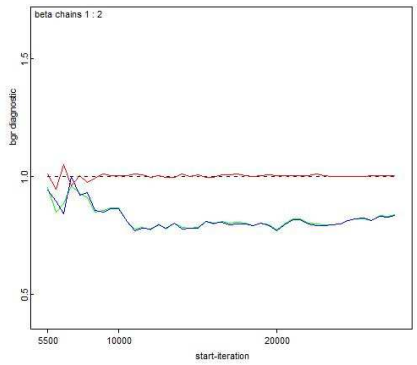

Figure A.14: BGR Diagnostic of Parameter β

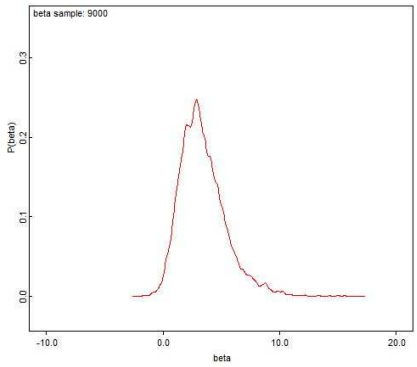

Figure A.15: Posterior Density of Parameter β

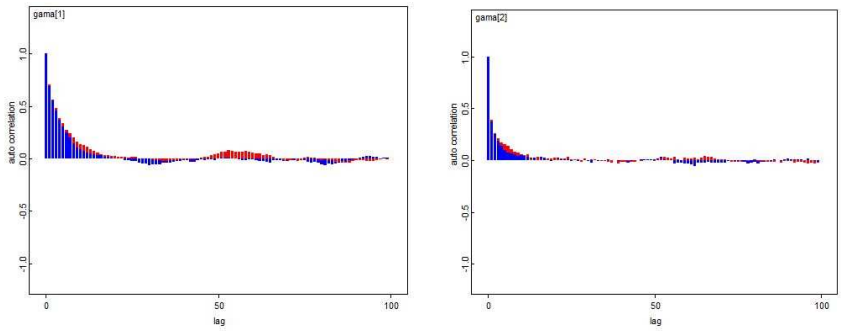

Figure A.16: Auto-Correlation of Parameter γ

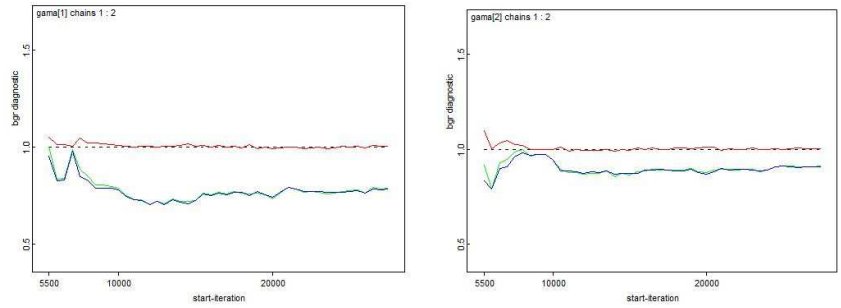

Figure A.17: BGR Diagnostic of Parameter γ

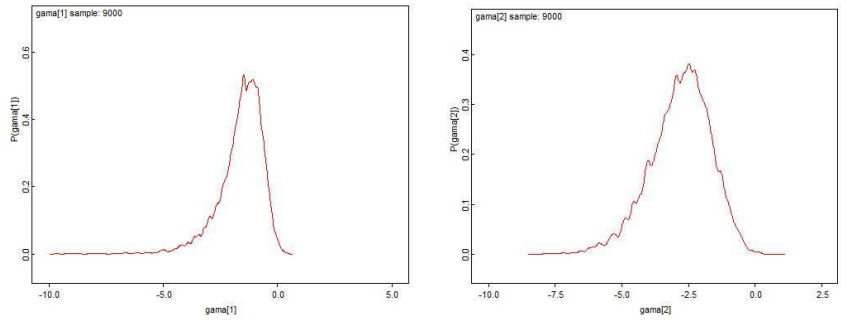

Figure A.18: Posterior Density of Parameter γ

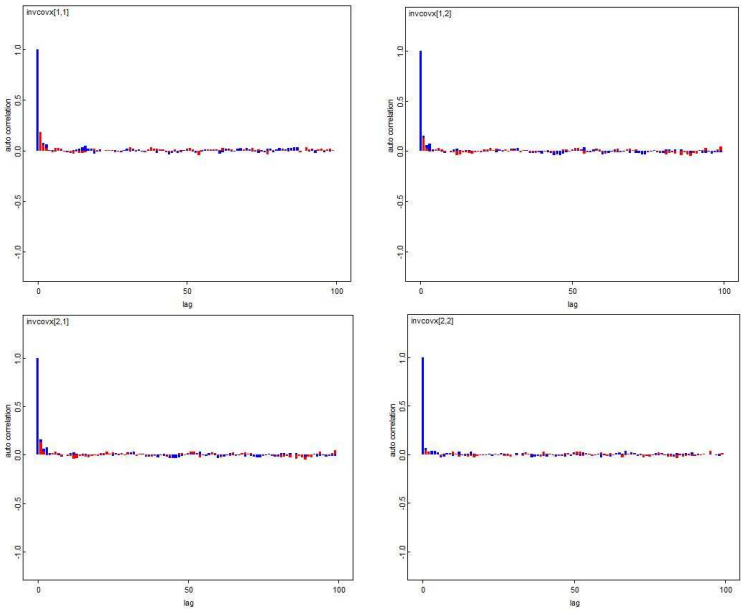

Figure A.19: Auto-Correlation of Parameter inverse Σ_X

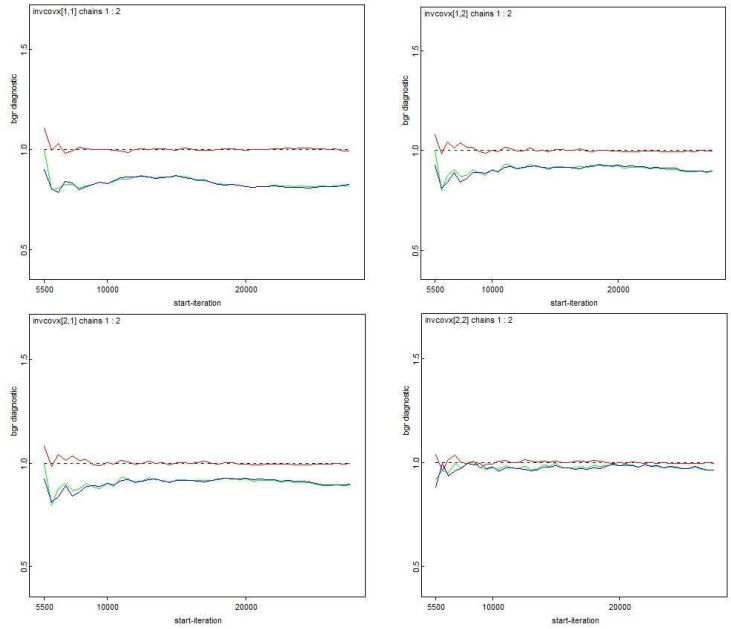

Figure A.20: BGR Diagnostic of Parameter inverse Σ_X

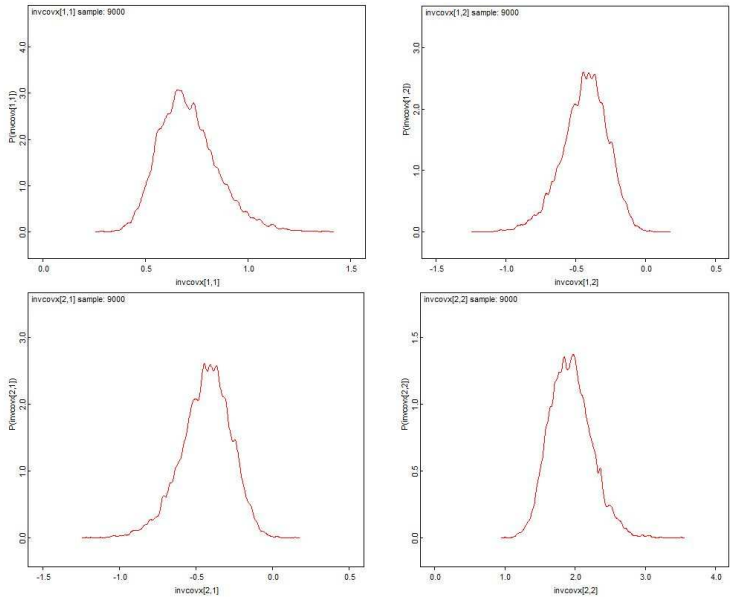

Figure A.21: Posterior Density of Parameter inverse $\boldsymbol{\Sigma_X}$

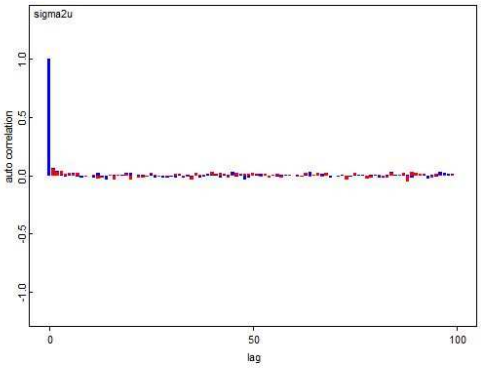

Figure A.22: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

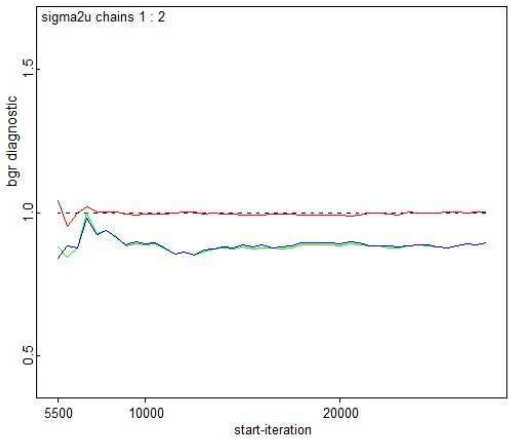

Figure A.23: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

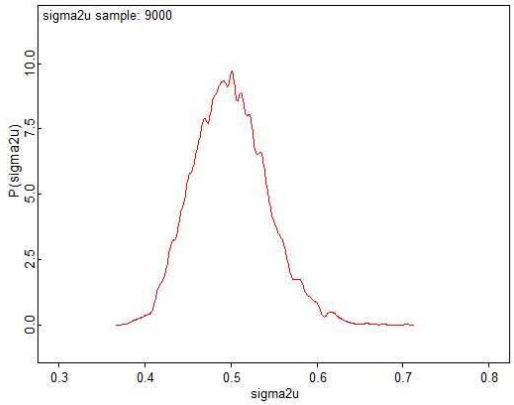

Figure A.24: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

A.3 Bivariate Normal Mixture Distribution

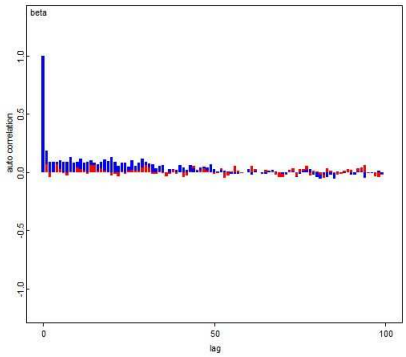

Figure A.25: Auto-Correlation of Parameter β

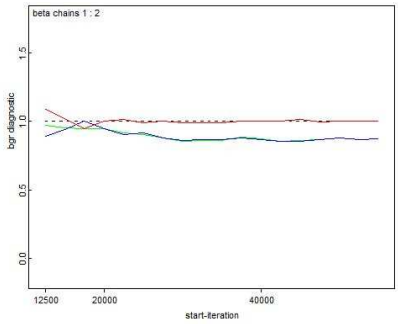

Figure A.26: BGR Diagnostic of Parameter β

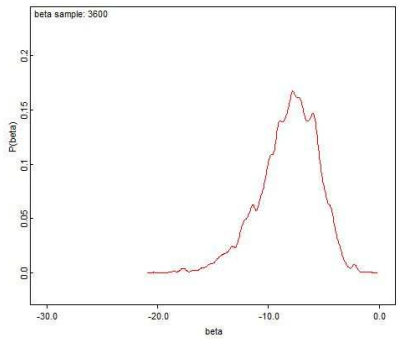

Figure A.27: Posterior Density of Parameter β

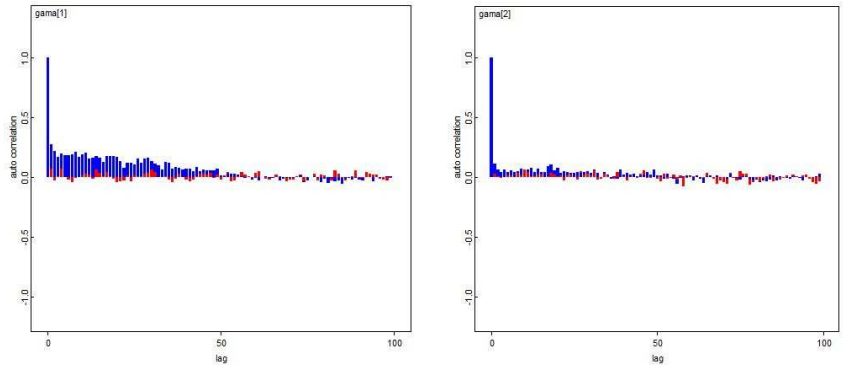

Figure A.28: Auto-Correlation of Parameter γ

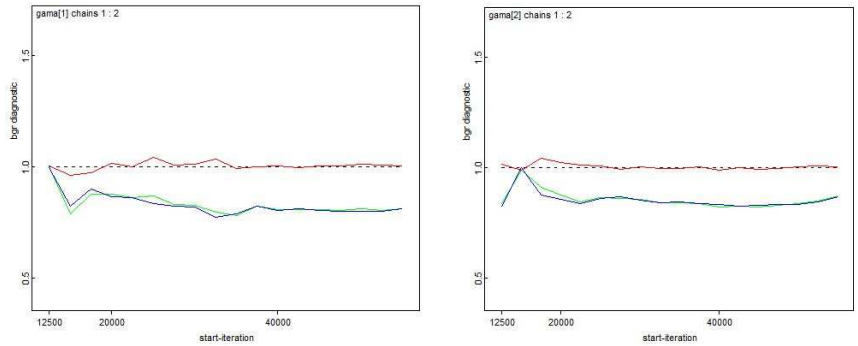

Figure A.29: BGR Diagnostic of Parameter γ

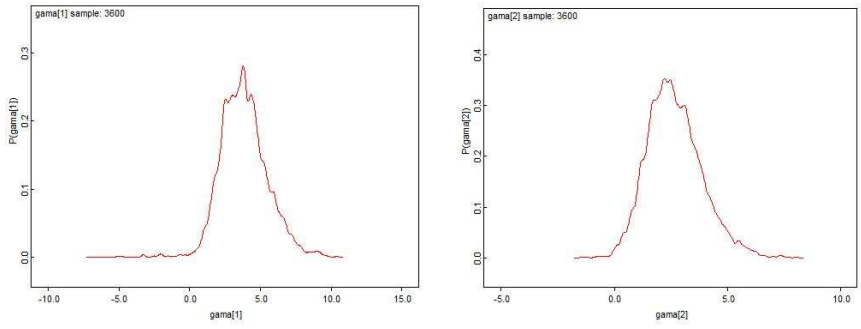

Figure A.30: Posterior Density of Parameter γ

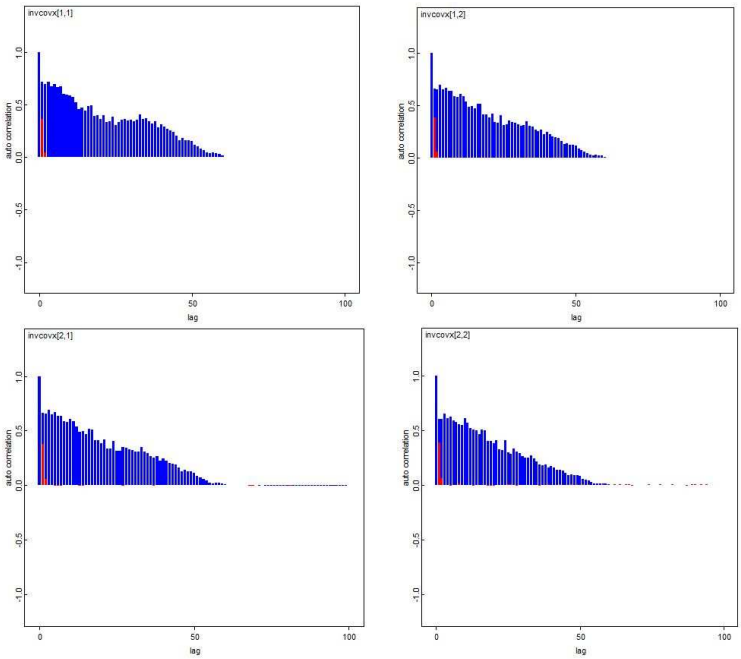

Figure A.31: Auto-Correlation of Parameter inverse Σ_X

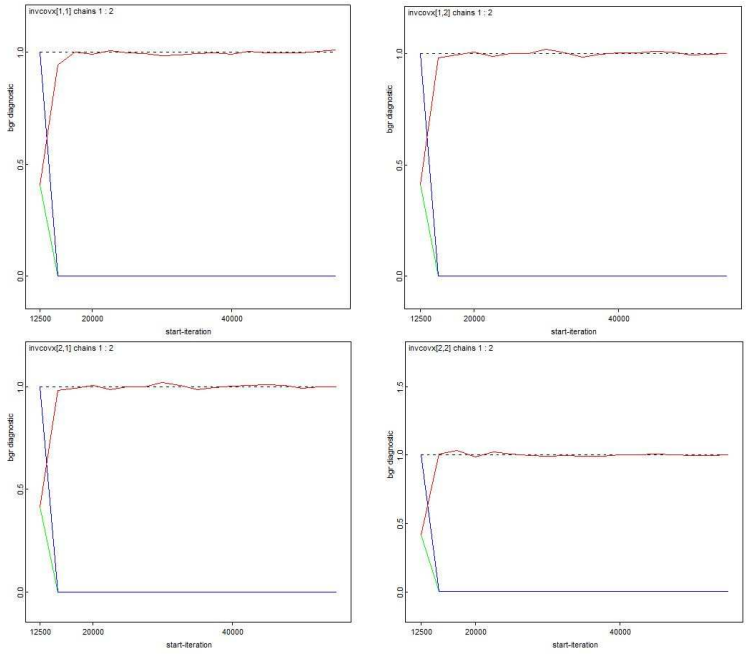

Figure A.32: BGR Diagnostic of Parameter inverse Σ_X

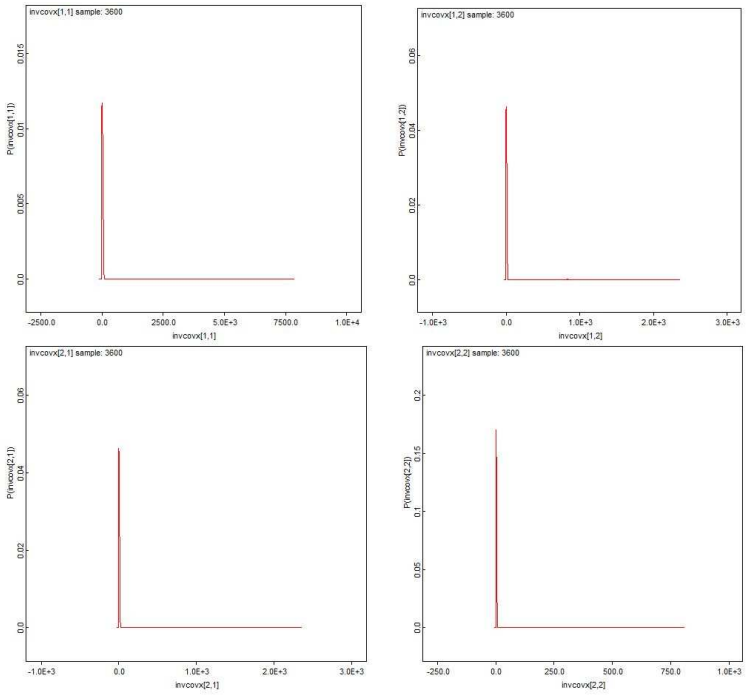

Figure A.33: Posterior Density of Parameter inverse $\boldsymbol{\Sigma_X}$

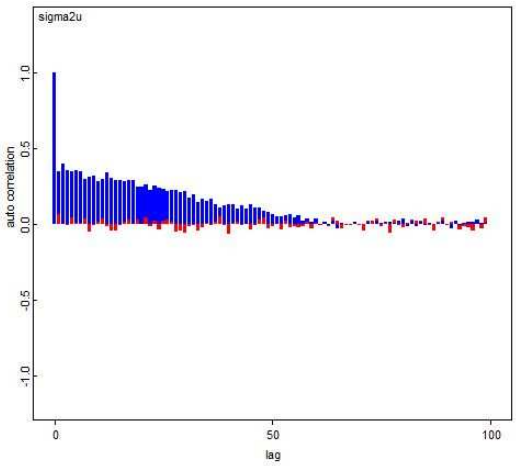

Figure A.34: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

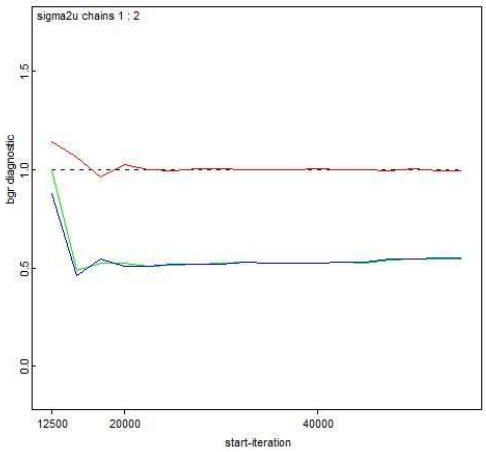

Figure A.35: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

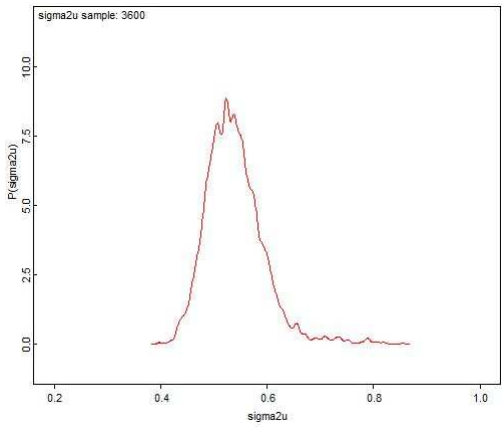

Figure A.36: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

A.4 Bivariate Skew-Normal Distribution

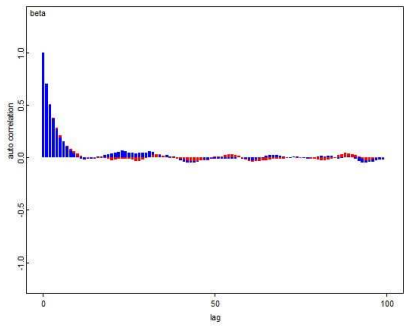

Figure A.37: Auto-Correlation of Parameter β

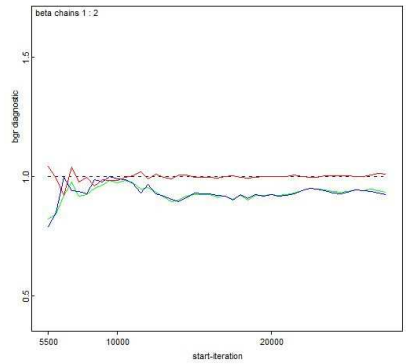

Figure A.38: BGR Diagnostic of Parameter β

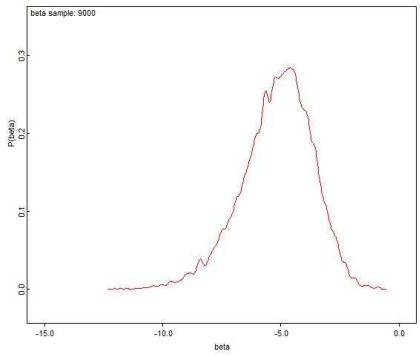

Figure A.39: Posterior Density of Parameter β

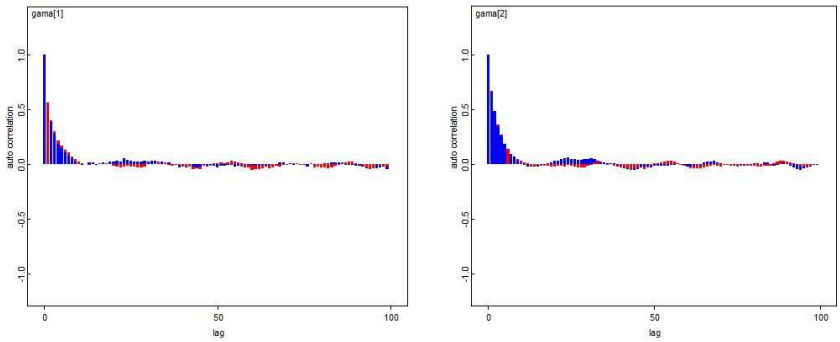

Figure A.40: Auto-Correlation of Parameter γ

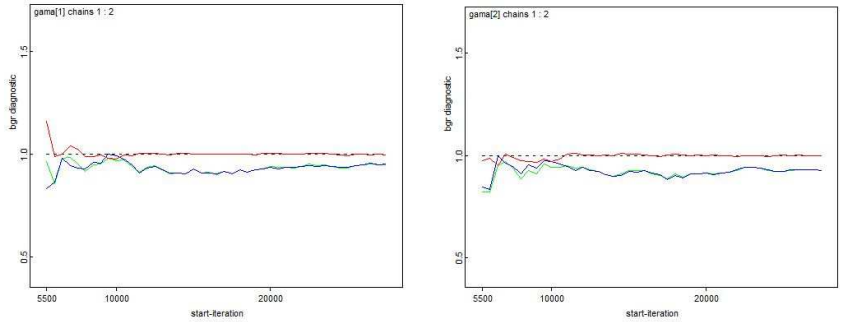

Figure A.41: BGR Diagnostic of Parameter γ

