Parameter estimation of multivariate multiple regression model using bayesian with non-informative Jeffreys' prior distribution

To cite this article: D R S Saputro *et al* 2018 *J. Phys.: Conf. Ser.* **1022** 012002

View the article online for updates and enhancements.

Parameter estimation of multivariate multiple regression model using bayesian with non-informative Jeffreys' prior distribution

D R S Saputro[1], F Amalia[2], P Widyaningsih[3] and R C Affan[4]

[1,2,3]Department of Mathematics, Universitas Sebelas Maret

Corresponding email: dewiretnoss@staff.uns.ac.id, firdaamalia021@gmail.com, purnami_w@staff.uns.ac.id

Abstract. Bayesian method is a method that can be used to estimate the parameters of multivariate multiple regression model. Bayesian method has two distributions, there are prior and posterior distributions. Posterior distribution is influenced by the selection of prior distribution. Jeffreys' prior distribution is a kind of Non-informative prior distribution. This prior is used when the information about parameter not available. Non-informative Jeffreys' prior distribution is combined with the sample information resulting the posterior distribution. Posterior distribution is used to estimate the parameter. The purposes of this research is to estimate the parameters of multivariate regression model using Bayesian method with Non-informative Jeffreys' prior distribution. Based on the results and discussion, parameter estimation of β and Σ which were obtained from expected value of random variable of marginal posterior distribution function. The marginal posterior distributions for β and Σ are multivariate normal and inverse Wishart. However, in calculation of the expected value involving integral of a function which difficult to determine the value. Therefore, approach is needed by generating of random samples according to the posterior distribution characteristics of each parameter using Markov chain Monte Carlo (MCMC) Gibbs sampling algorithm.

1. Introduction

Multivariate multiple regression model is used to modeling the relationship between *m* responses and a single set of predictor variables (Johnson and Wichern [8]). Multivariate multiple regression model has unkown parameters. Parameters are the characteristic of population. The parameters value is obtained from parameter estimation. According to Bolstad [4], there are two statistical approaches that generally are used to estimate the parameters. The first is the classical approach which based on all information from the random sample. The second approach is Bayesian method which combine the information from the random sample and the information of previous research to estimate the parameters. In Bayesian method the information from the random sample is likelihood function and the information of previous research we get prior distribution. According to Harris et al. [6], basic of Bayesian method is to multiplying between likelihood function and prior distribution which proportional with posterior distribution. Posterior distribution is used to estimate the parameters and it is influenced by the selection of prior distribution.

Generally, selecting prior is based on the parameter information available or not. If the parameter information is available, we use informative prior. Informative prior has significant effect on posterior distribution and more subjective (Gelman et al. [5]). If the parameter information is not available, we

use Non-informative prior. Non-informative prior are more objective than most classical analyzes (Yang and Berger [15]). Jeffreys' prior is a kind of Non-informative prior which commonly used in multivariate multiple Bayesian regression model. This prior is good to overcome invariant parameters.

Basically, Bayesian method states that future events can be predicted with terms of previous events that have occurred (Saputro et al. [11]). This research discuss about parameters estimation of multivariate multiple regression model using Bayesian method with Non-informative Jeffreys' prior.

2. The Previous Research

Here are some previous research about this material: Sun and Berger [13] research about the kind of Non-informative prior that could be used for normal multivariate models. Robert and Rousseau [10] research about Non-informative Jeffreys' prior in 2010. Iswari et al. [7] apply Bayesian method on simple linear regression model with Non-informative prior in 2014, while Prasdika et al. [9] in 2016 also examined apply Bayesian methods on simple regression model with informative prior. In the same year, Sinay and Hsu [12] apply the Bayesian method on multivariate regression model with uniform prior. While research about the application of Non-informative Jeffreys' prior on regression model is done by Amalia et al. [1].

3. Results and Discussion

3.1 Multivariate Multiple Regression Model

According to Johnson and Wichern [8], each response variable $Y_1, Y_2, \dots, Y_m$ is assumed to follow its own regression model, so that

$$Y_1 = \beta_{01} + \beta_{11}X_1 + \beta_{21}X_2 + \cdots + \beta_{r1}X_r + \varepsilon_1$$
$$Y_2 = \beta_{02} + \beta_{12}X_1 + \beta_{22}X_2 + \cdots + \beta_{r2}X_r + \varepsilon_2$$
$$\vdots$$
$$Y_m = \beta_{0m} + \beta_{1m}X_1 + \beta_{2m}X_2 + \cdots + \beta_{rm}X_r + \varepsilon_m.$$

the error term $\boldsymbol{\varepsilon}^T = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m]$ has $E[\boldsymbol{\varepsilon}] = 0$, and $Var[\boldsymbol{\varepsilon}] = \Sigma$. Σ is matrix variance ordo m × m. Let $\boldsymbol{X_j} = [X_{j0}, X_{j1}, \dots, X_{jr}]$ denote the value of predictor variables for the jth trial, let $\boldsymbol{Y_j^T} = [Y_{j1}, Y_{j2}, \dots, Y_{jm}]$ be the responses, and $\boldsymbol{\beta_j} = [\beta_{0j}, \beta_{1j}, \dots, \beta_{jr}]$ is regression parameter. $\boldsymbol{Y_j}$ has multivariate normal distribution $\boldsymbol{Y_j} \sim N_m(\boldsymbol{X_j\beta}, \boldsymbol{\Sigma})$. The density function for normal multivariate is

$$f(\boldsymbol{Y_j}|\boldsymbol{\beta}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{m}{2}} |\Sigma|^{-\frac{1}{2}} exp\left(-\frac{1}{2}\left(\boldsymbol{Y_j} - \boldsymbol{X_j\beta}\right)^T \Sigma^{-1}(\boldsymbol{Y_j} - \boldsymbol{X_j\beta})\right).$$

3.2 Bayesian Method

Basic of Bayesian method is to determine the likelihood function and prior distribusion, after that determine the posterior distribution from multiplying the likelihood function and prior distribusion. Parameter estimation is the expected value from posterior marginal distribution. Prior distribution function denote by $f(\theta)$ with θ is parameter. Likelihood function is joint probability function from Y as the random sampel, if θ is known and denote by $f(Y|\theta)$, and the posterior distribution function denote by $f(\theta|Y)$ as follows

$$f(\theta|Y) = \frac{f(\theta, Y)}{f(Y)} = \frac{f(\theta)f(Y|\theta)}{f(Y)} \propto f(\theta)f(Y|\theta).$$

3.3 Likelihood Function

The joint density function of m random variable $Y_{j1}, Y_{j2}, \dots, Y_{jm}$ evaluated at $y_{j1}, y_{j2}, \dots, y_{jm}$, with j=1,2,...,n say f(Y|θ), is referred to as a likelihood function, then

$$f(Y|\theta) = \prod_{j=1}^{n} f(Y_j|\theta) = f(Y_1|\theta)f(Y_2|\theta). \dots . f(Y_n|\theta).$$

The likelihood function with random variable has multivariate normal distribution $\boldsymbol{Y_j} \sim N_m(\boldsymbol{X_j}\beta, \boldsymbol{\Sigma})$, is

$$f(\boldsymbol{Y}|\boldsymbol{\beta},\boldsymbol{\Sigma}) = \prod_{j=1}^{n} (2\pi)^{-\frac{m}{2}} |\Sigma|^{-\frac{1}{2}} exp\left(-\frac{1}{2}(\boldsymbol{Y_j} - \boldsymbol{X_j\beta})^T \Sigma^{-1}(\boldsymbol{Y_j} - \boldsymbol{X_j\beta})\right)$$

$$= (2\pi)^{-\frac{nm}{2}} |\Sigma|^{-\frac{n}{2}} exp\left(tr\left(\sum_{j=1}^{n} -\frac{1}{2}(\boldsymbol{Y_j} - \boldsymbol{X_j\beta})^T \Sigma^{-1}(\boldsymbol{Y_j} - \boldsymbol{X_j\beta})\right)\right)$$

$$= (2\pi)^{-\frac{nm}{2}} |\Sigma|^{-\frac{n}{2}} exp\left(tr\left(-\frac{n}{2}(\overline{\boldsymbol{Y}} - \boldsymbol{X_j\beta})^T \Sigma^{-1}(\overline{\boldsymbol{Y}} - \boldsymbol{X_j\beta}) - \frac{1}{2}\Sigma^{-1}\boldsymbol{S}\right)\right)$$

with $\overline{\boldsymbol{Y}} = \frac{1}{n}\sum_{j=1}^{n}\boldsymbol{Y_j}$ and $\boldsymbol{S} = \sum_{j=1}^{n}(\boldsymbol{Y_j} - \overline{\boldsymbol{Y}})^T (\boldsymbol{Y_j} - \overline{\boldsymbol{Y}})$.