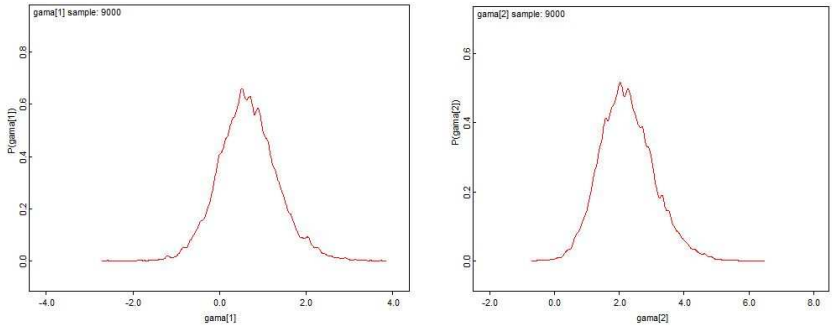

Figure A.42: Posterior Density of Parameter γ

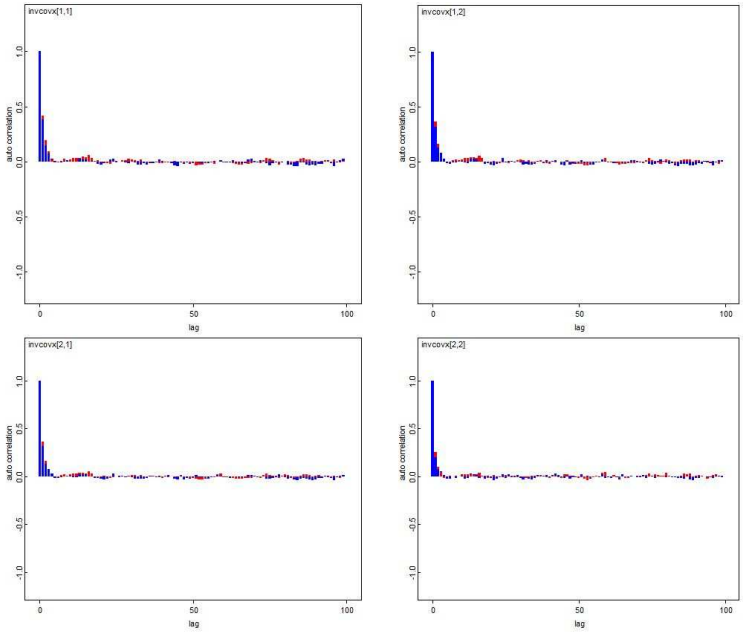

Figure A.43: Auto-Correlation of Parameter inverse Σ_X

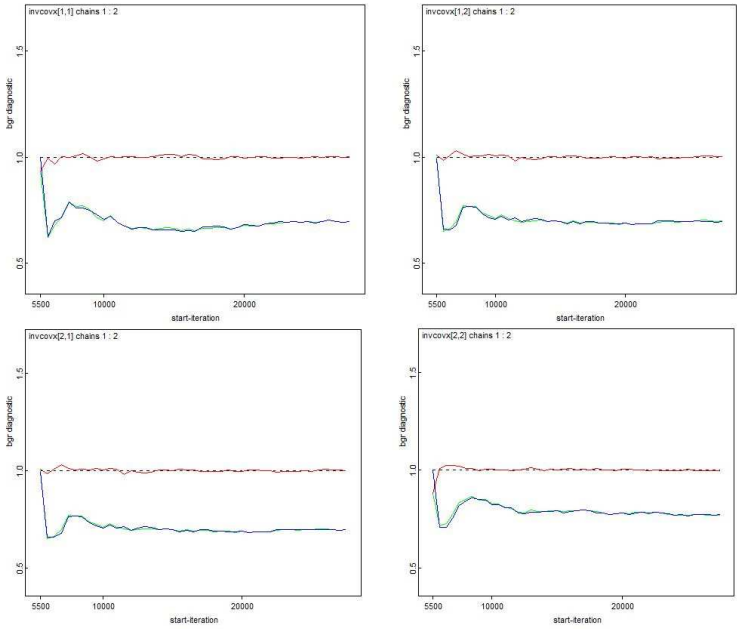

Figure A.44: BGR Diagnostic of Parameter inverse Σ_X

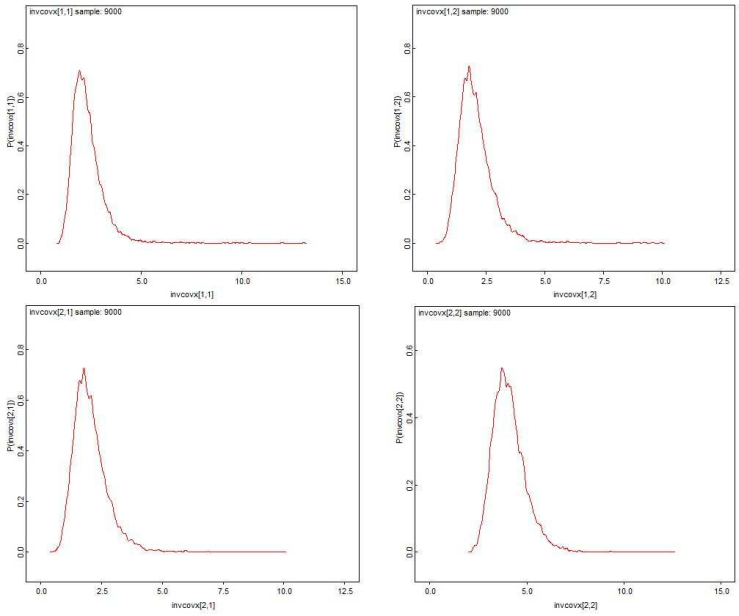

Figure A.45: Posterior Density of Parameter inverse $\boldsymbol{\Sigma_X}$

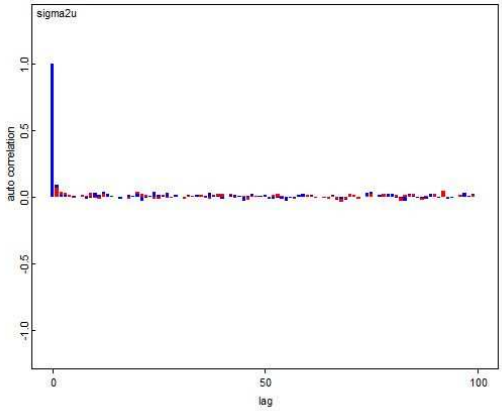

Figure A.46: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

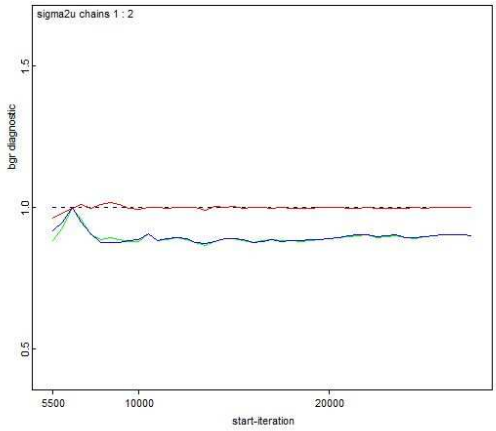

Figure A.47: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

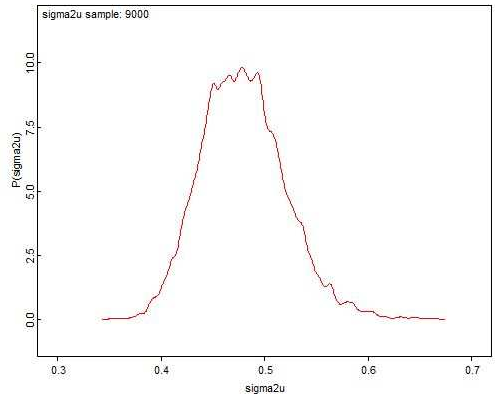

Figure A.48: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

A.5 Bivariate Skew-Normal Mixture Distribution

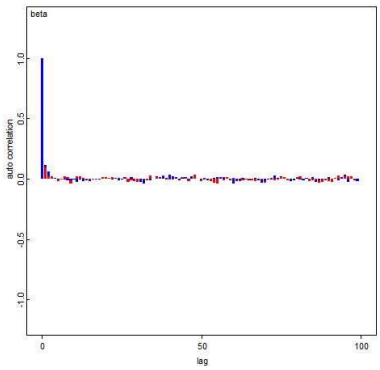

Figure A.49: Auto-Correlation of Parameter β

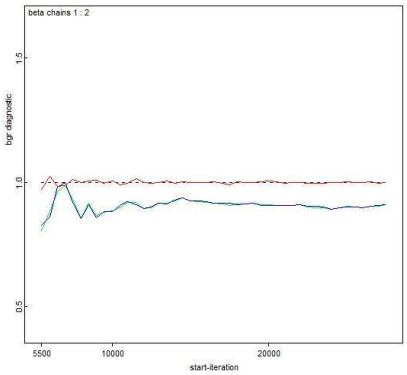

Figure A.50: BGR Diagnostic of Parameter β

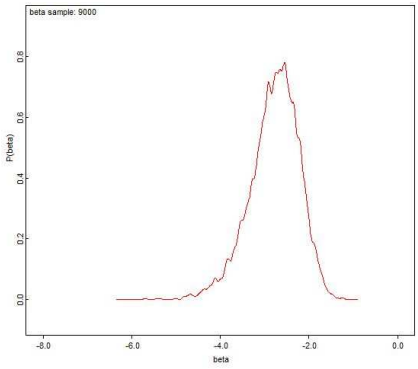

Figure A.51: Posterior Density of Parameter β

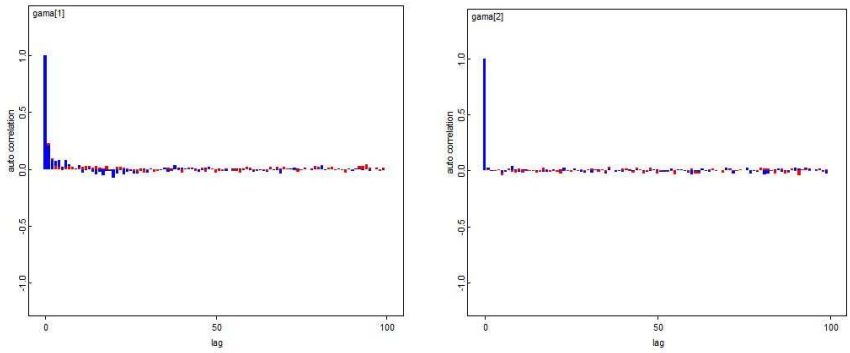

Figure A.52: Auto-Correlation of Parameter γ

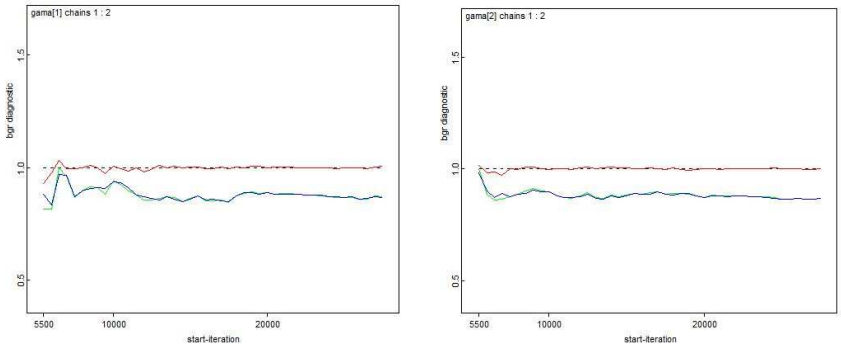

Figure A.53: BGR Diagnostic of Parameter γ

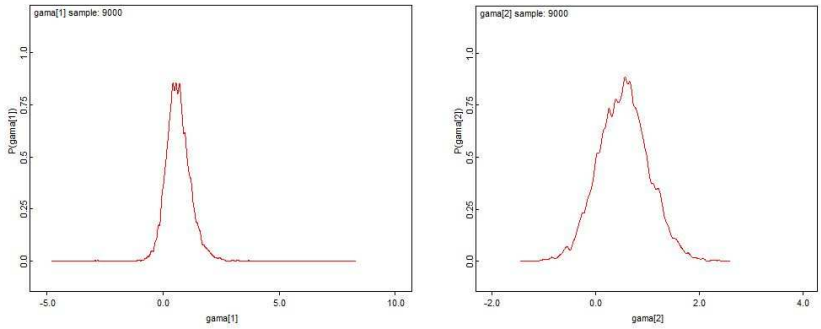

Figure A.54: Posterior Density of Parameter γ

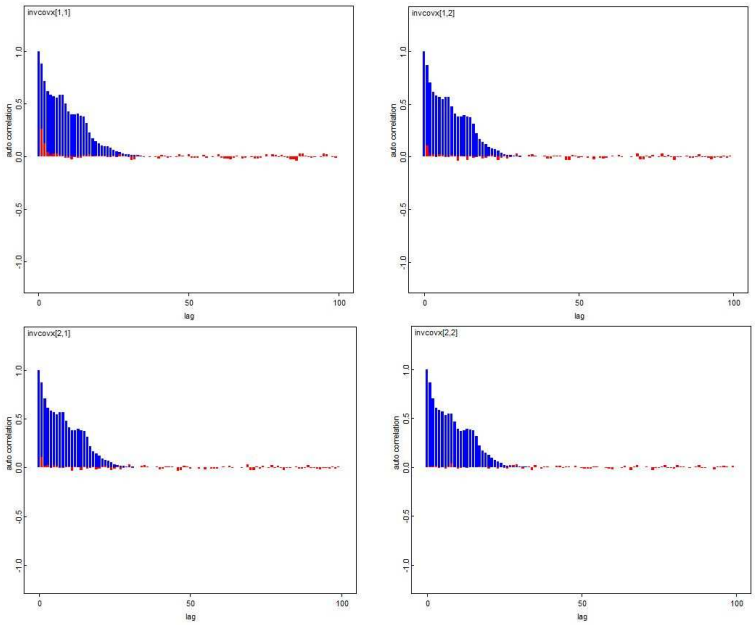

Figure A.55: Auto-Correlation of Parameter inverse Σ_X

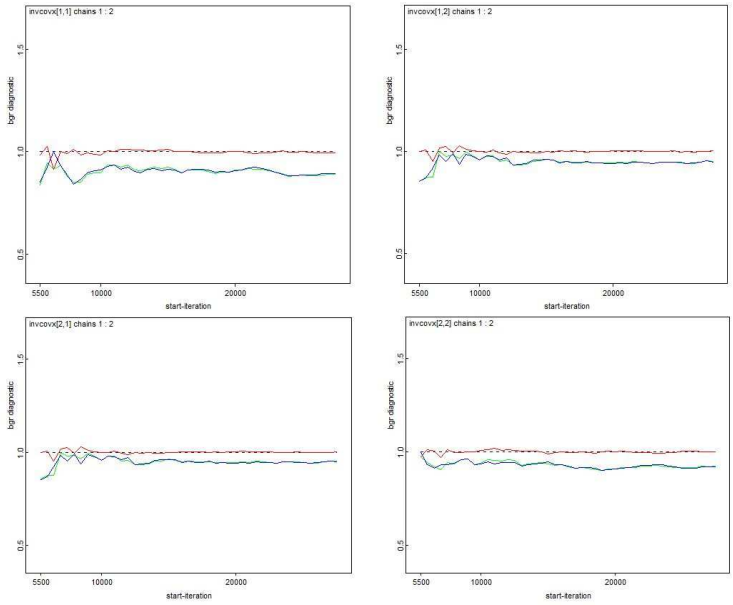

Figure A.56: BGR Diagnostic of Parameter inverse Σ_X

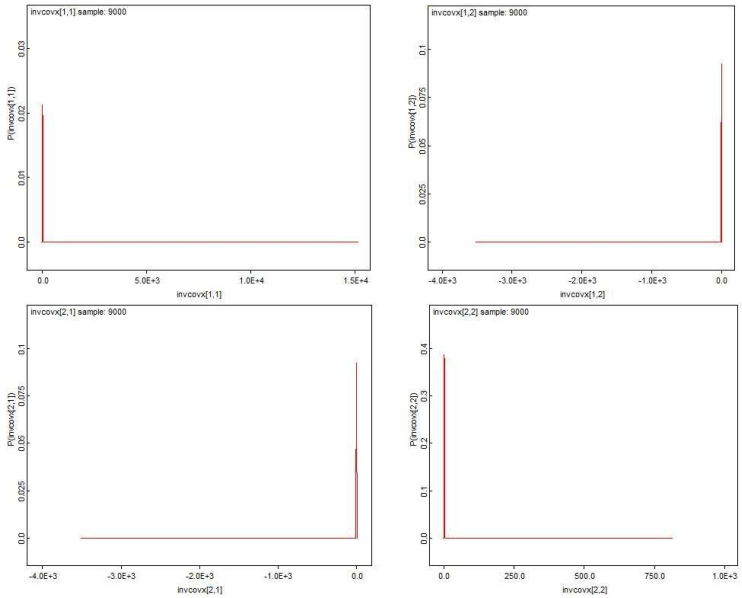

Figure A.57: Posterior Density of Parameter inverse $\boldsymbol{\Sigma_X}$

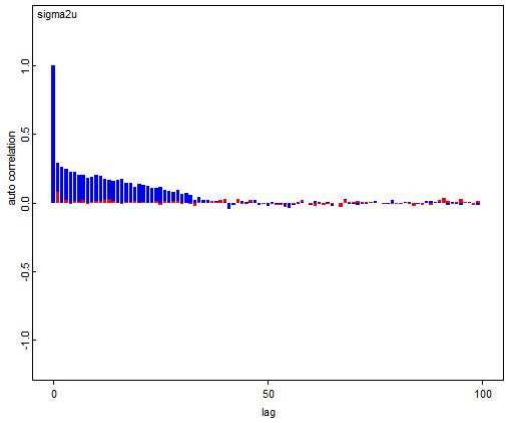

Figure A.58: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

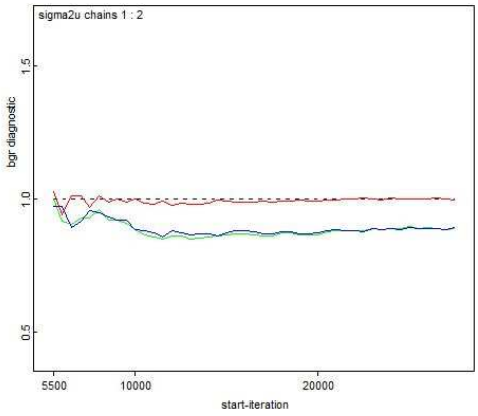

Figure A.59: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

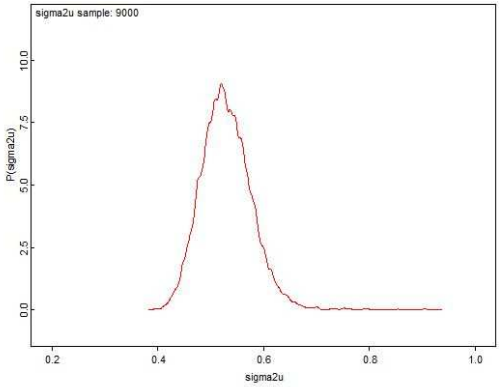

Figure A.60: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

APPENDIX B

CONVERGENCE DIAGNOSTICS OF JMM-KD MODEL

In this chapter, convergence diagnostics are presented for each interested parameters. Thinning, burn-in (warm up) iteration number and MC chain size in chapter 3 Simulation Study were determined in the light of these results. There are three images to observe the convergence in the posterior density. They are Auto-correlation between iterations, BGR (Brooks-Gelman-Rubin) (Gelman and Rubin (1992))diagnostic and Posterior density. These images are created with the help of OpenBUGS. The simulation procedure is same in simulation study part. There are two chains of initials for BGR Diagnostics, sample size for each chain is 50,000 which means 100,000 total samples. In each sample, sample size is $\boldsymbol{n = 100}$. Visual outputs of thinning=10 and burn-up=5,000 are given.

B.1 Bivariate Gamma Distribution

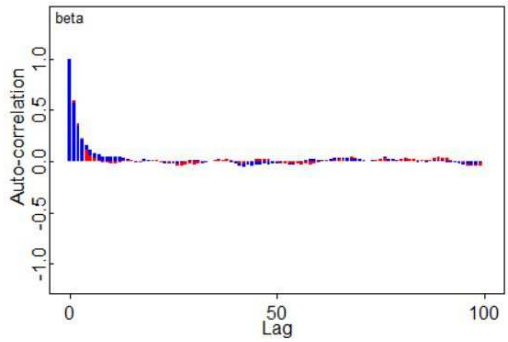

Figure B.1: Auto-Correlation of Parameter $\boldsymbol{\beta}$

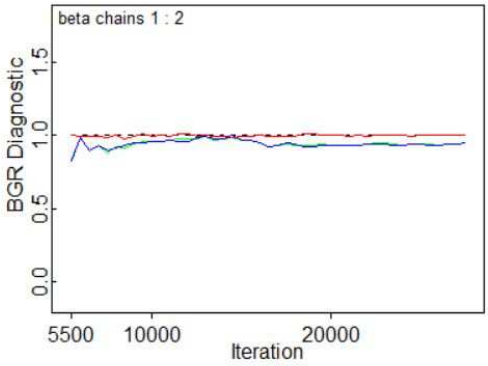

Figure B.2: BGR Diagnostic of Parameter β

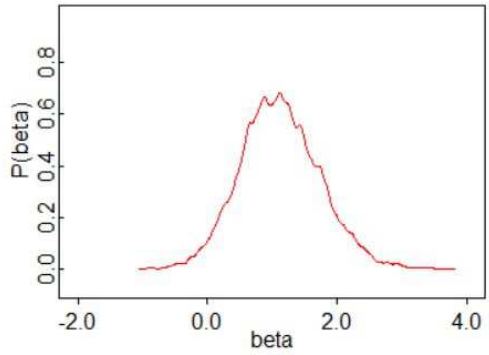

Figure B.3: Posterior Density of Parameter β

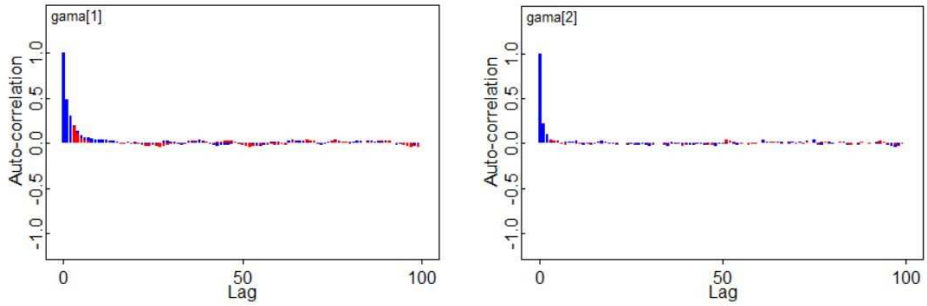

Figure B.4: Auto-Correlation of Parameter γ

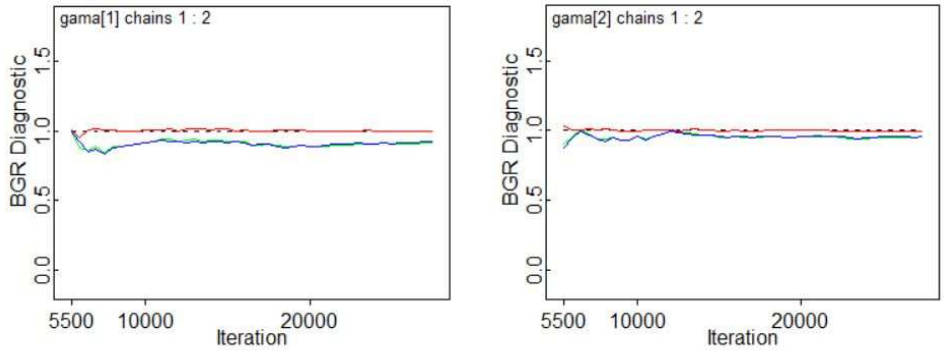

Figure B.5: BGR Diagnostic of Parameter γ

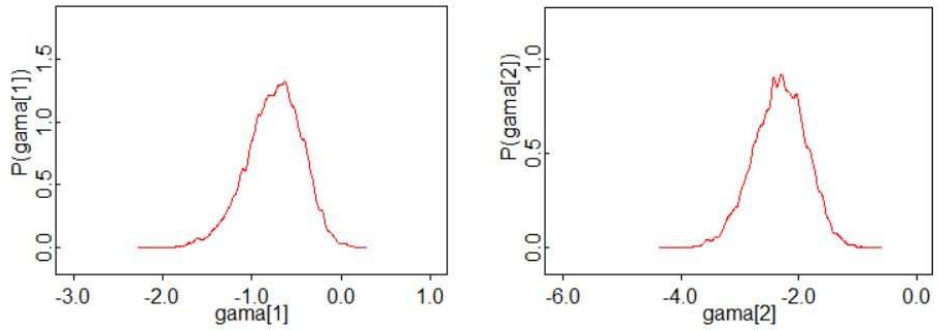

Figure B.6: Posterior Density of Parameter γ

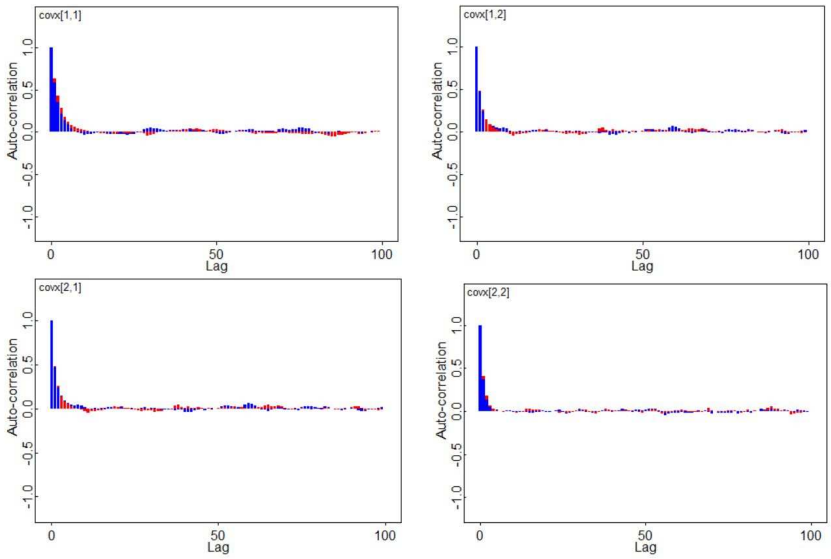

Figure B.7: Auto-Correlation of Parameter Σ_X

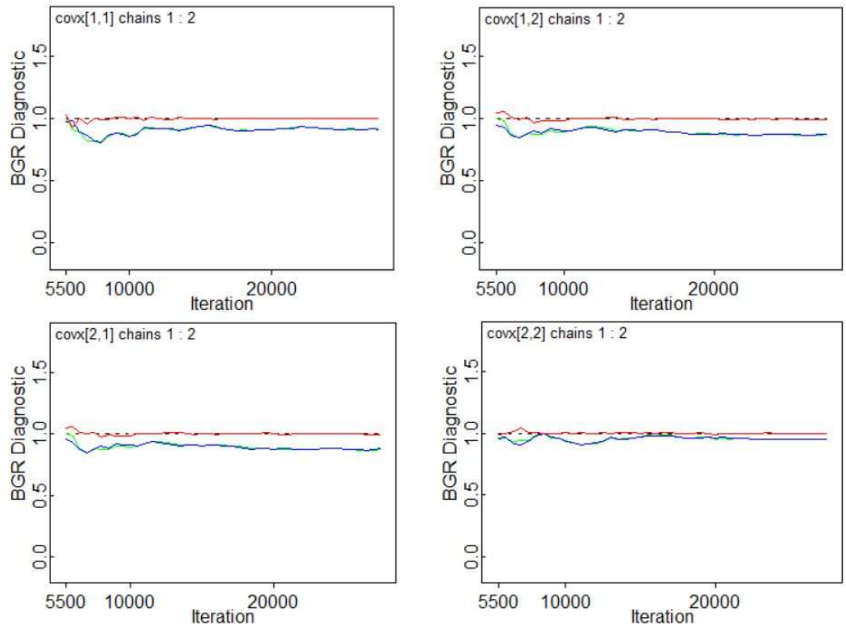

Figure B.8: BGR Diagnostic of Parameter Σ_X

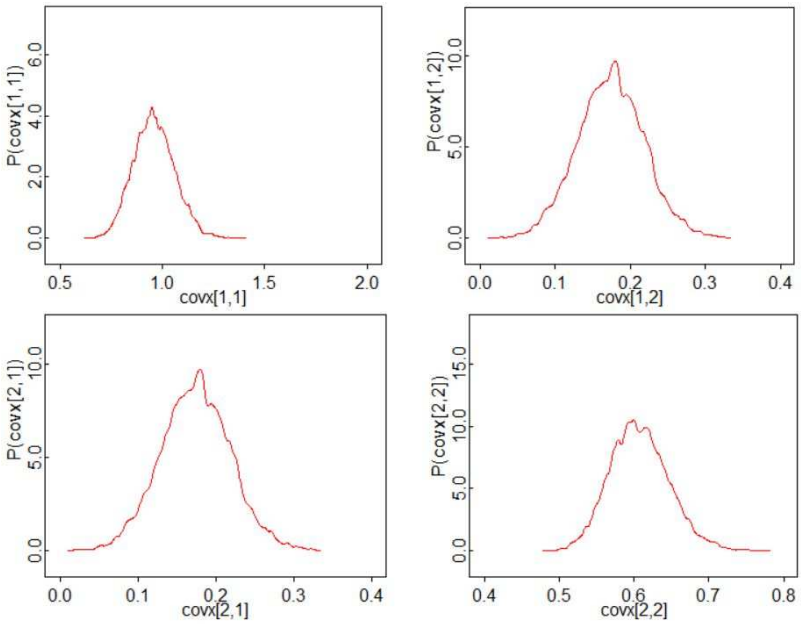

Figure B.9: Posterior Density of Parameter $\Sigma_{\boldsymbol{X}}$

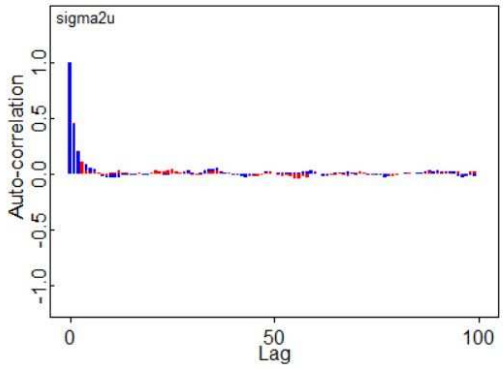

Figure B.10: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

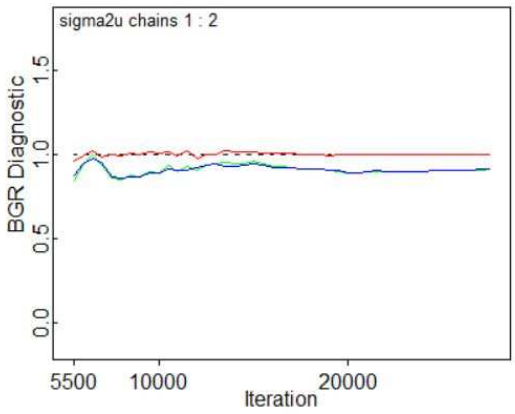

Figure B.11: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

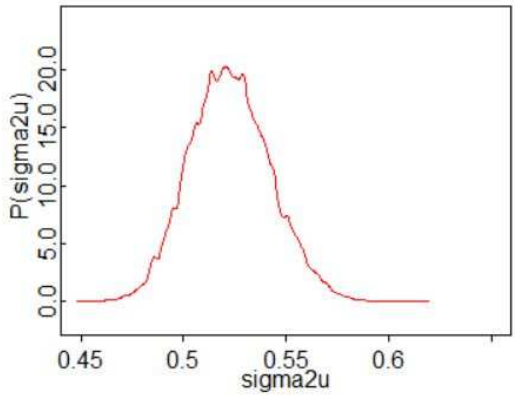

Figure B.12: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

B.2 Bivariate Gamma Mixture Distribution

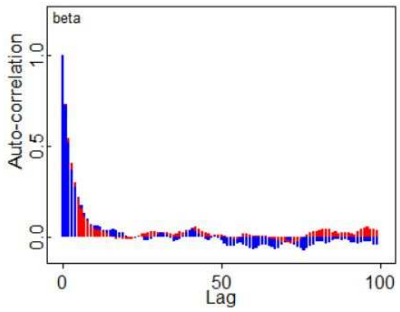

Figure B.13: Auto-Correlation of Parameter β

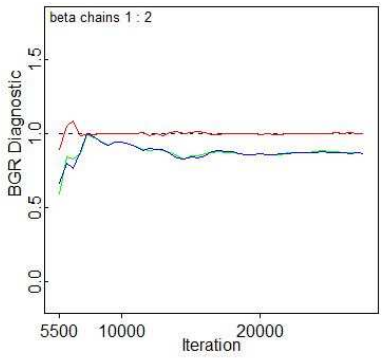

Figure B.14: BGR Diagnostic of Parameter β

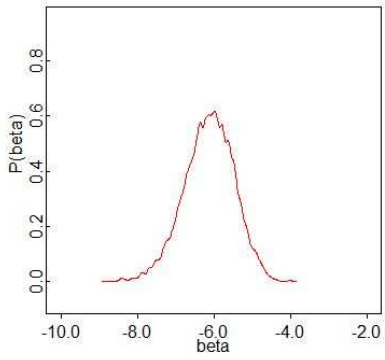

Figure B.15: Posterior Density of Parameter β

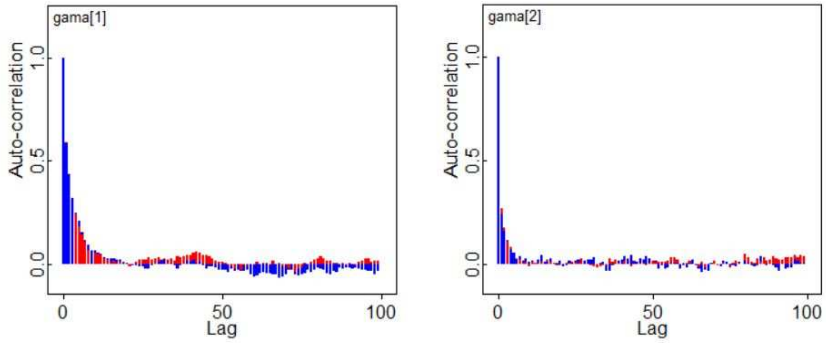

Figure B.16: Auto-Correlation of Parameter γ

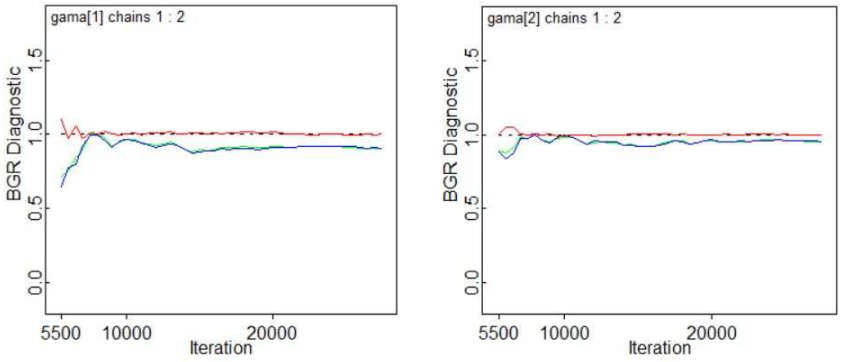

Figure B.17: BGR Diagnostic of Parameter γ

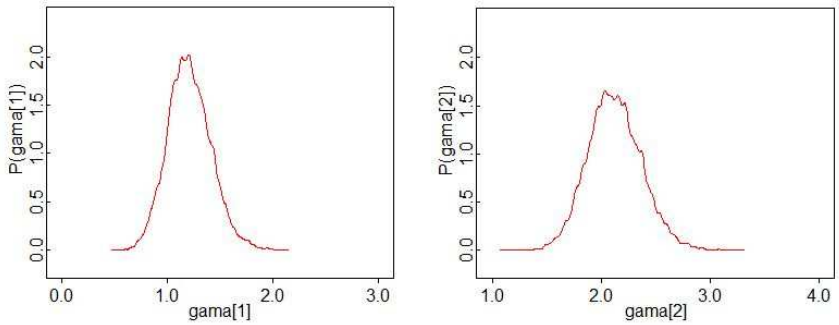

Figure B.18: Posterior Density of Parameter γ

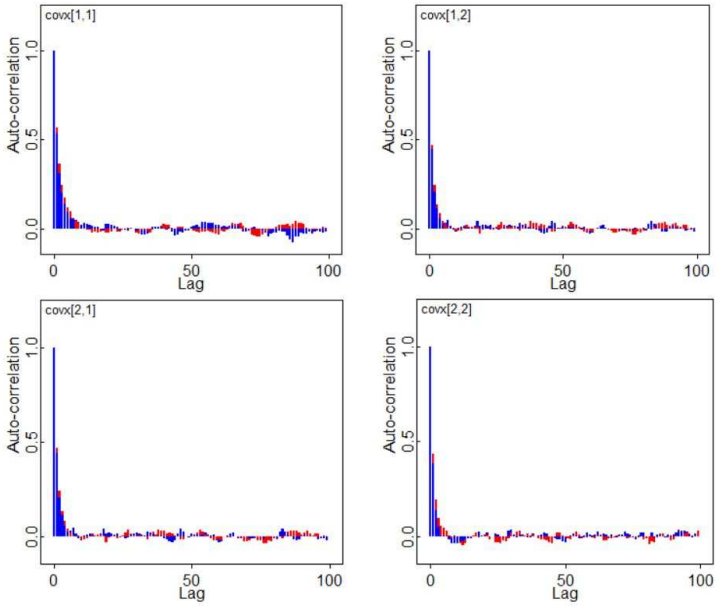

Figure B.19: Auto-Correlation of Parameter Σ_X

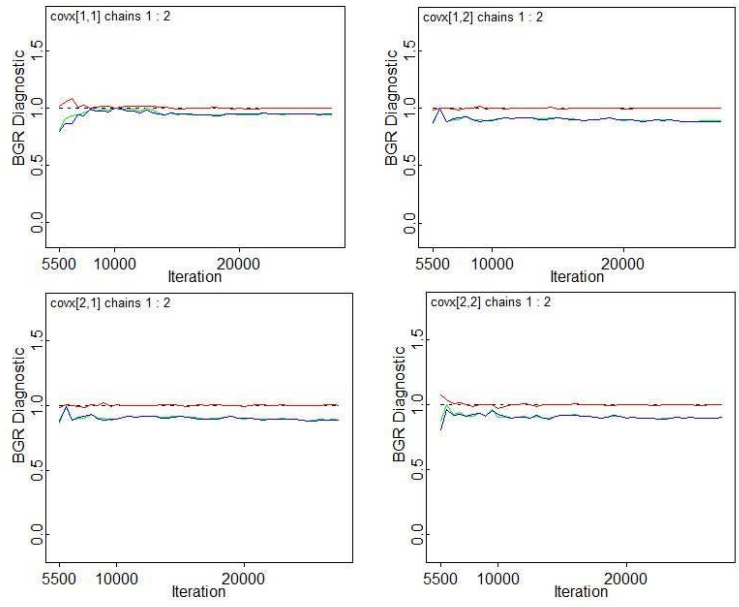

Figure B.20: BGR Diagnostic of Parameter Σ_X

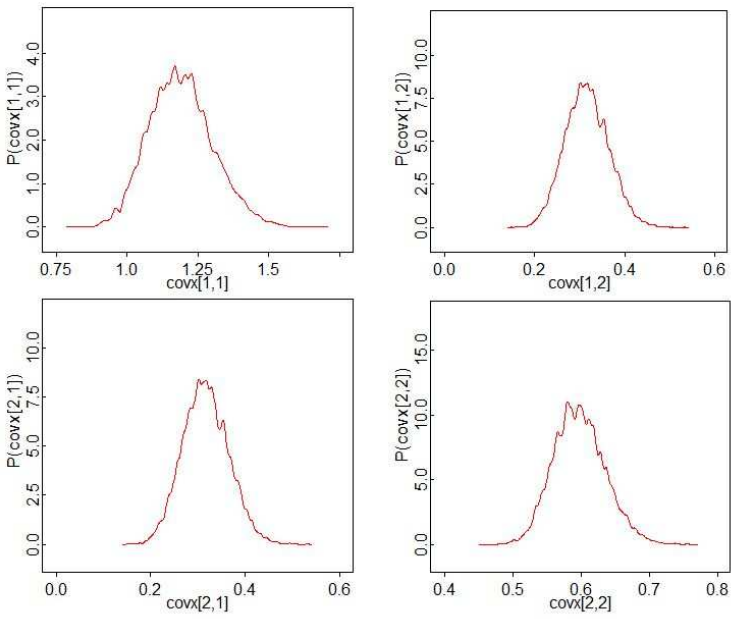

Figure B.21: Posterior Density of Parameter $\boldsymbol{\Sigma_X}$

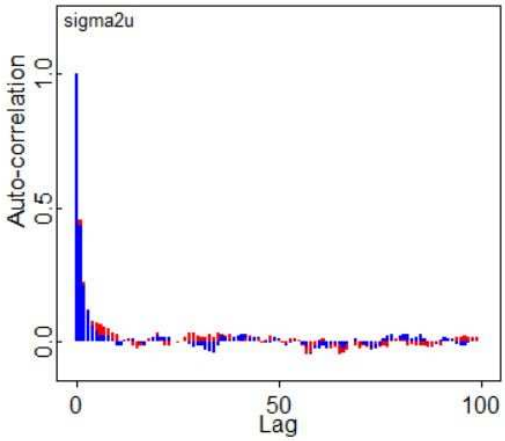

Figure B.22: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

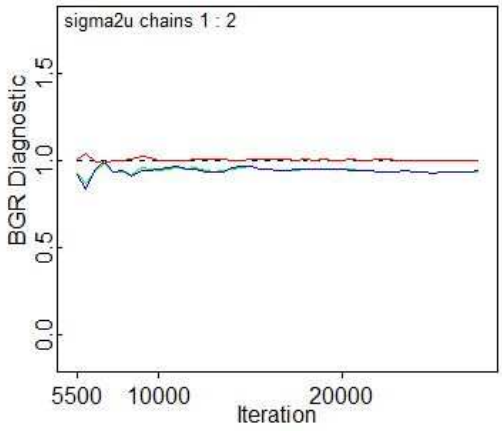

Figure B.23: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

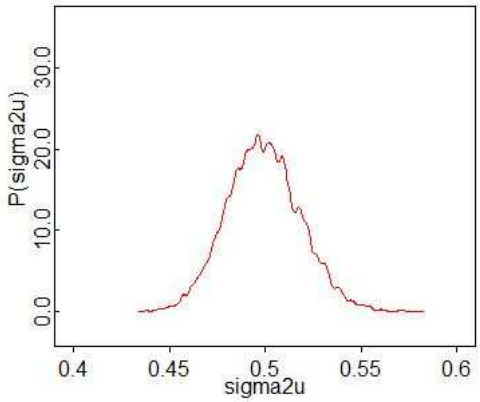

Figure B.24: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

B.3 Bivariate Normal Mixture Distribution

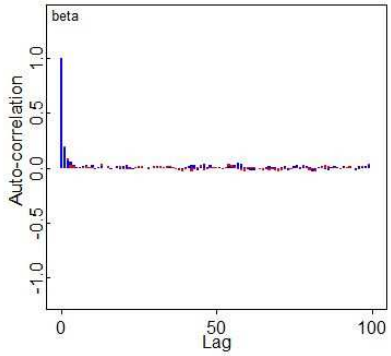

Figure B.25: Auto-Correlation of Parameter β

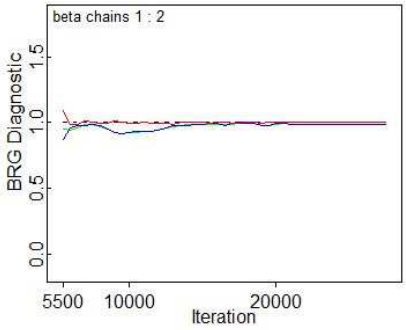

Figure B.26: BGR Diagnostic of Parameter β

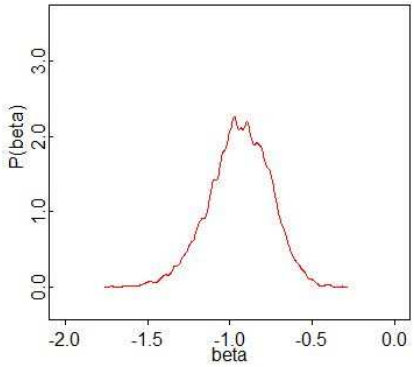

Figure B.27: Posterior Density of Parameter β

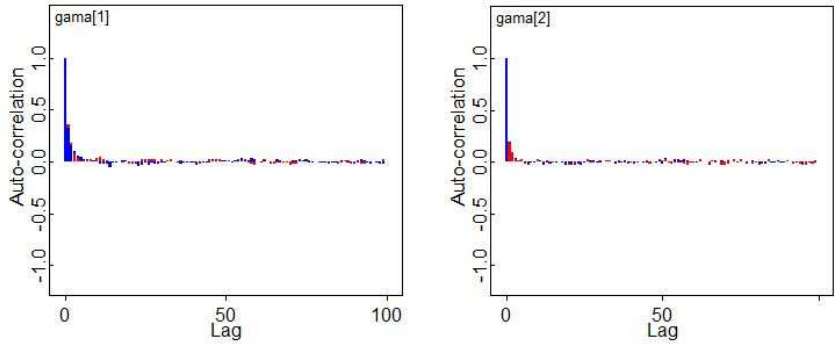

Figure B.28: Auto-Correlation of Parameter γ

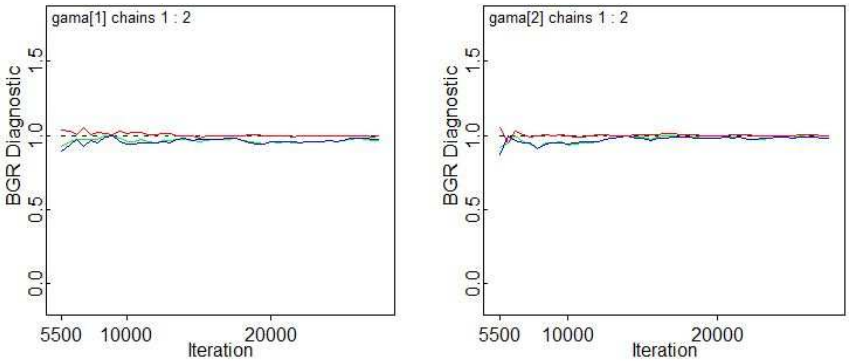

Figure B.29: BGR Diagnostic of Parameter γ

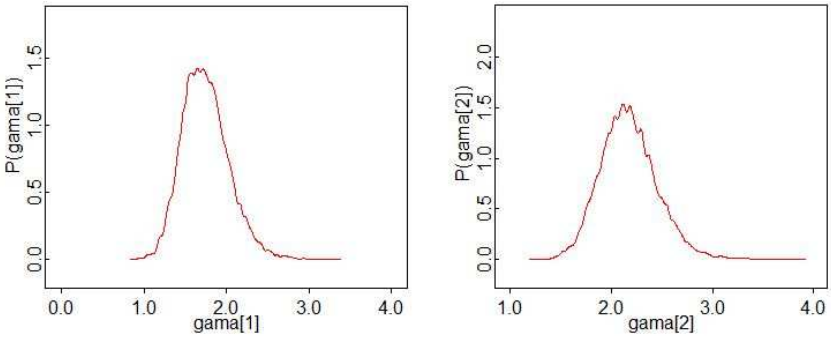

Figure B.30: Posterior Density of Parameter γ

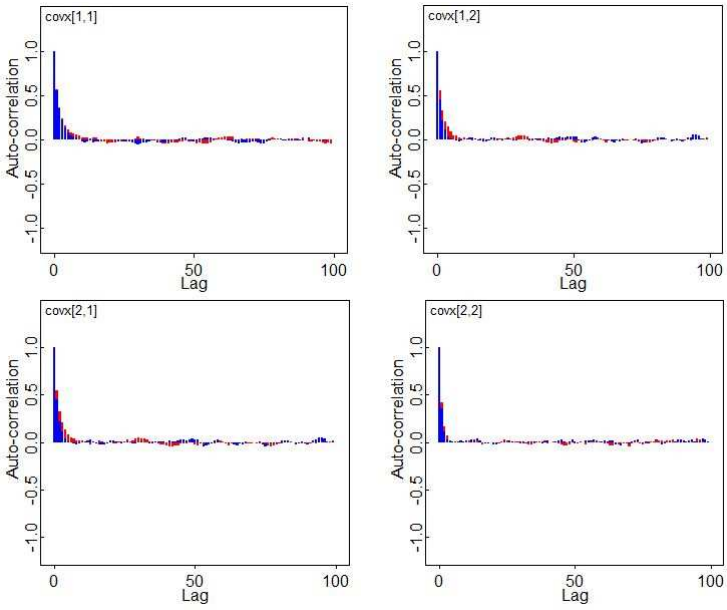

Figure B.31: Auto-Correlation of Parameter Σ_X

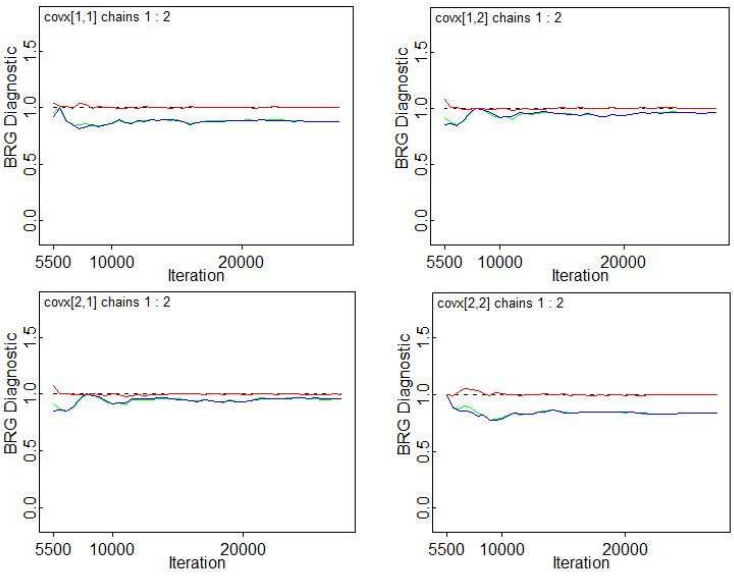

Figure B.32: BGR Diagnostic of Parameter Σ_X

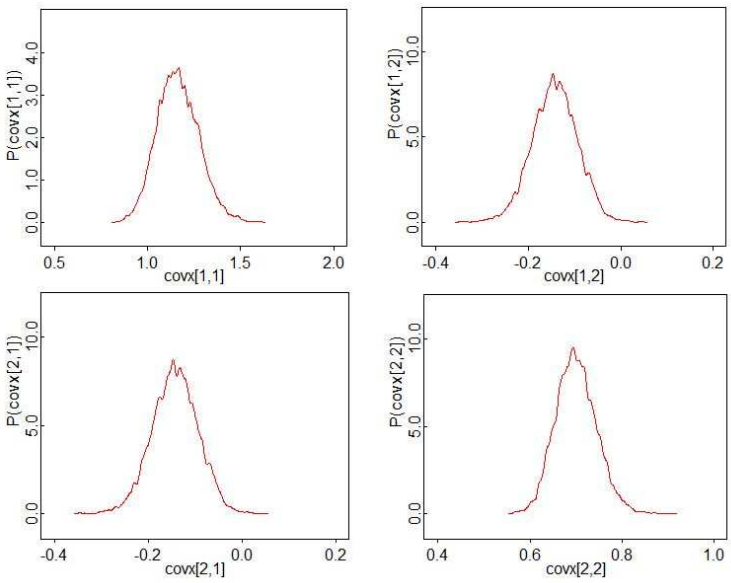

Figure B.33: Posterior Density of Parameter $\boldsymbol{\Sigma_X}$

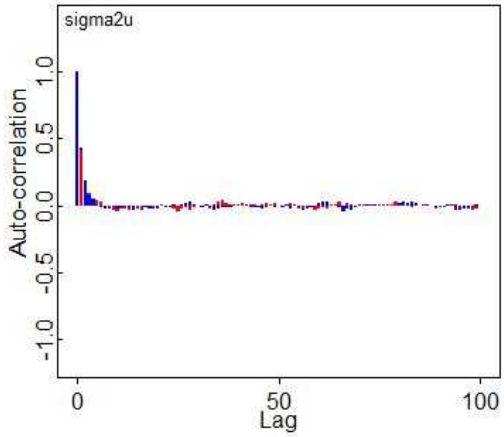

Figure B.34: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

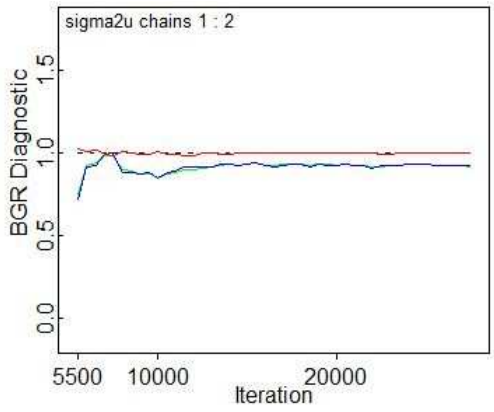

Figure B.35: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