3.4 Non-informative Jeffreys' prior

According to Harris et al. [6] prior distribution is previous distribution information. Usually, knowledge of the prior will be available and included in the analysis as informative prior. Knowledge of the prior may be obtained from the opinions of experts or because it reuses the posterior distribution from the previous research, but if the knowledge of the prior is uncertain, lost, or ignored, non-informative prior is used.

Jeffreys' prior is kind of Non-informative prior. Jeffreys' prior is square root of Fisher information written by $f(\boldsymbol{\beta}) = [I(\boldsymbol{\beta})]^{\frac{1}{2}}$ with $I(\boldsymbol{\beta})$ is Fisher's information. According to Bain and Engelhardt [3], Fisher's information is written as

$$I(\boldsymbol{\beta}) = -E_\beta\left(\frac{\partial^2 \log f(\boldsymbol{Y}|\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2}\right)$$

If $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)^T$ is a vector, it is used $f(\boldsymbol{\theta}) = [\det \mathbf{I}(\boldsymbol{\theta})]^{\frac{1}{2}}$ with $\mathbf{I}(\boldsymbol{\theta})$ is Fisher's information matrix ordo (pxp). The Fisher Information is

$$\mathbf{I_{ij}}(\boldsymbol{\theta}) = -E_\theta\left(\frac{\partial^2 \log f(\boldsymbol{Y}|\boldsymbol{\theta})}{\partial \theta_i \, \partial \theta_j}\right)$$

with $i = 1,2,\ldots,p$ and $j = 1,2,\ldots,p$.

Non-informative prior of multivariate multiple regression is $f(\boldsymbol{\theta})$ with $\boldsymbol{\theta} = (\boldsymbol{\beta}, \Sigma)$, assumsed $\boldsymbol{\beta}$ and Σ is independent so $f(\boldsymbol{\theta}) = f(\boldsymbol{\beta})f(\Sigma)$. Furthermore, a Non-informative prior distribution $f(\boldsymbol{\theta})$ is determined from the following steps.

Non-informative Jeffreys' prior for $\boldsymbol{\beta}$ is

$$\mathbf{I}(\boldsymbol{\beta}) = -E\left(\frac{\partial^2}{\partial \boldsymbol{\beta}^2}\log f(\boldsymbol{Y}|\boldsymbol{\theta})\right) = -E(-\Sigma^{-1}) = \Sigma^{-1}$$

$$f(\boldsymbol{\beta}) = \sqrt{\mathbf{I}(\boldsymbol{\beta})} = \sqrt{\Sigma^{-1}} = \Sigma^{-\frac{1}{2}}.$$

Multivariate multiple regression model with *m* dimensi has Non-informative Jeffreys' prior for $\boldsymbol{\beta}$

$$f(\boldsymbol{\beta}) = f(\beta_1) \times f(\beta_2) \times \ldots \times f(\beta_m) = \Sigma^{-\frac{1}{2}} \times \Sigma^{-\frac{1}{2}} \times \ldots \times \Sigma^{-\frac{1}{2}} = \Sigma^{-\frac{m}{2}}$$

Non-informative Jeffreys' prior for Σ is

$$\mathbf{I}(\Sigma) = -E\left(\frac{\partial^2}{\partial \Sigma^2}\log f(\boldsymbol{Y}|\boldsymbol{\theta})\right) = -E(-(\boldsymbol{Y_j} - \boldsymbol{X_j\beta})^T \Sigma^{-3}(\boldsymbol{Y_j} - \boldsymbol{X_j\beta})) = \Sigma^{-2}$$

$$f(\Sigma) = \sqrt{\mathbf{I}(\Sigma)} = \sqrt{\Sigma^{-2}} = \Sigma^{-1}.$$

Non-informative Jeffreys' prior for $\boldsymbol{\theta} = (\boldsymbol{\beta}, \Sigma)$ is

$$f(\boldsymbol{\theta}) = f(\beta_1, \beta_2, \dots, \beta_m)f(\Sigma) = \boldsymbol{\Sigma}^{-\frac{m}{2}} \boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}^{-\frac{m+2}{2}}$$

Thus Non-informative Jeffreys' prior for multivariate multiple regression models is $f(\boldsymbol{\beta}, \Sigma) = \Sigma^{-\frac{m+2}{2}}$.

3.5   Posterior Distribution

After obtaining likelihood function and prior distribution, we determine the posterior distribution. The posterior distribution function for multivariate multiple regression function is given below

$$f(\boldsymbol{\beta}, \boldsymbol{\Sigma}|\mathbf{Y}) \propto |\boldsymbol{\Sigma}|^{-\frac{m+2}{2}} (2\pi)^{-\frac{nm}{2}} |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left(\text{tr}\left(-\frac{n}{2}(\overline{\mathbf{Y}} - \mathbf{X_j}\boldsymbol{\beta})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}(\overline{\mathbf{Y}} - \mathbf{X_j}\boldsymbol{\beta}) - \frac{1}{2}\boldsymbol{\Sigma}^{-1}\mathbf{S} \right) \right)$$

$$\propto |\boldsymbol{\Sigma}|^{-\frac{1}{2}}|\acute{\mathbf{O}}|^{-\frac{m+n+1}{2}} \exp\left(\text{tr}\left(-\frac{n}{2}(\overline{\mathbf{Y}} - \mathbf{X_j}\boldsymbol{\beta})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}(\overline{\mathbf{Y}} - \mathbf{X_j}\boldsymbol{\beta}) \right) \right) \exp\left(\text{tr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\mathbf{S} \right) \right)$$

$$\propto \left(|\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(\text{tr}\left(-\frac{n}{2}(\overline{\mathbf{Y}} - \mathbf{X_j}\boldsymbol{\beta})^{\mathrm{T}} \Sigma^{-1}(\overline{\mathbf{Y}} - \mathbf{X_j}\boldsymbol{\beta}) \right) \right) \right) \left(|\boldsymbol{\Sigma}|^{-\frac{m+n+1}{2}} \exp\left(\text{tr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\mathbf{S} \right) \right) \right)$$

$$\propto \left(\left|\frac{\Sigma}{n}\right|^{-\frac{1}{2}} \exp\left(\text{tr}\left(-\frac{1}{2}(\overline{\mathbf{Y}} - \mathbf{X_j}\boldsymbol{\beta})^{\mathrm{T}} \left(\frac{\Sigma}{n}\right)^{-1}(\overline{\mathbf{Y}} - \mathbf{X_j}\boldsymbol{\beta}) \right) \right) \right)$$

$$\left(|\boldsymbol{\Sigma}|^{-\frac{m+n+1}{2}} \exp\left(\text{tr}\left(-\frac{1}{2}\mathbf{S}\,\boldsymbol{\Sigma}^{-1} \right) \right) \right)$$

The posterior marginal distribution for $\boldsymbol{\beta}$ is $N_m\left(\overline{Y}, \frac{\Sigma}{n}\right)$ and posterior marginal distribution for Σ is $W^{-1}(\boldsymbol{S}^{-1}, n)$.

3.6 Parameter Estimation

The parameter estimation is determined from the expected value or the average value of marginal posterior distribution for the mean $\boldsymbol{\beta}$ and variance $\boldsymbol{\Sigma}$. Given $\boldsymbol{\beta}$ is the matrix of regression parameter of size $(r+1) \times m$ so the parameter estimate for $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = \text{E}[\boldsymbol{\beta}] = \text{E}[\beta_{01}, \beta_{11}, \dots, \beta_{(r+1)\times m}]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \boldsymbol{\beta} \left|\frac{\Sigma}{n}\right|^{-\frac{1}{2}} exp\left(tr\left(-\frac{1}{2}(\overline{Y} - \boldsymbol{X_j}\boldsymbol{\beta})^T \left(\frac{\Sigma}{n}\right)^{-1}(\overline{Y} - \boldsymbol{X_j}\boldsymbol{\beta}) \right) \right) d\beta_{(r+1)\times m} \dots d\beta_{11} d\beta_{01}$$

and the parameter estimate for Σ is

$$\hat{\boldsymbol{\Sigma}} = \text{E}[\boldsymbol{\Sigma}] = \text{E}[\boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{12}, \dots, \boldsymbol{\Sigma}_{m\times m}]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \boldsymbol{\Sigma}|\boldsymbol{\Sigma}|^{-\frac{m+n+1}{2}} exp\left(tr\left(-\frac{1}{2}\boldsymbol{S}\,\boldsymbol{\Sigma}^{-1} \right) \right) d\boldsymbol{\Sigma}_{m\times m} \dots d\boldsymbol{\Sigma}_{12} d\boldsymbol{\Sigma}_{11}.$$

When we estimate $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ using integral dimensions with complex functions that are difficult to determine the solution. It is necessary to approach the complex functions by generating random samples using MCMC Gibbs sampling algorithm. According to Walsh [14] MCMC is used to simulate sampling from posterior distributions. MCMC algorithms has two kind of algorithms, there are Metropolis-Hastings and Gibbs sampling. The Gibbs sampling algorithm is a special occurrence of Metropolis-Hastings, since the parameter is estimated to be more than one (Asriadi [2]). In the Bayesian multivariate regression model there are two parameters that are estimated so that the algorithm used is Gibbs sampling algorithm.