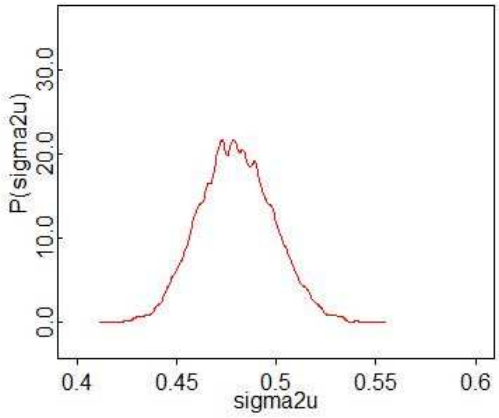

Figure B.36: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

B.4 Bivariate Skew-Normal Distribution

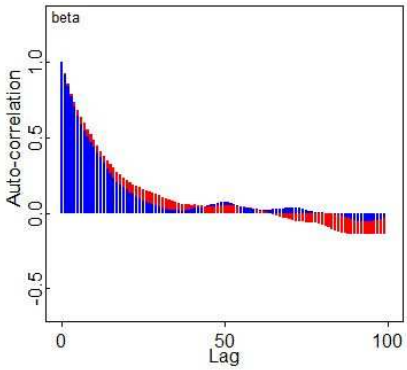

Figure B.37: Auto-Correlation of Parameter β

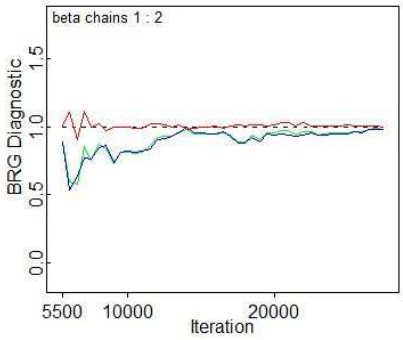

Figure B.38: BGR Diagnostic of Parameter β

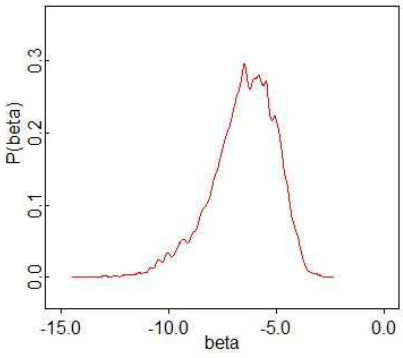

Figure B.39: Posterior Density of Parameter β

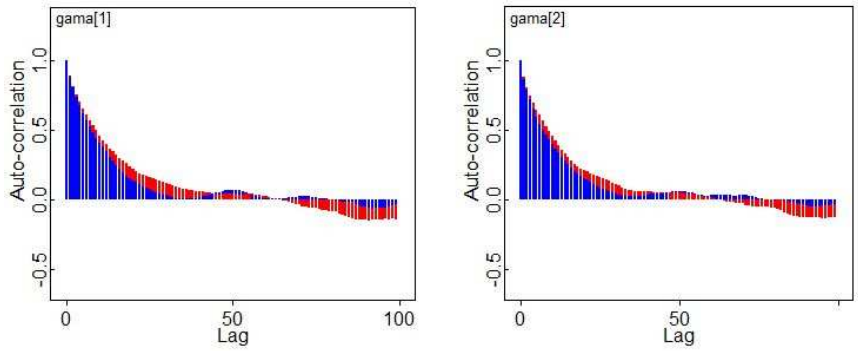

Figure B.40: Auto-Correlation of Parameter γ

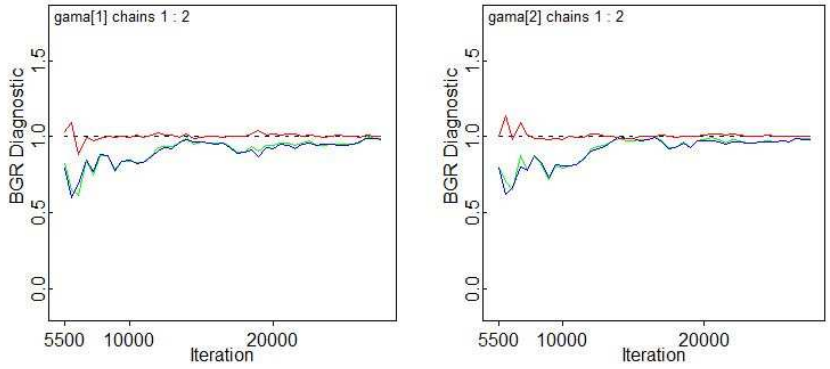

Figure B.41: BGR Diagnostic of Parameter γ

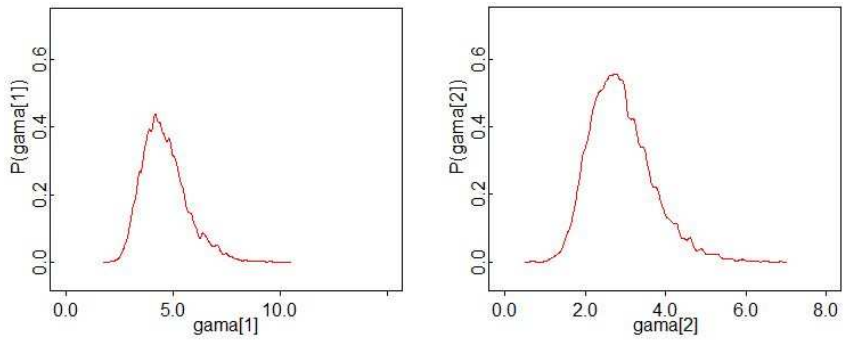

Figure B.42: Posterior Density of Parameter γ

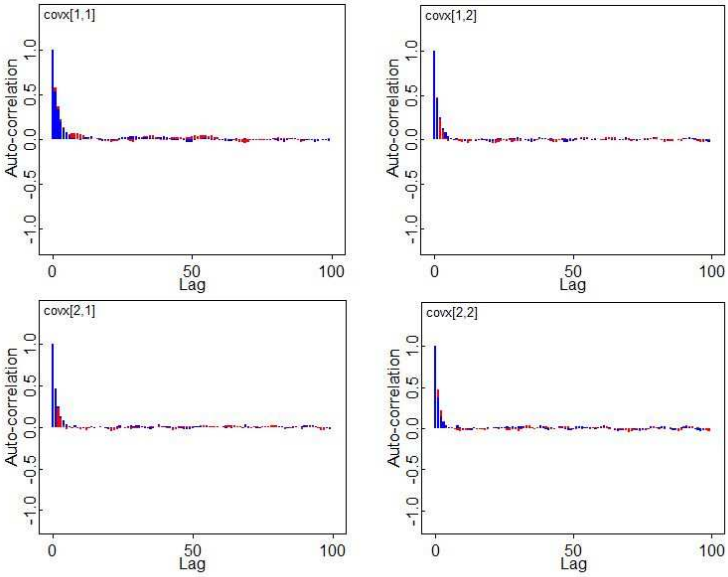

Figure B.43: Auto-Correlation of Parameter Σ_X

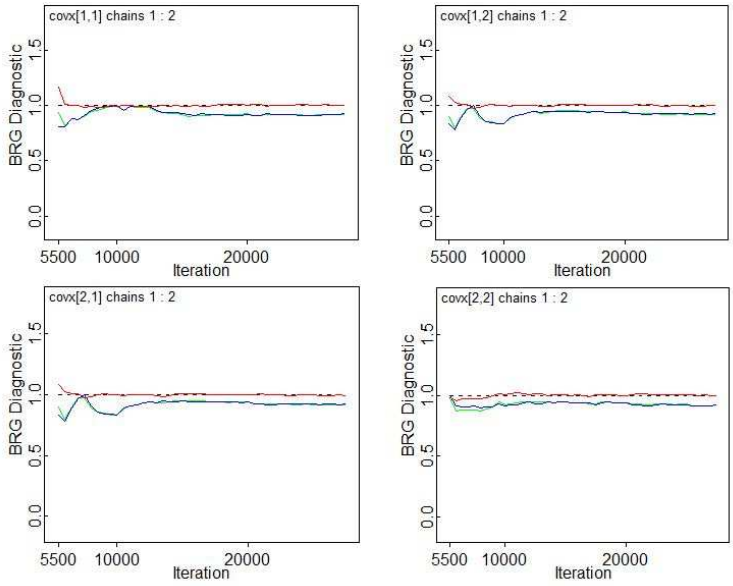

Figure B.44: BGR Diagnostic of Parameter Σ_X

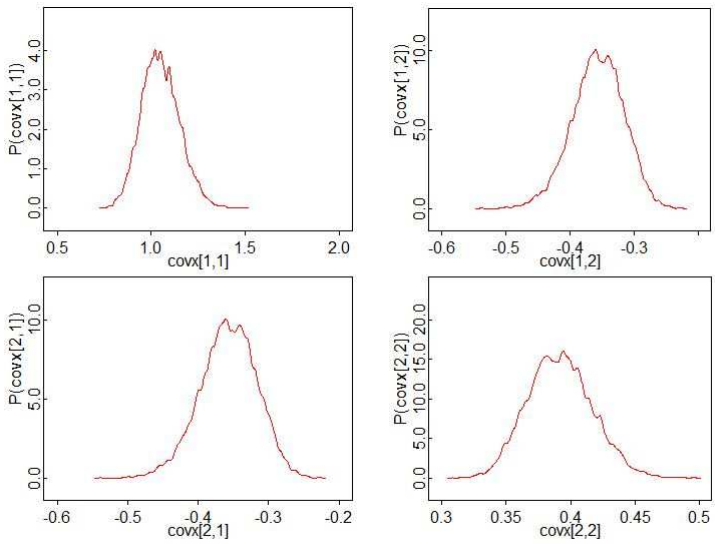

Figure B.45: Posterior Density of Parameter $\boldsymbol{\Sigma_X}$

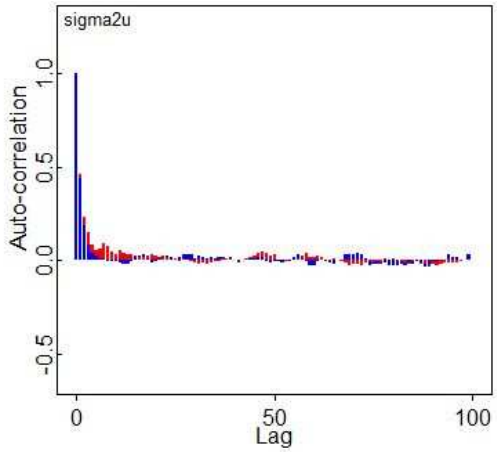

Figure B.46: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

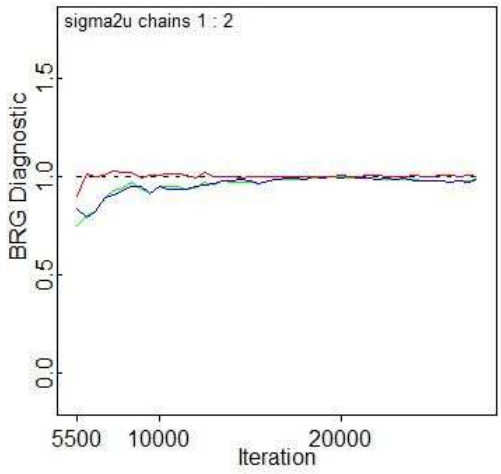

Figure B.47: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

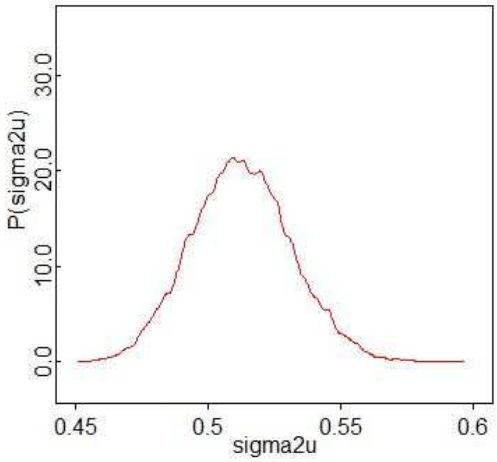

Figure B.48: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

B.5 Bivariate Skew-Normal Mixture Distribution

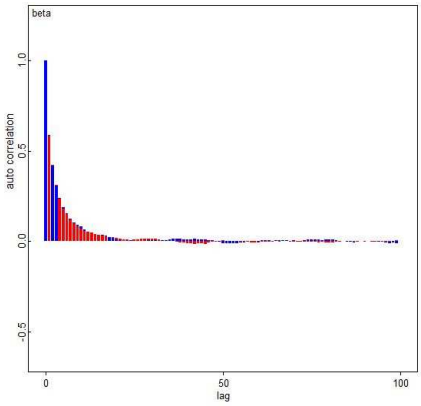

Figure B.49: Auto-Correlation of Parameter β

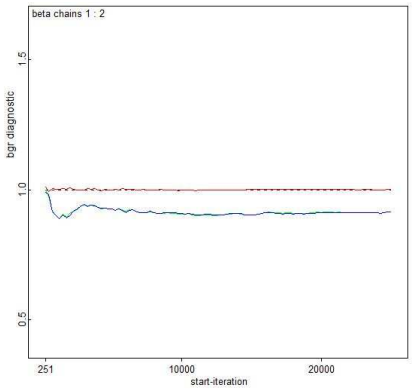

Figure B.50: BGR Diagnostic of Parameter β

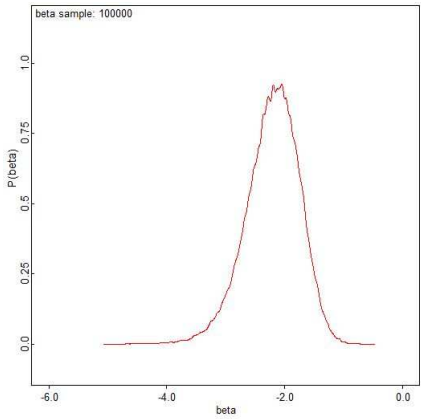

Figure B.51: Posterior Density of Parameter β

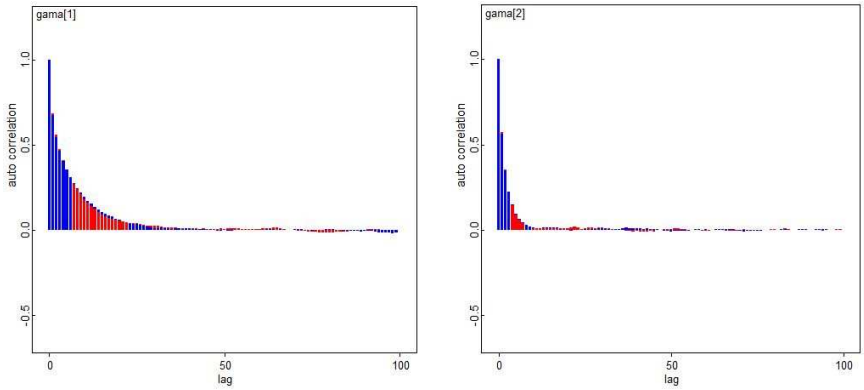

Figure B.52: Auto-Correlation of Parameter γ

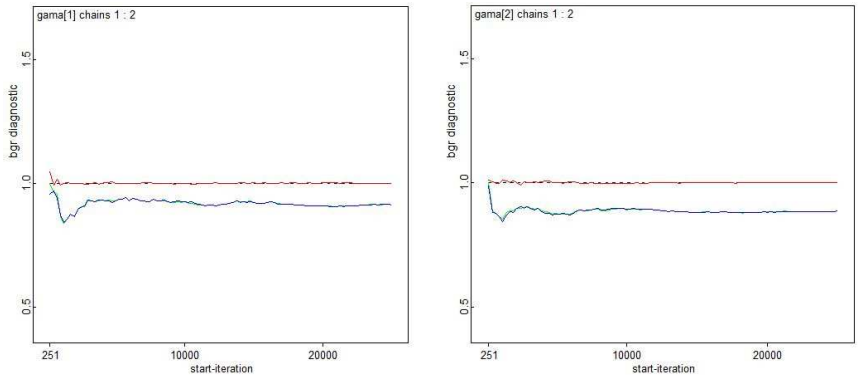

Figure B.53: BGR Diagnostic of Parameter γ

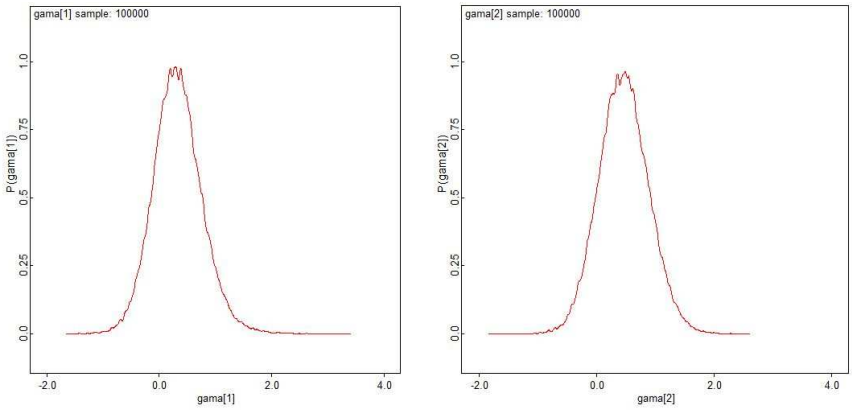

Figure B.54: Posterior Density of Parameter γ

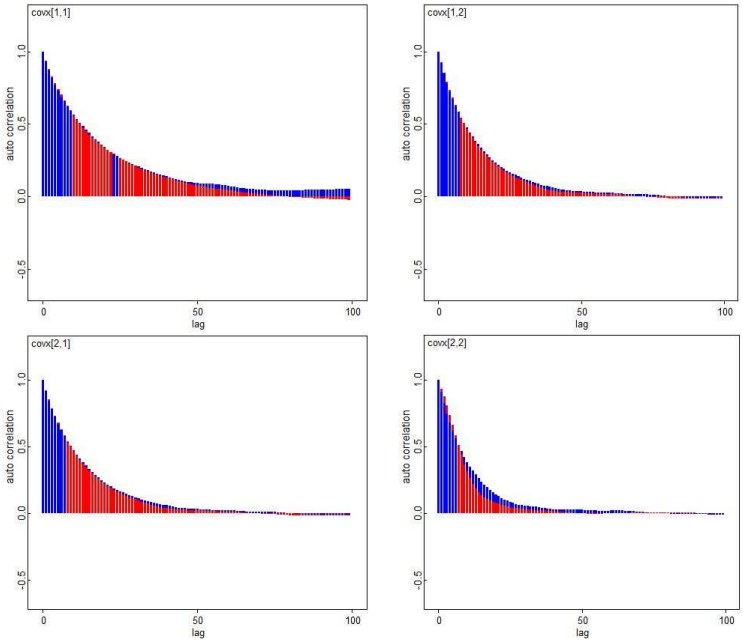

Figure B.55: Auto-Correlation of Parameter Σ_X

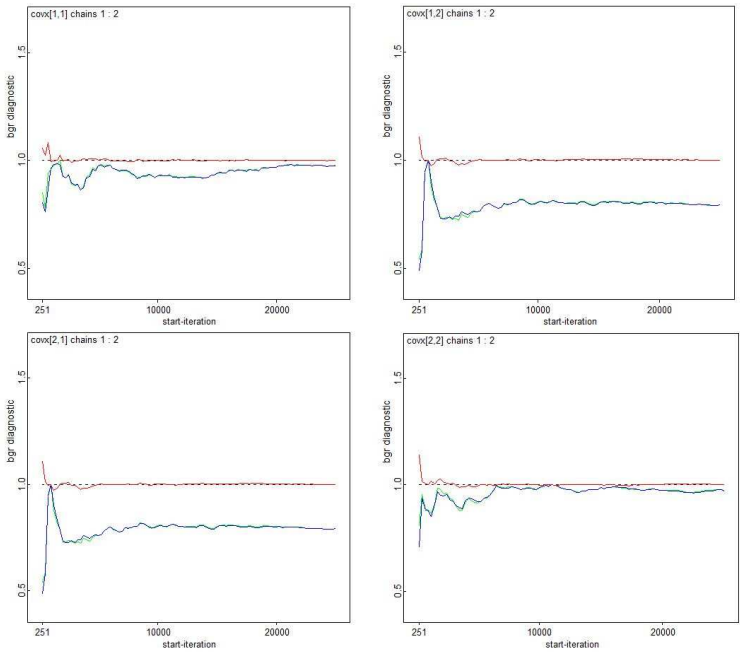

Figure B.56: BGR Diagnostic of Parameter $\boldsymbol{\Sigma_X}$

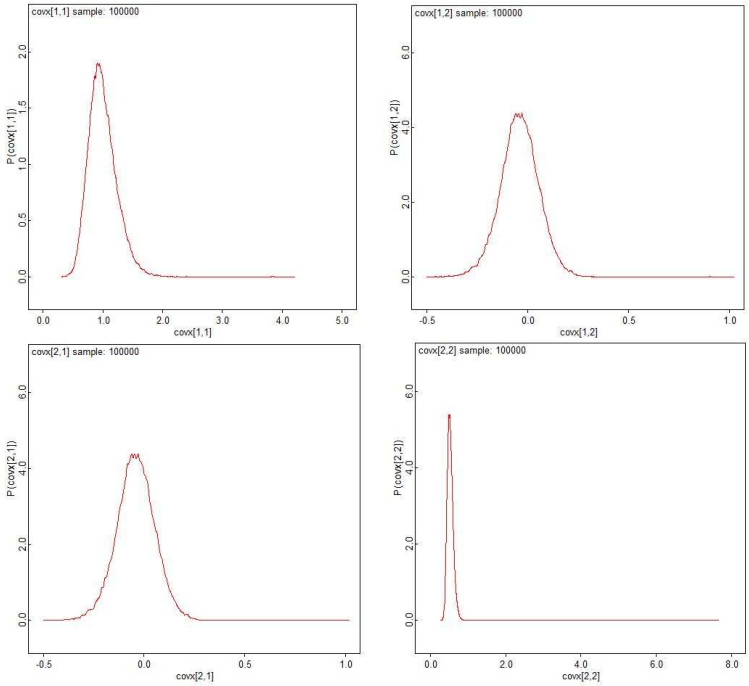

Figure B.57: Posterior Density of Parameter $\boldsymbol{\Sigma_X}$

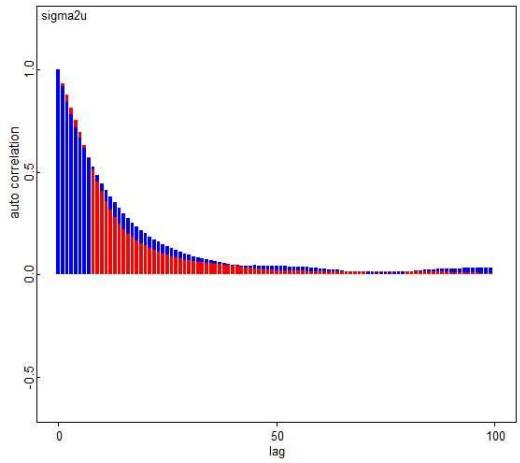

Figure B.58: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

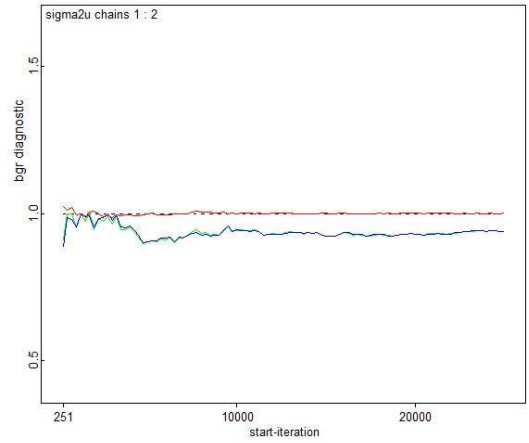

Figure B.59: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

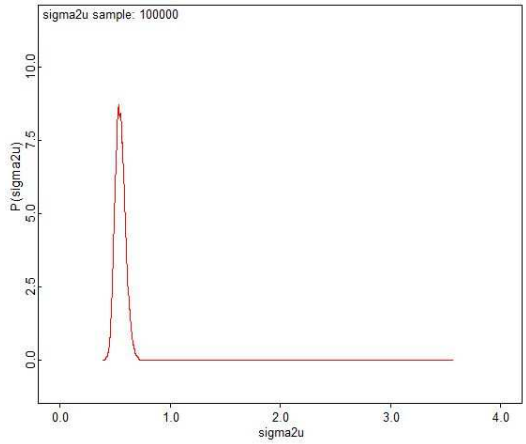

Figure B.60: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

APPENDIX C

MC STANDARD ERROR TABLES OF SIMULATIONS

Table C.1: MC Standard Error of JMM-St

		MC Standard Error					
n	Parameter	$X \sim N$	$X \sim N$ Mixture	$X \sim SN$ (Skewness=0.2)	$X \sim SN$ Mixture (Skewness=0.1)	$X \sim Ga$	$X \sim Ga$ Mixture
$P(Y=1 \mid x)=0.1$							
100	β	0.8953	1.2437	1.7833	0.7041	1.4645	1.5930
	γ_1	0.7061	0.9605	0.9144	0.6197	0.9767	0.9023
	γ_2	0.5821	0.7692	0.8641	0.5335	1.0037	1.1690
	σ_U^2	0.0428	0.0444	0.0590	0.0456	0.0489	0.0480
	$var(X_{i0})$	0.2579	0.2636	0.2883	0.2476	0.2655	0.3010
	$var(X_{i1})$	0.0838	0.1040	0.0639	0.0883	0.0936	0.0953
	$cov(X_{i0}, X_{i1})$	0.1054	0.1156	0.0875	0.0965	0.1100	0.1242
500	β	0.2650	0.8004	0.7395	0.2358	0.5745	0.6155
	γ_1	0.1901	0.4777	0.3567	0.2037	0.2772	0.2865
	γ_2	0.2138	0.3908	0.3519	0.2070	0.3813	0.3633
	σ_U^2	0.0203	0.0181	0.0179	0.0195	0.0205	0.0211
	$Var(X_{i0})$	0.0981	0.1010	0.0923	0.1089	0.1257	0.1253