The next step is to generate a random sample of posterior marginal distribution using MCMC Gibbs sampling algorithm. The parameter estimate is obtained after calculating the mean value of the random sample of the generated result.

The Gibbs sampling algorithm for multivariate multiple regression model is

(1) Construck a multivariate multiple regression model to obtain $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma} \cdot$

(2) Initialize value $\boldsymbol{\beta^{(0)}}$ and $\boldsymbol{\Sigma^{(0)}}$ based on the information from first step.

(3) Obtain $\mathbf{Y_j^{(1)}} \sim \text{N}_m\big(\mathbf{X_j}\boldsymbol{\beta^{(0)}}, \boldsymbol{\Sigma^{(0)}}\big)$.

(4) Generate $\boldsymbol{\beta^{(1)}}|\mathbf{Y_j^{(1)}} \sim \text{N}_m\left(\overline{\text{Y}}, \frac{\Sigma}{\text{n}}|\boldsymbol{\beta^{(0)}}\right)$ with $\overline{\text{Y}} = \frac{1}{\text{n}}\sum_{\text{j}=1}^{\text{n}} \mathbf{Y_j}$.

(5) Generate $\mathbf{\Sigma^{(1)}}|\mathbf{Y_j^{(1)}} \sim IW(\mathbf{S^{-1}}, n|\mathbf{\Sigma^{(0)}})$.

(6) Repeat step 3,4, and 5 as much as M repetition until we get $\mathbf{\beta^{(M)}}|\mathbf{Y_j^{(M)}} \sim N_m\left(\overline{Y}, \frac{\Sigma}{n}|\mathbf{\beta^{(M-1)}}\right)$ and

$\mathbf{\Sigma^{(M)}}|\mathbf{Y_j^{(M)}} \sim IW(\mathbf{S^{-1}}, n|\mathbf{\Sigma^{(M-1)}})$.

The Gibbs sampling algorithm generates a sample sequence $\left(\boldsymbol{\beta^{(1)}}, \boldsymbol{\beta^{(2)}}, .., \boldsymbol{\beta^{(M)}}\right)$ dan $\left(\boldsymbol{\Sigma^{(1)}}, \boldsymbol{\Sigma^{(2)}}, ..., \boldsymbol{\Sigma^{(M)}}\right)$. The parameter estimation for $\boldsymbol{\beta}$ is $\widehat{\boldsymbol{\beta}} = \frac{1}{M}\sum_{i=1}^{M} \boldsymbol{\beta^{(i)}}$ and the parameter estimation for $\boldsymbol{\Sigma}$ is $\widehat{\boldsymbol{\Sigma}} = \frac{1}{M}\sum_{i=1}^{M} \boldsymbol{\Sigma^{(i)}}$.

3.7 Application

Bayesian method with Non-informative Jeffreys' prior is used to estimate multivariat multiple regression model from the data. The data is generated from multivariate normal distribution with number of data n = 88, has two dimensions (m = 2), the mean $\overline{y}^T = (19715, \ 6937)$, and the variance $\Sigma = \begin{pmatrix} 2495993 & 11264928 \\ 11264928 & 5329880 \end{pmatrix}$. Data generated using R software. Here are the steps in estimating the parameters.

From the mean and variance we can determine the generate data of posterior distribution function using MCMC Gibbs sampling, so we get $\boldsymbol{\beta^{(i)}}|Y_j^{(i)} \sim N_m\left(\overline{Y}, \frac{\Sigma}{n}|\boldsymbol{\beta^{(i-1)}}\right)$ and $\boldsymbol{\Sigma^{(i)}}|Y_j^{(i)} \sim IW\left(\boldsymbol{S^{-1}}, n|\boldsymbol{\Sigma^{(i-1)}}\right)$, with i=1,2,3,...,M. The parameters estimation in determine from average of the generate data of posterior distribution function. Parameters estimation is shown by Tabel 1.