	$Var(X_{i1})$	0.0394	0.0424	0.0298	0.0375	0.0540	0.0429
	$Cov(X_{i0}, X_{i1})$	0.0454	0.0512	0.0420	0.0487	0.0447	0.0464
$P(Y=1 \mid x)=0.5$							
100	β	0.9471	0.5383	1.1952	0.7055	0.6788	1.6622
	γ_1	0.9707	0.9545	0.8122	0.8153	0.3716	0.5738
	γ_2	0.6710	0.6626	0.7385	0.6303	0.3017	0.6000
	σ_U^2	0.0428	0.0430	0.0452	0.0417	0.0532	0.0514
	$var(X_{i0})$	0.2415	0.3703	0.2149	0.1962	0.2750	0.3078
	$var(X_{i1})$	0.0964	0.1346	0.0625	0.0833	0.1144	0.0950
	$cov(X_{i0}, X_{i1})$	0.1148	0.2366	0.0865	0.1132	0.0997	0.1064
500	β	0.6449	0.2089	0.9387	0.4847	0.2759	0.6937
	γ_1	0.7274	0.3765	0.5999	0.7576	0.1686	0.2212
	γ_2	0.4344	0.2719	0.4811	0.4821	0.1727	0.2267
	σ_U^2	0.0196	0.0195	0.0182	0.0179	0.0257	0.0186
	$Var(X_{i0})$	0.1115	0.0961	0.0874	0.1108	0.1576	0.1496
	$Var(X_{i1})$	0.0406	0.0431	0.0257	0.0394	0.0499	0.0524
	$Cov(X_{i0}, X_{i1})$	0.0443	0.0414	0.0353	0.0495	0.0472	0.0521

Table C.2: MC Standard Error of JMM-KD

		MC Standard Error					
n	Parameter	$X \sim N$	$X \sim N$ Mixture	$X \sim SN$ (Skewness=0.2)	$X \sim SN$ Mixture (Skewness=0.1)	$X \sim Ga$	$X \sim Ga$ Mixture
$P(Y=1 \mid x)=0.1$							
100	β	0.8886	1.4495	1.6982	0.6583	1.7781	1.4442
	γ_1	0.6980	1.0409	0.8364	0.5745	1.0846	0.7027
	γ_2	0.5840	0.7404	0.8755	0.4404	1.0180	0.9818
	σ_U^2	0.0416	0.0430	0.0407	0.0416	0.0375	0.0459
	$var(X_{i0})$	0.2534	0.2245	0.2380	0.2442	0.2648	0.2802
	$var(X_{i1})$	0.0863	0.1004	0.0596	0.0883	0.1201	0.0865
	$cov(X_{i0}, X_{i1})$	0.1065	0.0994	0.0953	0.1136	0.0928	0.1021
500	β	0.2965	0.8568	0.8147	0.2502	0.5276	0.6444
	γ_1	0.2344	0.5263	0.3166	0.1989	0.2796	0.2580
	γ_2	0.1967	0.4229	0.4379	0.2125	0.3672	0.4428
	σ_U^2	0.0181	0.0185	0.0186	0.0204	0.0168	0.0216
	$Var(X_{i0})$	0.1069	0.0946	0.0953	0.1084	0.1151	0.1401
	$Var(X_{i1})$	0.0419	0.0455	0.0270	0.0363	0.0486	0.0471
	$Cov(X_{i0}, X_{i1})$	0.0529	0.0480	0.0434	0.0458	0.0429	0.0611
$P(Y=1 \mid x)=0.5$							
100	β	0.7306	0.5320	1.3600	0.7300	0.6018	1.7852
	γ_1	0.8958	0.9380	0.7925	0.9936	0.3054	0.5517
	γ_2	0.6900	0.6632	0.7689	0.7540	0.3199	0.6532
	σ_U^2	0.0407	0.0426	0.0411	0.0443	0.0464	0.0440
	$Var(X_{i0})$	0.2604	0.2185	0.1900	0.2744	0.2888	0.3364
	$Var(X_{i1})$	0.0891	0.1002	0.0577	0.0775	0.1295	0.1026
	$Cov(X_{i0}, X_{i1})$	0.0974	0.0995	0.0774	0.1038	0.1202	0.1199
500	β	0.5856	0.2242	1.1749	0.4791	0.2501	0.6601
	γ_1	0.7074	0.3847	0.7143	0.7203	0.1145	0.1861
	γ_2	0.4437	0.3179	0.5831	0.4662	0.1367	0.2505
	σ_U^2	0.0202	0.0175	0.0206	0.0186	0.0176	0.0177
	$Var(X_{i0})$	0.1000	0.1016	0.0989	0.1085	0.1199	0.1336
	$Var(X_{i1})$	0.0436	0.0424	0.0275	0.0388	0.0572	0.0481
	$Cov(X_{i0}, X_{i1})$	0.0432	0.0458	0.0381	0.0454	0.0494	0.0563

APPENDIX D

CONVERGENCE DIAGNOSTICS OF JMM-ST WITH DPM

In this chapter, convergence diagnostics of JMM-St with DPM are presented for each interested parameters. Thinning, burn-in (warm up) iteration number and MC chain size in chapter 3 Simulation Study were determined in the light of these results. There are three images to observe the convergence in the posterior density. They are Auto-correlation between iterations, BGR (Brooks-Gelman-Rubin) diagnostic and Posterior density. These images are created with the help of OpenBUGS. The simulation procedure is same in simulation study part. There are two chains of initials for BGR Diagnostics, sample size for each chain is 50,000 which means 100,000 total samples. In each sample, sample size is $\boldsymbol{n = 100}$.

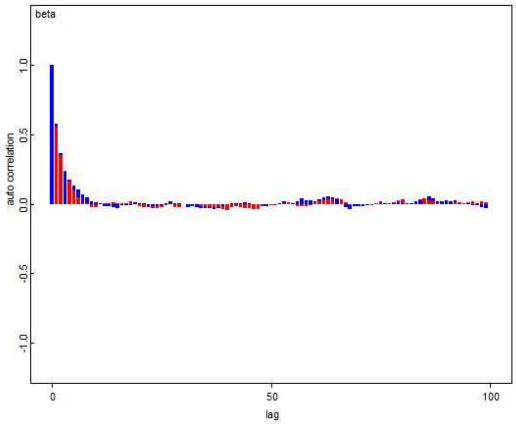

Figure D.1: Auto-Correlation of Parameter $\boldsymbol{\beta}$

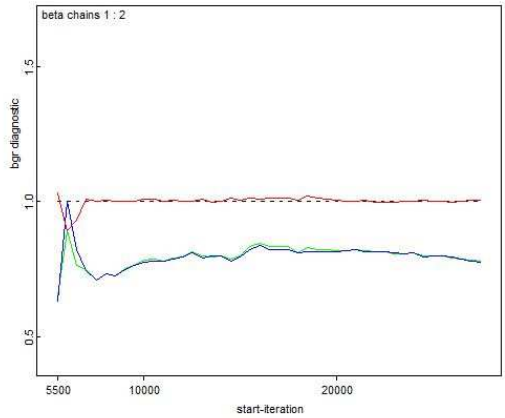

Figure D.2: BGR Diagnostic of Parameter β

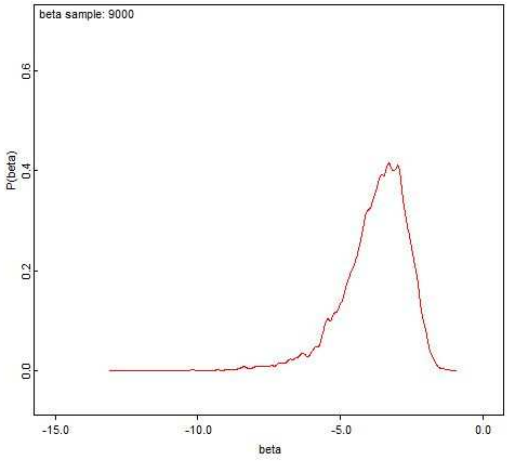

Figure D.3: Posterior Density of Parameter β

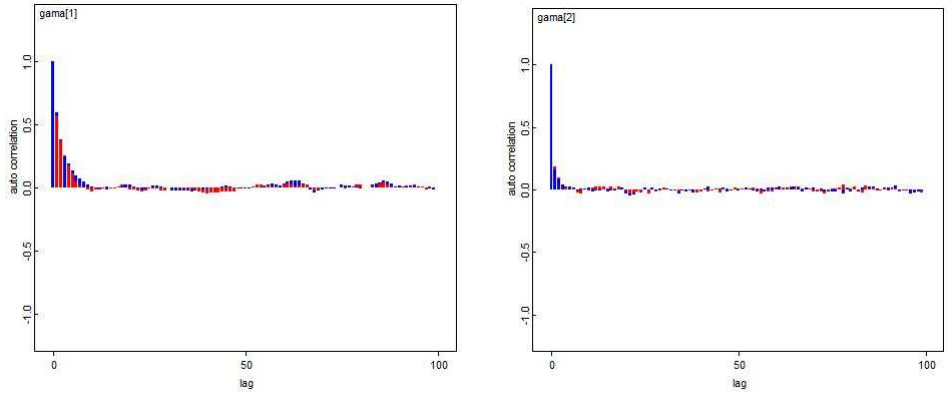

Figure D.4: Auto-Correlation of Parameter γ

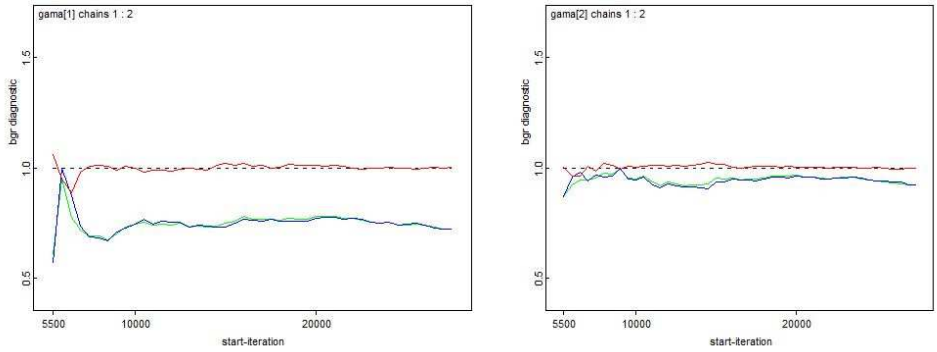

Figure D.5: BGR Diagnostic of Parameter γ

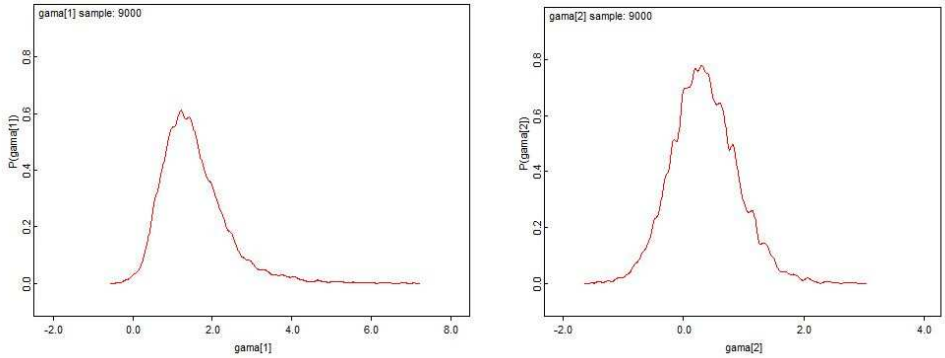

Figure D.6: Posterior Density of Parameter γ

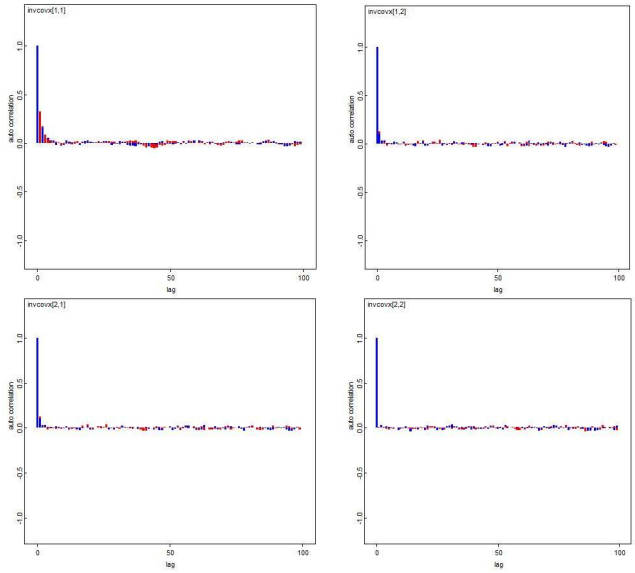

Figure D.7: Auto-Correlation of Parameter inverse $\boldsymbol{\Sigma_X}$

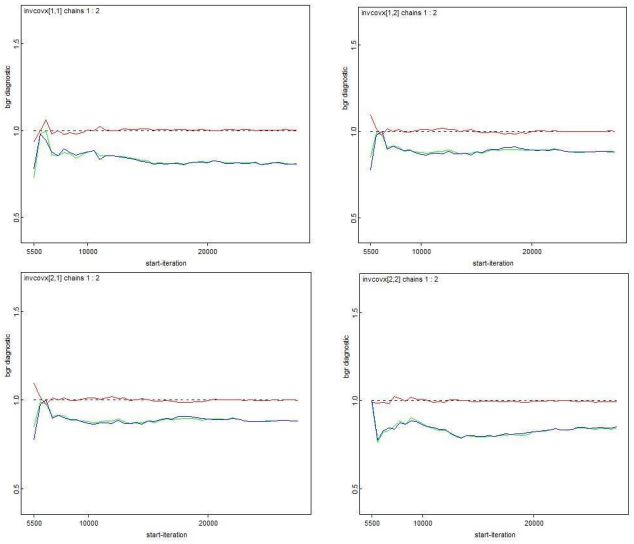

Figure D.8: BGR Diagnostic of Parameter inverse $\boldsymbol{\Sigma_X}$

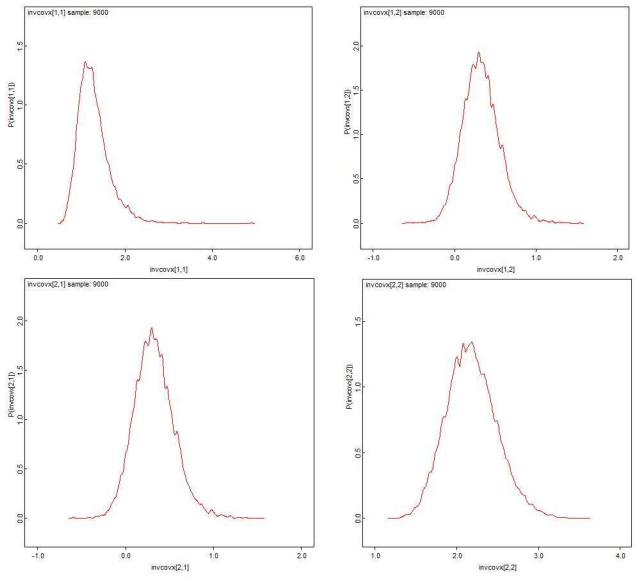

Figure D.9: Posterior Density of Parameter inverse $\boldsymbol{\Sigma_X}$

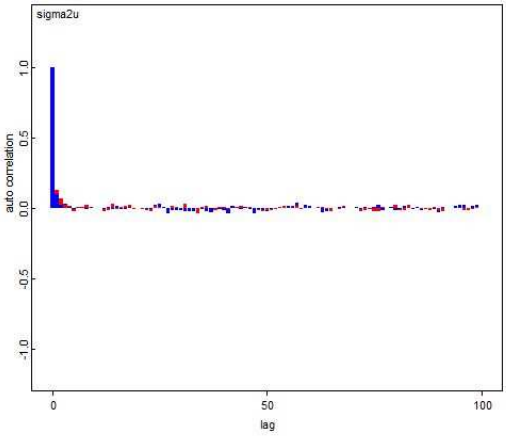

Figure D.10: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

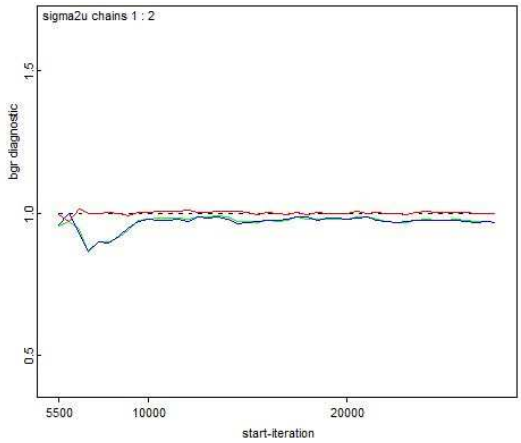

Figure D.11: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

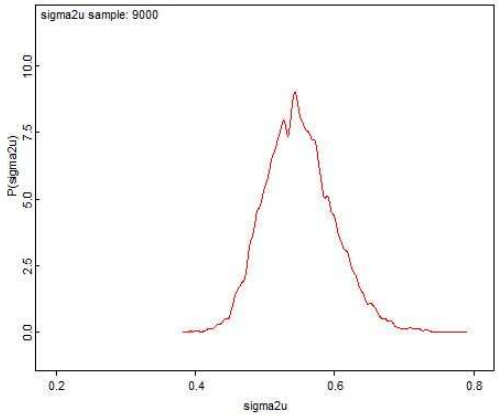

Figure D.12: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

APPENDIX E

CONVERGENCE DIAGNOSTICS OF JMM-KD WITH DPM

In Appendix B, convergence diagnostics of JMM-KD with DPM are presented for each interested parameters. Thinning, burn-in (warm up) iteration number and MC chain size in chapter 3 Simulation Study were determined in the light of these results. There are three images to observe the convergence in the posterior density. They are Auto-correlation between iterations, BGR (Brooks-Gelman-Rubin) diagnostic and Posterior density. These images are created with the help of OpenBUGS. The simulation procedure is same in simulation study part. There are two chains of initials for BGR Diagnostics, sample size for each chain is 50,000 which means 100,000 total samples. In each sample, sample size is $\boldsymbol{n = 100}$.

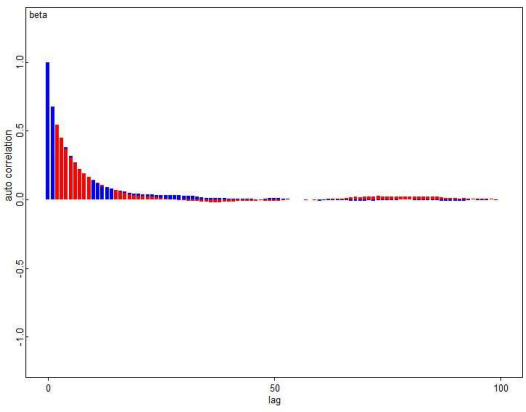

Figure E.1: Auto-Correlation of Parameter $\boldsymbol{\beta}$

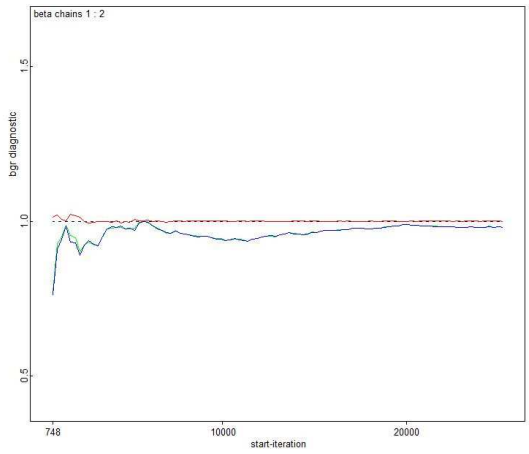

Figure E.2: BGR Diagnostic of Parameter β

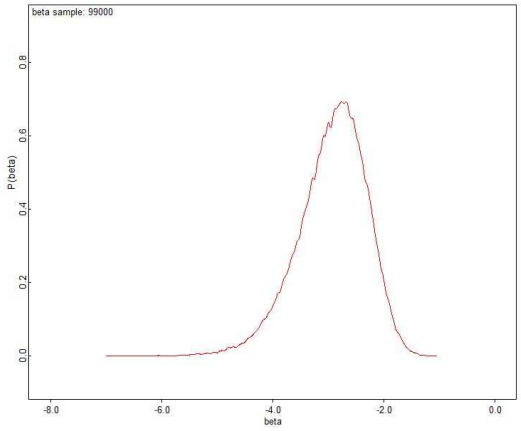

Figure E.3: Posterior Density of Parameter β

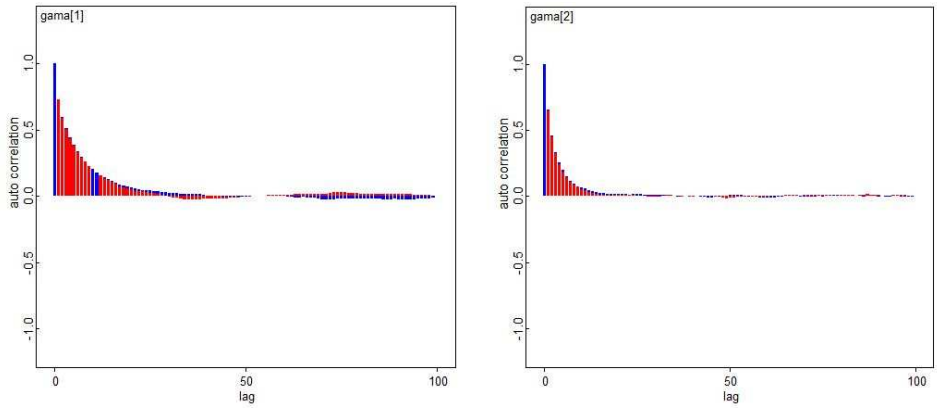

Figure E.4: Auto-Correlation of Parameter γ

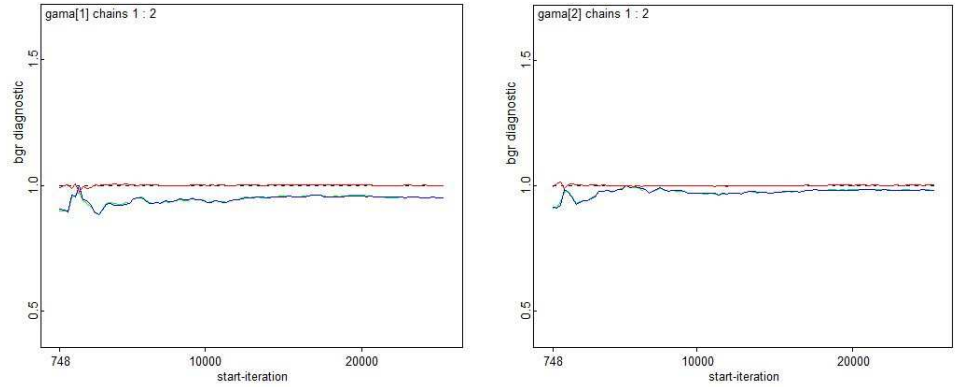

Figure E.5: BGR Diagnostic of Parameter γ

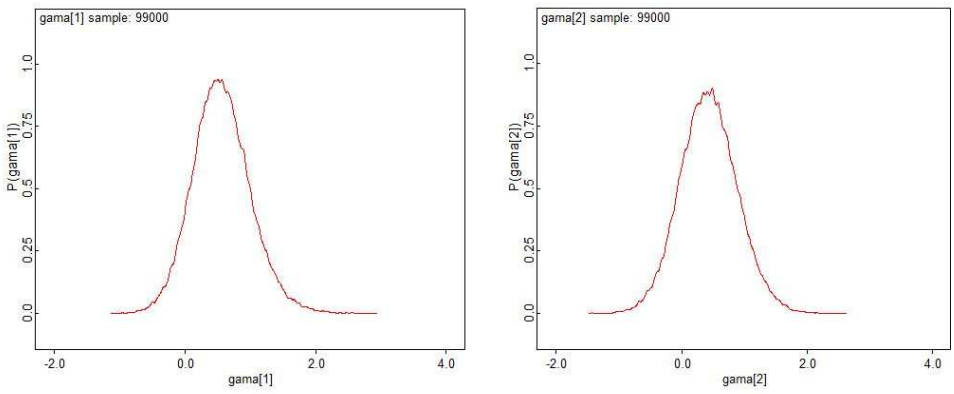

Figure E.6: Posterior Density of Parameter γ

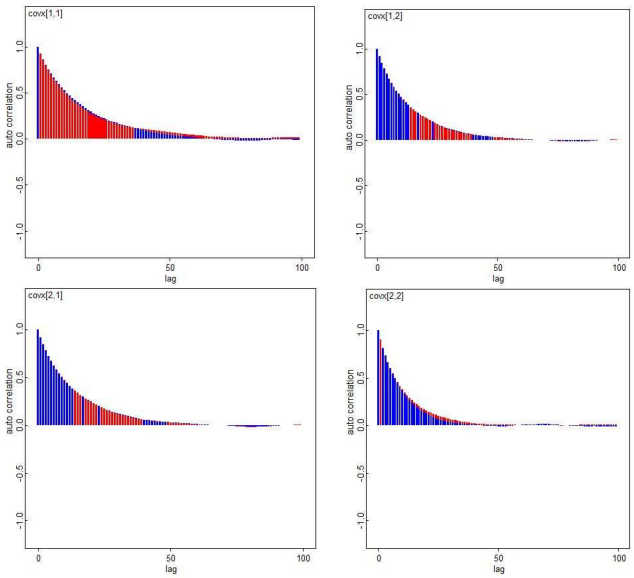

Figure E.7: Auto-Correlation of Parameter $\boldsymbol{\Sigma_X}$

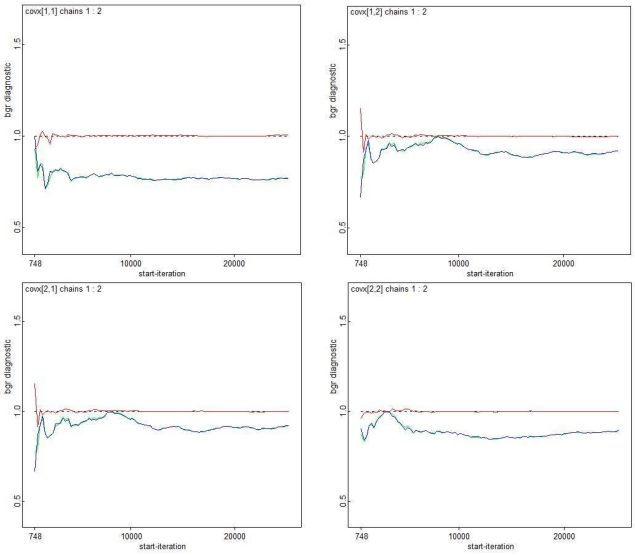

Figure E.8: BGR Diagnostic of Parameter $\boldsymbol{\Sigma_X}$

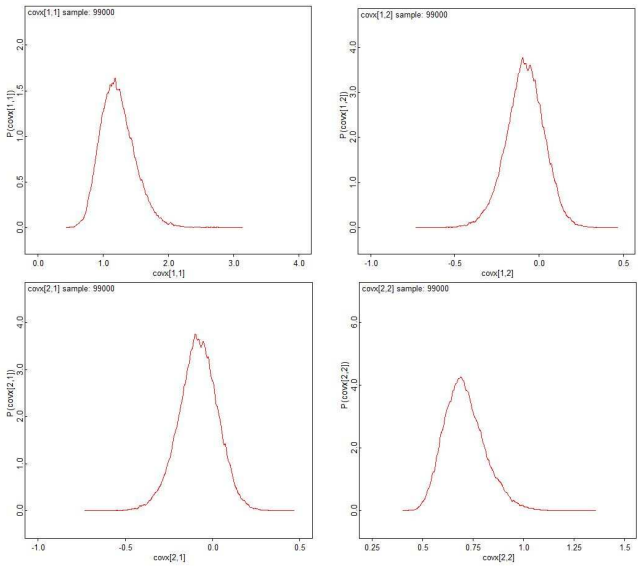

Figure E.9: Posterior Density of Parameter $\boldsymbol{\Sigma_X}$

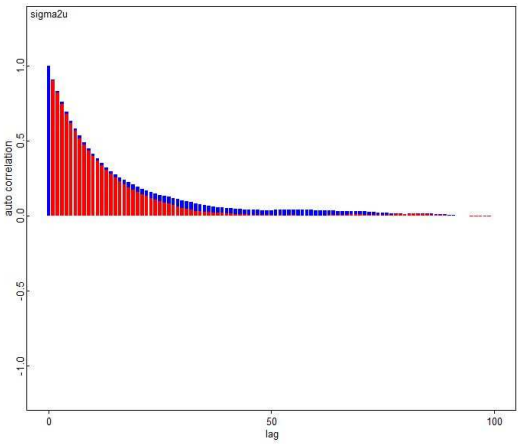

Figure E.10: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

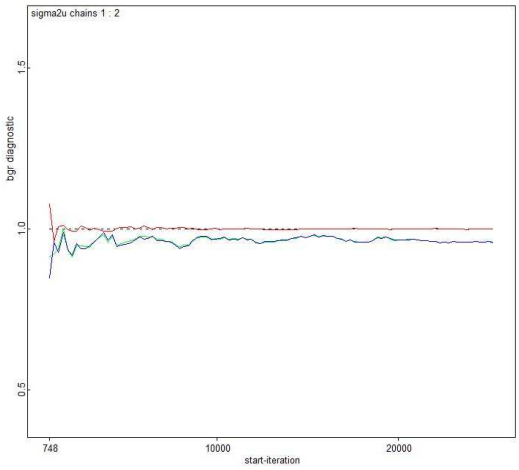

Figure E.11: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

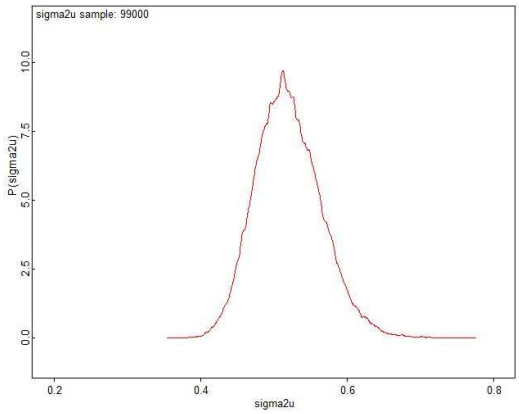

Figure E.12: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

APPENDIX F

CONVERGENCE DIAGNOSTICS FOR JMM-ST WITH DPM APPLICATION

JMM-St with DPM model convergence diagnostics and posterior density of parameters are given in this chapter.

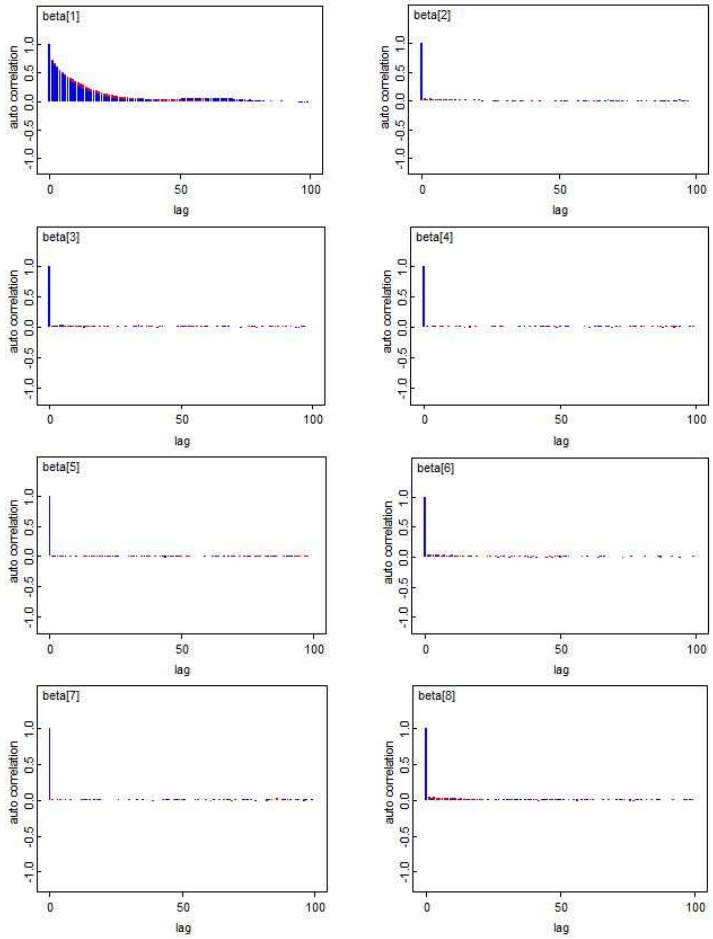

Figure F.1: Auto-Correlation of Parameter β

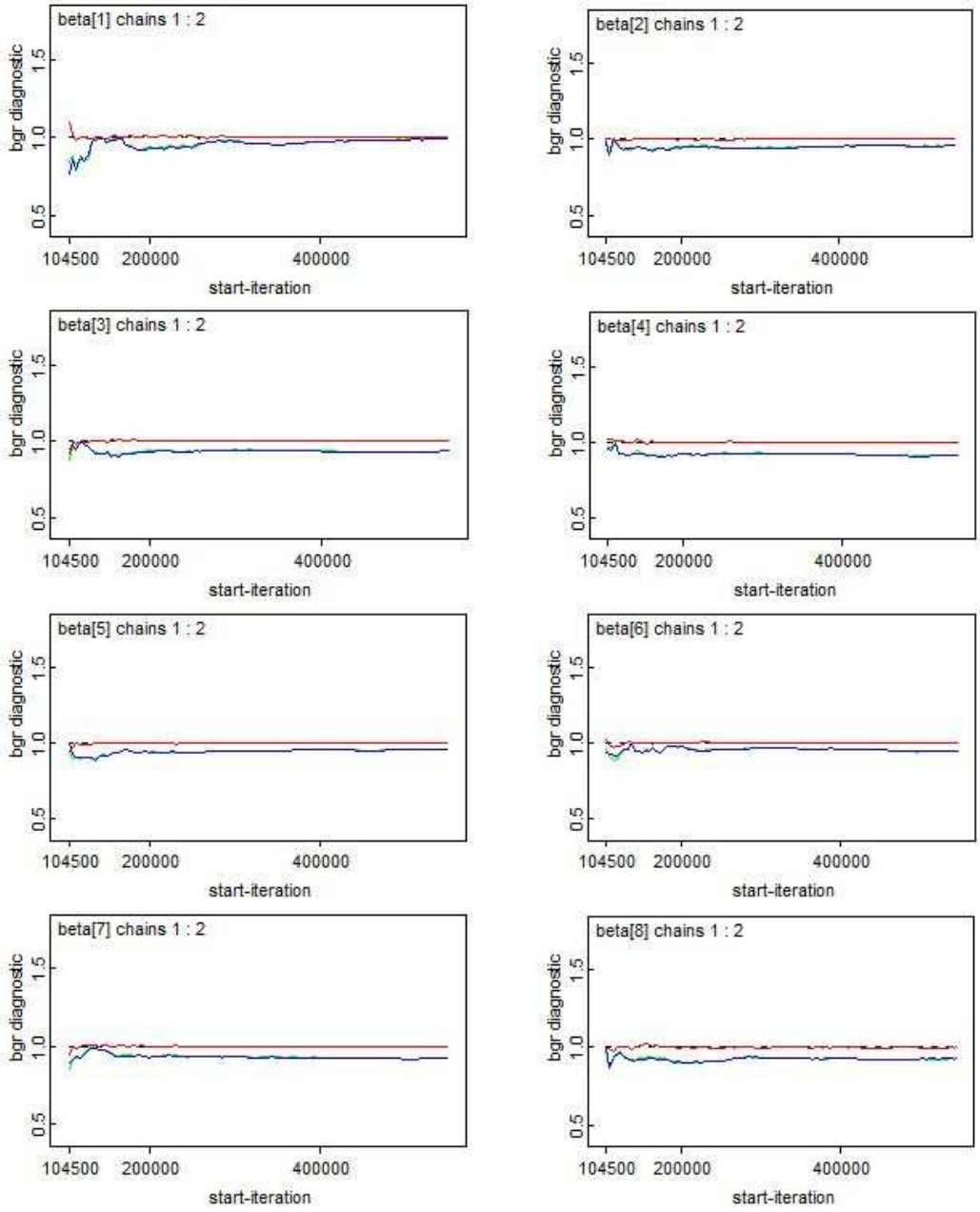

Figure F.2: BGR Diagnostic of Parameter β

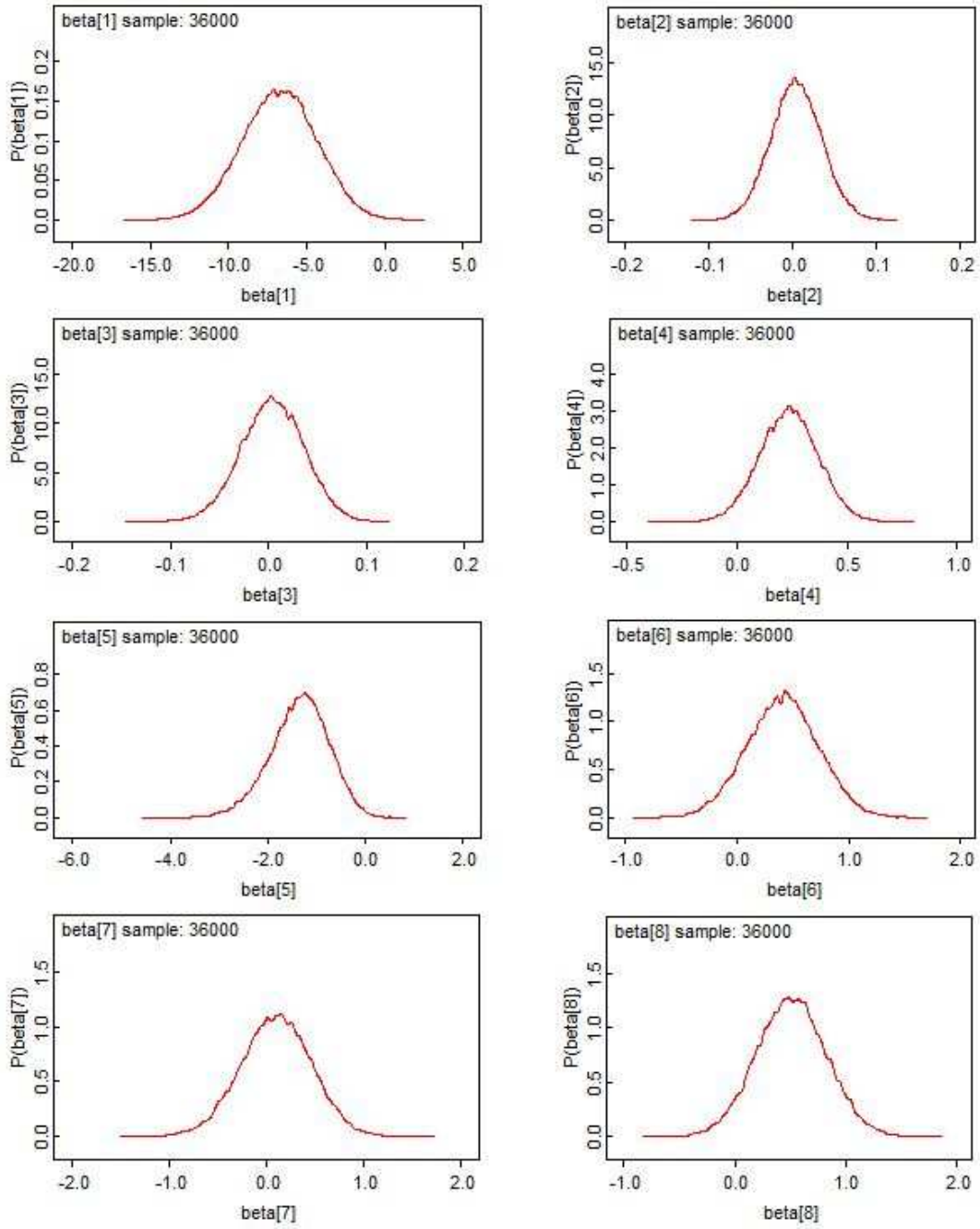

Figure F.3: Posterior Density of Parameter β

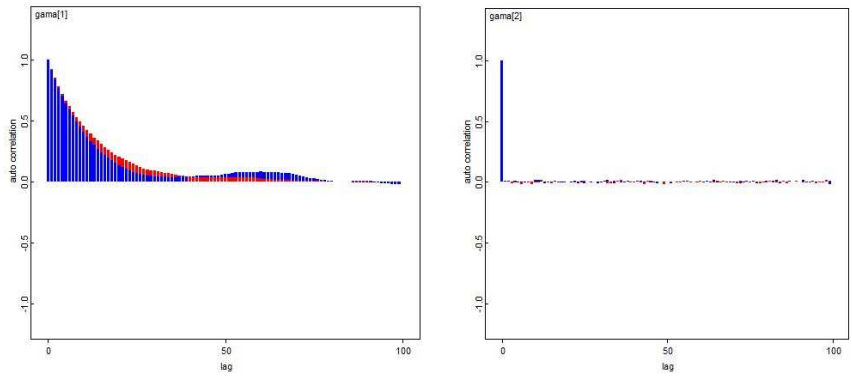

Figure F.4: Auto-Correlation of Parameter γ

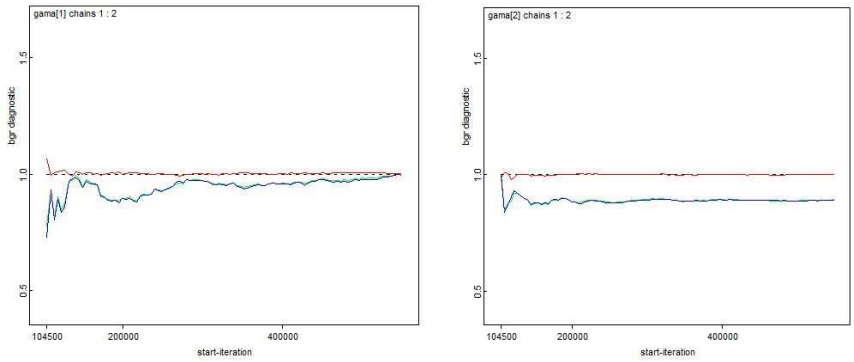

Figure F.5: BGR Diagnostic of Parameter γ

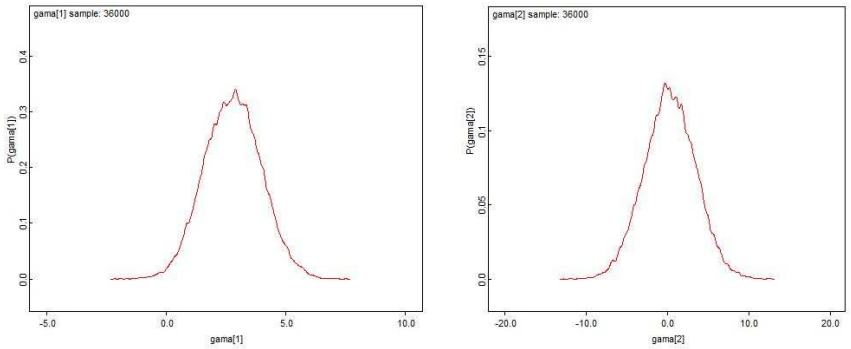

Figure F.6: Posterior Density of Parameter γ

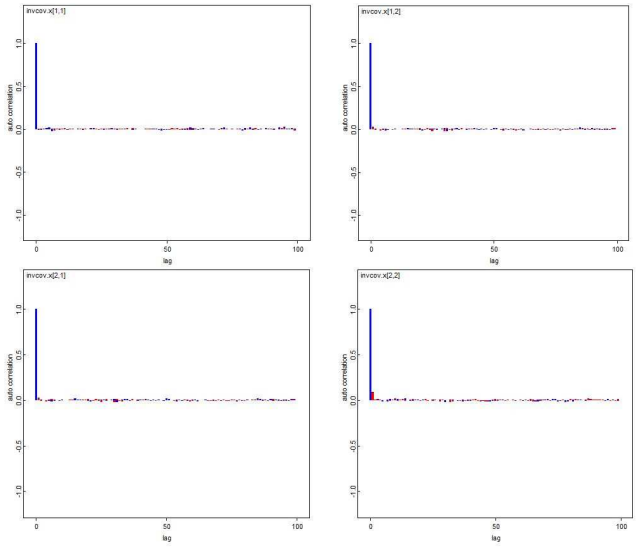

Figure F.7: Auto-Correlation of Parameter inverse $\boldsymbol{\Sigma_X}$

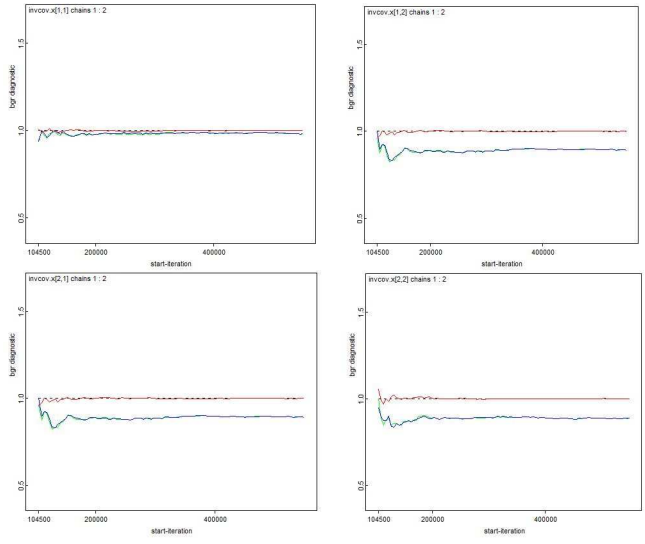

Figure F.8: BGR Diagnostic of Parameter inverse $\boldsymbol{\Sigma_X}$

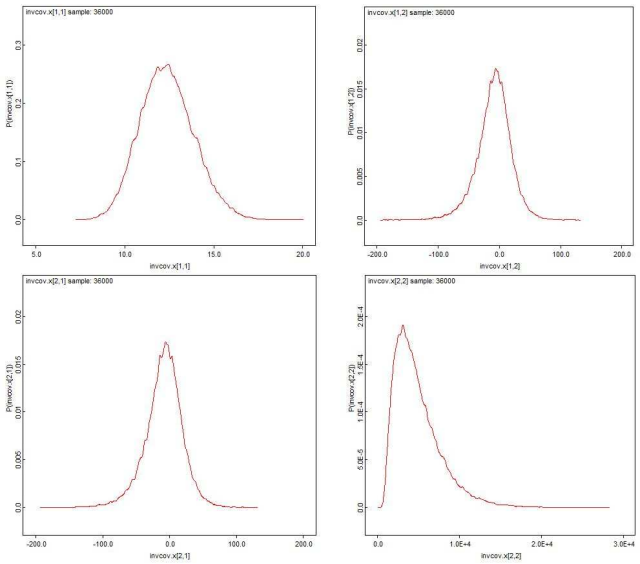

Figure F.9: Posterior Density of Parameter inverse $\boldsymbol{\Sigma_X}$

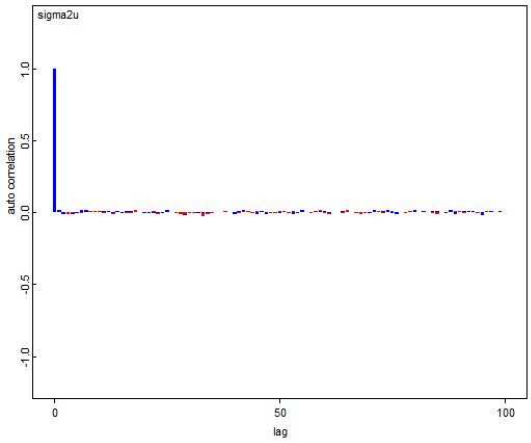

Figure F.10: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

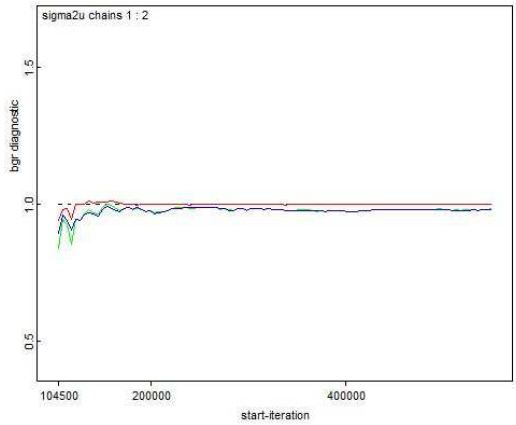

Figure F.11: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

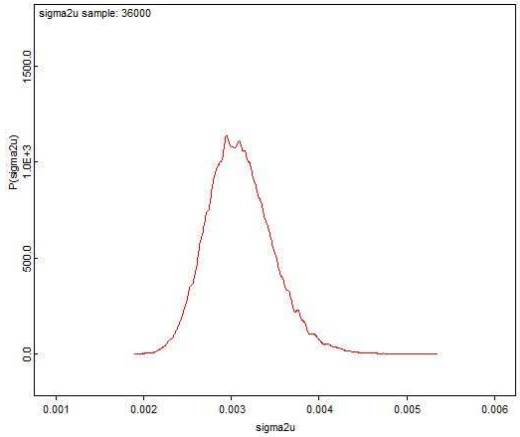

Figure F.12: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

APPENDIX G

CONVERGENCE DIAGNOSTICS FOR JMM-KD WITH DPM APPLICATION

JMM-KD with DPM model convergence diagnostics and posterior density of parameters are given in this chapter.

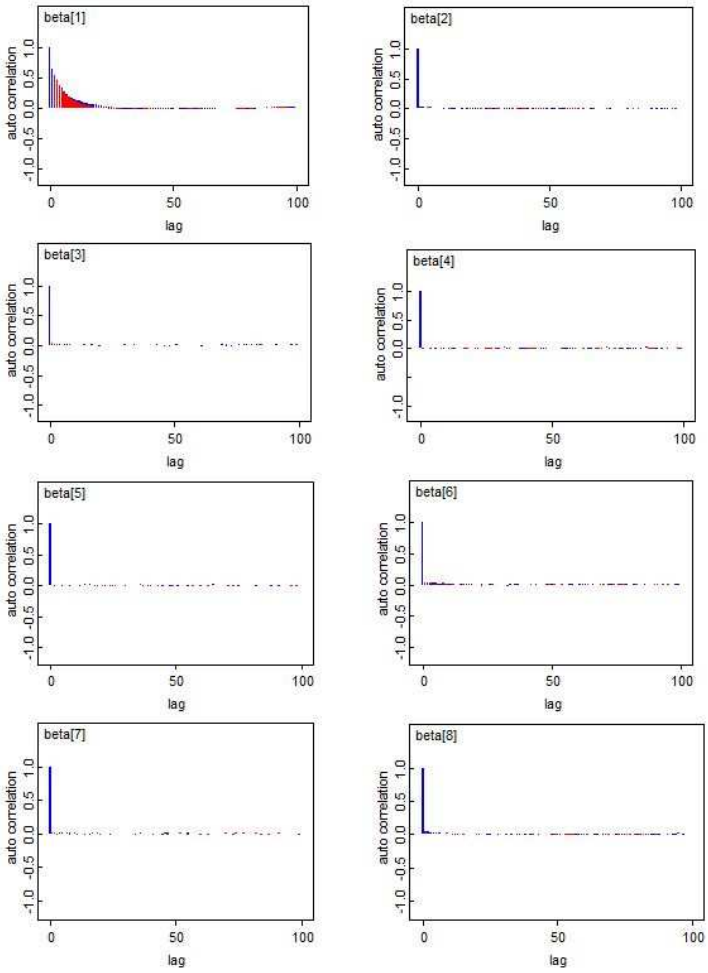

Figure G.1: Auto-Correlation of Parameter β

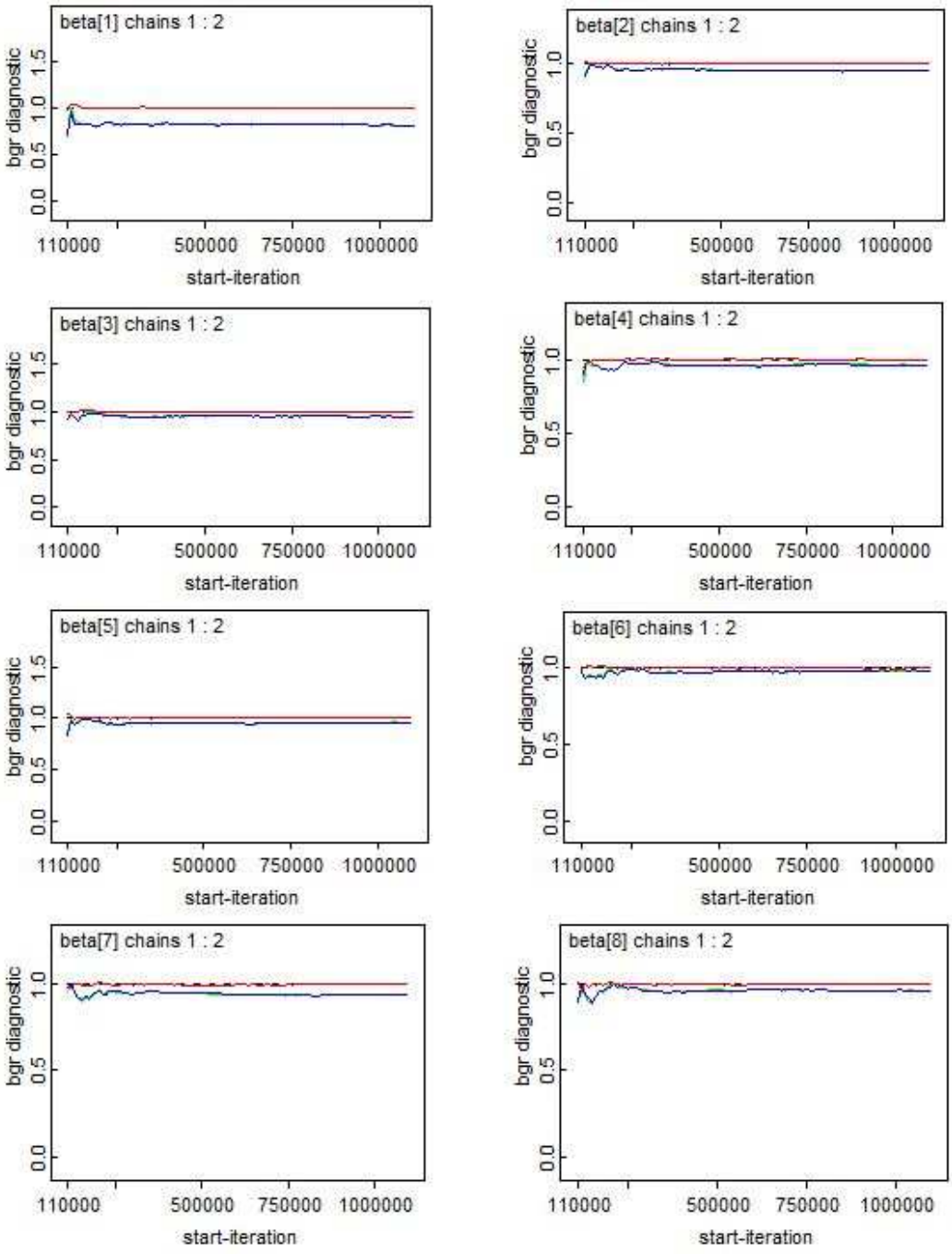

Figure G.2: BGR Diagnostic of Parameter β

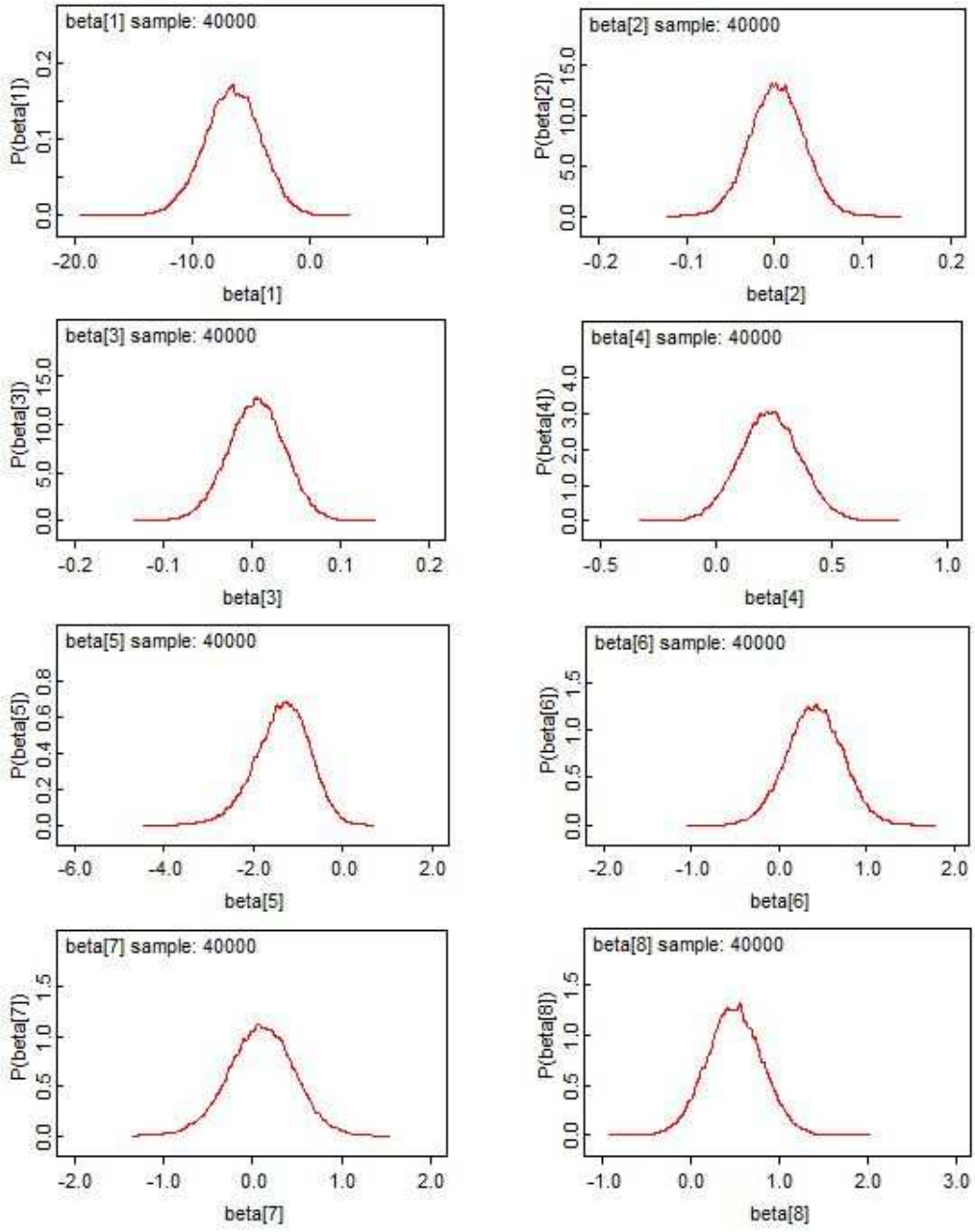

Figure G.3: Posterior Density of Parameter β

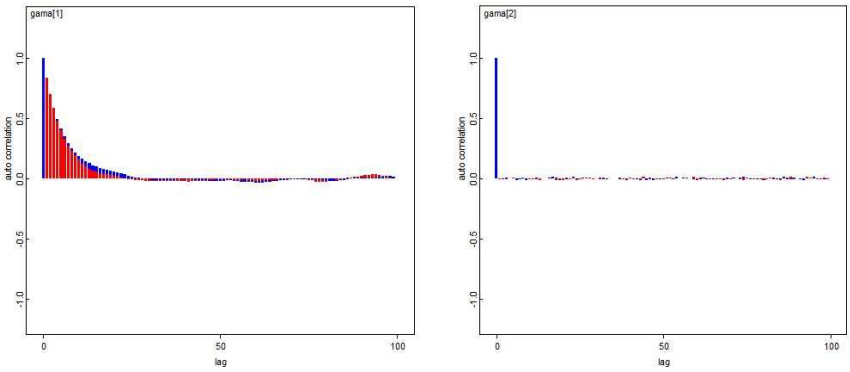

Figure G.4: Auto-Correlation of Parameter γ

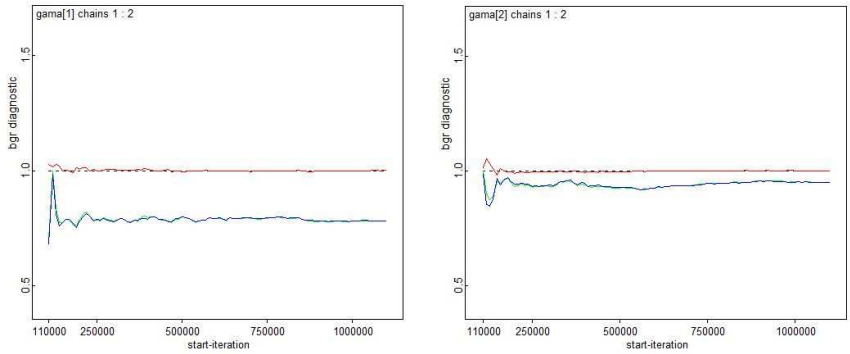

Figure G.5: BGR Diagnostic of Parameter γ

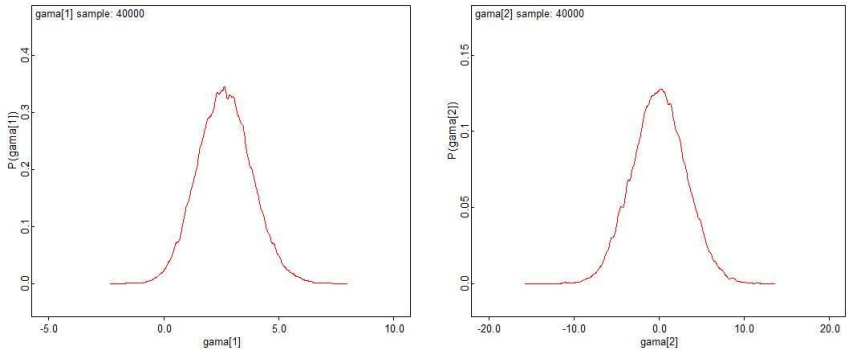

Figure G.6: Posterior Density of Parameter γ

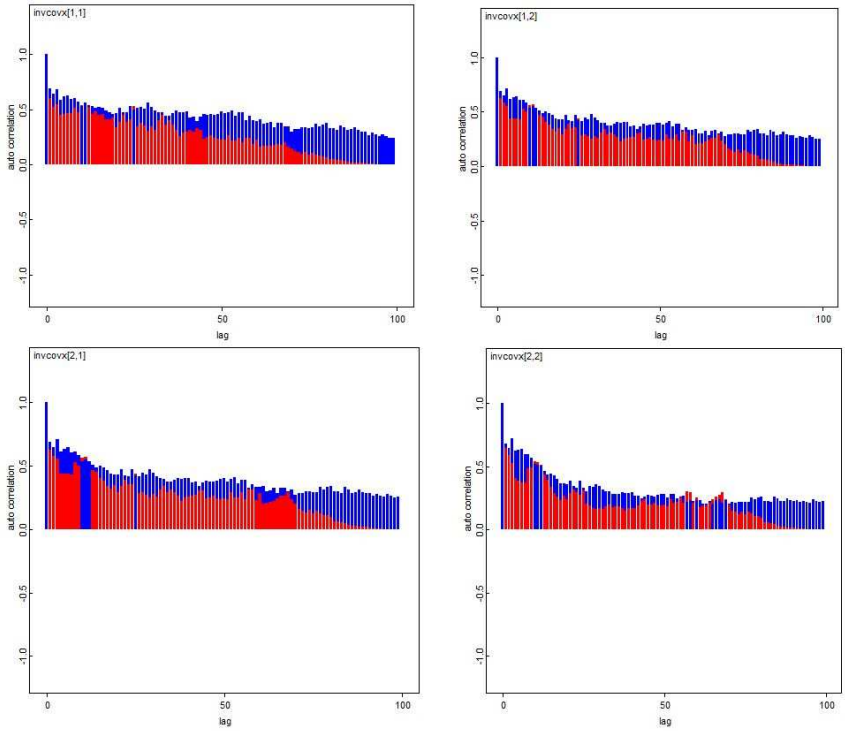

Figure G.7: Auto-Correlation of Parameter $\boldsymbol{\Sigma_X}$

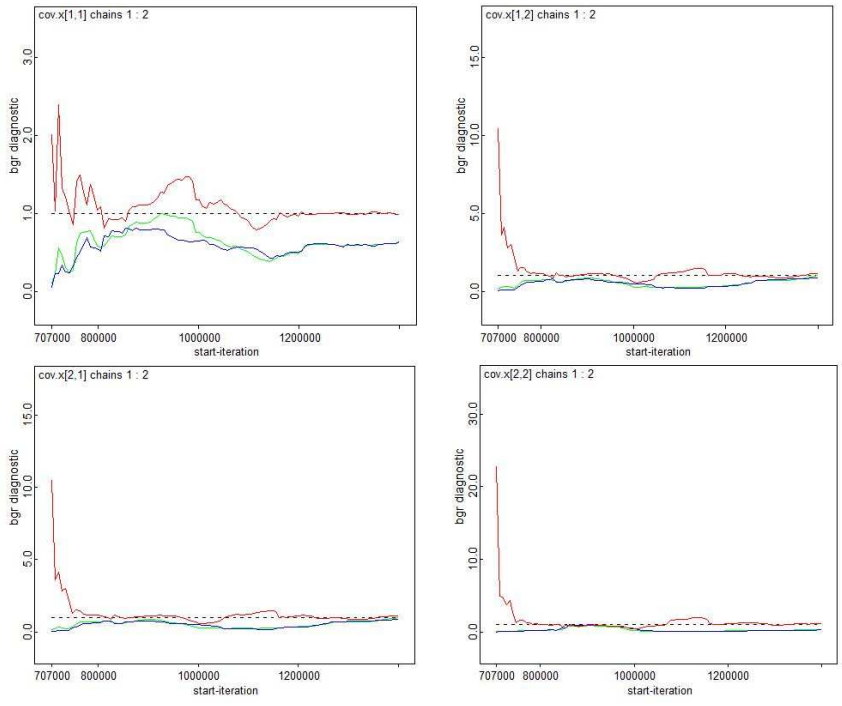

Figure G.8: BGR Diagnostic of Parameter $\boldsymbol{\Sigma_X}$

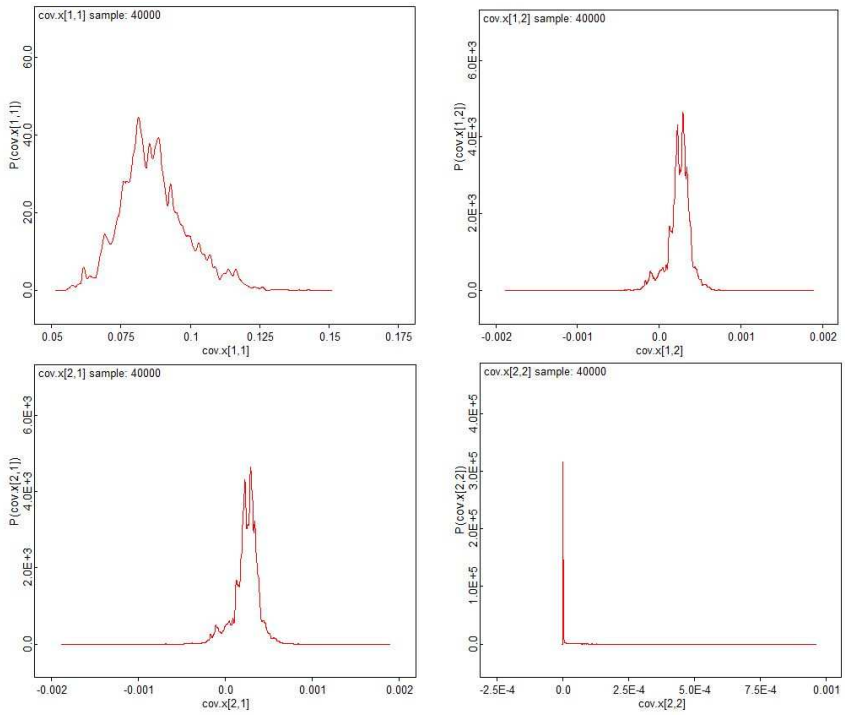

Figure G.9: Posterior Density of Parameter $\boldsymbol{\Sigma_X}$

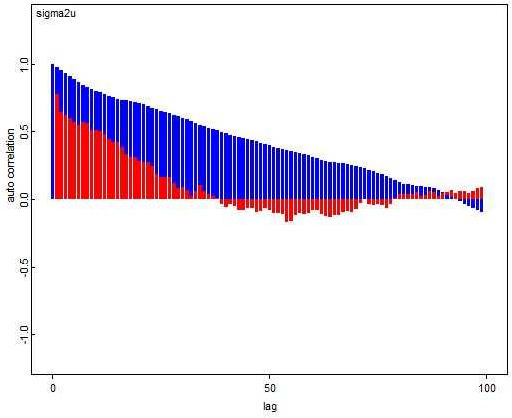

Figure G.10: Auto-Correlation of Parameter $\boldsymbol{\sigma_U^2}$

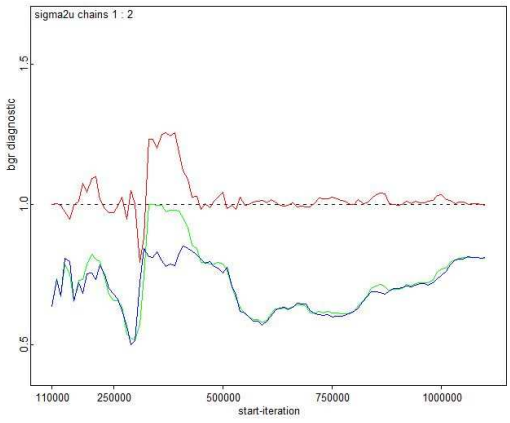

Figure G.11: BGR Diagnostic of Parameter $\boldsymbol{\sigma_U^2}$

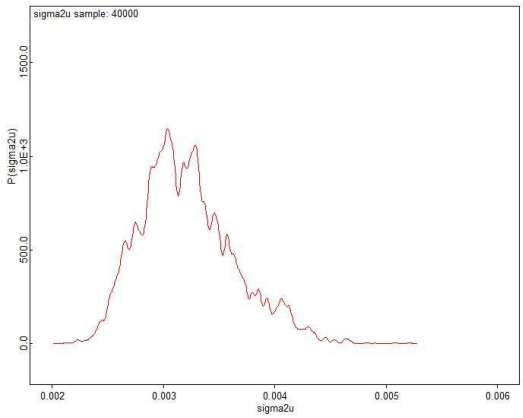

Figure G.12: Posterior Density of Parameter $\boldsymbol{\sigma_U^2}$

APPENDIX H

OPENBUGS CODES

H.1 JMM-St with DPM

```
model
    {
    for( i in 1 : N ) { #over the clusters(persons)

        #model for y (binary cross sectional response)
        y[i] ~ dbern(p[i])
        logit(p[i]) <- beta + gama[1] * x[i,1] + gama[2] * x[i,2]

    #model for random effects
            S[i] ~ dcat(pi[])
            for (d in 1:2){
            muz[i,d] <-  kappa[S[i],d]
                }

            x[i,1:2] ~ dmnorm(muz[i,1:2],invcovx[1:2,1:2])
            for (j in 1:C){
            SC[i,j] <- equals(j,S[i])
                            }
                        } #i
        #Precision Parameter
        # alpha <- 1
```

```
        alpha ~ dgamma(1,1)

              #Constructive DPP
              pz[1] <- r[1]
              for (j in 2:C){
              pz[j] <- r[j] * (1-r[j - 1]) * pz[j - 1 ]/r[j-1]
                                            }
              p.sum <- sum(pz[])
              for (j in 1:C){
              kappa[j,1:2] ~ dmnorm(mu[ ],invcovx[1:2,1:2])


              #r[j] ~ dbeta(1,alpha)
r[j] <- 1 - exp(-rstar[j])
              rstar[j] ~ dexp(alpha)
              #Scaling to Ensure Sum to 1
              pi[j] <- pz[j]/p.sum
                       }

              #Total Cluster
                       K <- sum(c1[])
                       for (j in 1:C){
                              sumSC[j] <- sum(SC[ ,j])
                              c1[j] <- step(sumSC[j]-1)
                       }

    for( k in 1 : Ntot ) { #over all the repeated  observations


    #model for w (continuous longitudinal covariate):
        w[k] ~ dnorm(meanw[k],tauu)
```

```
        meanw[k] <-    x[ids[k],1] + x[ids[k],2]*t[k]
             } #k

        beta ~ dnorm(0,0.01)
        gama[1:2] ~ dmnorm(mugama[1:2],invcovgama[1:2,1:2])

tauu ~ dgamma(a,b)
sigma2u <- 1/tauu

invcovx[1:2,1:2] ~ dwish(R[1:2,1:2],dof)


         } #model
```

H.2 JMM-KD with DPM

```
model
    {
        for( i in 1 : N ) { #over the clusters(persons)

        #model for y (binary cross sectional response)
        y[i] ~ dbern(p[i])
        logit(p[i]) <- beta + gama[1] * x[i,1] + gama[2] * x[i,2]

            #model for random effects (DPM)
                S[i] ~ dcat(pi[])
                for (d in 1:2){
                muz[i,d] <-  kappa[S[i],d]
                              }

                x[i,1:2] ~ dmnorm(muz[i,1:2],invcovx[1:2,1:2])
```

```
                for (j in 1:C){
                        SC[i,j] <- equals(j,S[i])
                            }
                         } #i
                #Precision Parameter
                  #alpha <- 1
            alpha ~ dgamma(1,1)

             #Constructive DPP
                pz[1] <- r[1]
                for (j in 2:C){
                pz[j] <- r[j] * (1-r[j - 1]) * pz[j - 1 ]/r[j-1]
                            }
                p.sum <- sum(pz[])
                for (j in 1:C){
                kappa[j,1:2] ~ dmnorm(mu[ ],invcovx[1:2,1:2])

                #r[j] ~ dbeta(1,alpha)
r[j] <- 1 - exp(-rstar[j])
                rstar[j] ~ dexp(alpha)
                #Scaling to Ensure Sum to 1
                pi[j] <- pz[j]/p.sum
                        }

                 #Total Cluster
                K <- sum(c1[])
                for (j in 1:C){
                sumSC[j] <- sum(SC[ ,j])
                c1[j] <- step(sumSC[j]-1)
                         }
```

```
for( k in 1 : Ntot ) { #over all the repeated  observations

 #model for w (continuous longitudinal covariate):
 w[k] ~ dnorm(meanw[k],tauu)
 meanw[k] <-    x[ids[k],1] + x[ids[k],2]*t[k]
              } #k

 beta ~ dnorm(0,0.01)
 gama[1:2] ~ dmnorm(mugama[1:2],invcovgama[1:2,1:2])

     #G-MVLG prior for variances of x and w model

     tauu <- 1/sigma2u
     sigma2u <- exp(theta[4])
     invcovx[1:2,1:2] <- inverse(covx[1:2,1:2])
     covx[1,1] <- sigma2.11
     covx[1,2] <- sigma2.12
     covx[2,1] <- sigma2.12
     covx[2,2] <- sigma2.22

     #for positive definitness of covx
     sigma2.11 <- l11 * l11
     sigma2.12 <-  l21 * l11
     sigma2.22 <- l21*l21+ l22*l22

     l11 <- exp(theta[1])
     l21 <- theta[2]
     l22 <- exp(theta[3])

     theta[1] ~ dflat()
     theta[2] ~ dflat()
```

```
        theta[3] ~ dflat()
        theta[4] ~ dflat()

        dummy <- 0
        dummy ~ dloglik(phi)
        phi <- nu * log(delta) + log(sum(v[1:60]))

for( i in 1 : 60 ) {
    v[i]  <- (v1[i]  / v2[i]) *  v3[i]
    v1[i]  <- pow((1-delta),nu) * (mus[1] *
pow(lambdas[1],-nu-i)) * (mus[2] * pow(lambdas[2],-nu-i)) * (mus[3] *
pow(lambdas[3],-nu-i)) * (mus[4] * pow(lambdas[4],-nu-i))
    v2[i]  <- (pow(exp(loggam(nu+i)),3) *
exp(loggam(nu)) * exp(logfact(i)))    #
    v3[i]  <- exp((nu+i) * (mus[1]*theta[1] +
mus[2]*theta[2] + mus[3]*theta[3] + mus[4]*theta[4]) - ((1/lambdas[1])
* exp(mus[1]*theta[1]) + (1/lambdas[2]) * exp(mus[2]*theta[2]) +
(1/lambdas[3]) * exp(mus[3]*theta[3]) + (1/lambdas[4]) *
exp(mus[4]*theta[4])))
                                          }

        } #model
```