Table 1. Parameter estimation

Parameter	Mean	percentil 0.25	Percentil 0.975	Explanation
$\widehat{\beta_{01}}$	0.0056794	0.0046285	0.0068708	Significant
$\widehat{\beta_{11}}$	0.0055101	0.0050309	0.0059481	Significant
$\widehat{\beta_{21}}$	79.69864	29.7474	129.1843	Significant
$\widehat{\beta_{02}}$	-0.000897	-0.0008824	-0.001864	Significant
$\widehat{\beta_{12}}$	0.0031068	0.0028224	0.003394	Significant
$\widehat{\beta_{22}}$	82.15982	40.84126	124.8749	Significant
$\widehat{\Sigma_{11}}$	807413.4	598997.2	1089146	Significant
$\widehat{\Sigma_{21}}$	211616.5	100467.5	337936.4	Significant
$\widehat{\Sigma_{22}}$	314092.5	233559.6	423550.0	Significant

The parameter estimation of Bayesian multivariate multiple regression model given below.

$$\hat{\beta} = \begin{pmatrix} \beta_{01} & \beta_{02} \\ \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} = \begin{pmatrix} 0.0056794 & -0.000897 \\ 0.0055101 & 0.0031068 \\ 79.69864 & 82.15982 \end{pmatrix}$$

$$\widehat{\Sigma} = \begin{pmatrix} \Sigma_{11} & \Sigma_{21} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} 807413.4 & 211616.5 \\ 211616.5 & 314092.5 \end{pmatrix}$$

4. Conclusions

Based on results and discussion, the following conclusions is obtained. Parameter estimation of multivariate multiple regression model using Bayesian method was obtained by determining the expected value of the posterior marginal distribution $f(\beta|Y)$ and $f(\Sigma|Y)$. However, in calculation of parameter estimation, involving integral function which is difficult to determine the solution, so we use sampling with Markov chain Monte Carlo (MCMC) Gibbs sampling algorithm as approach in estimating its parameters. The result of MCMC Gibbs sampling algorithm is obtained by the matrix random sample $\left(\beta^{(1)}, \beta^{(2)}, .., \beta^{(M)}\right)$ and $\left(\Sigma^{(1)}, \Sigma^{(2)}, ..., \Sigma^{(M)}\right)$. The parameter estimation for β is $\hat{\beta} = \frac{1}{M}\sum_{i=1}^{M}\beta^{(i)}$ and the parameter estimation for Σ is $\hat{\Sigma} = \frac{1}{M}\sum_{i=1}^{M}\Sigma^{(i)}$.

5. References

[1]   Amalia F, D R S Saputro and T J Parmaningsih 2016 *Estimasi Parameter Model Regresi Linier Sederhana Bayes dengan Distribusi Prior Noninformatif* Seminar Nasional 2016 Matematika dan Pendidikan Matematika

[2]   Asriadi A 2007 Simulasi Stokastik menggunakan Algoritma Metropolis Hastings *JMAP (Jurnal Matematika dan Pembelajaran) Jurusan Matematika FMIPA UNJ* **6**(2) 45-52

[3]   Bain L J and M Engelhardt 1992 *Introduction to probability and mathematical statistics* Second ed. (Buxbury Press, Inc., California)

[4]   Bolstad W M 2007 *Introduction to Bayesian statistics* Second ed. (New Jersey: Wiley)

[5]   Gelman A, J B Carlin, H S Stern, D B Dunson, A Vehtari and D B Rubin, *Bayesian Data Analysis* Second ed. (New York: Chapman and Hall)

[6]   Harris P, M Cox, C Matthews, I Smith and L Wrigh 2015 *a guide to Bayesian inference for regression problems* European Metrologu Research Programme (EMRP)

[7]   Iswari A A I A C, I W Sumarjaya and I G A M Srinadi 2014 Analisis Regresi Bayes Linear Sederhana dengan Prior Noninformatif *E-Jurnal Matematika* **3**(2) 38-44

[8]   Johnson R and D W Wichern 2007 *Applied multivariate statistical analysis* Sixth ed. (Pearson Education, Inc., America)

[9]   Prasdika D A, D R S Saputro and T J Parmaningsih 2016 *Estimasi parameter model regresi linier sederhana bayes dengan distribusi prior informative* Seminar Nasional 2016 Matematika dan Pendidikan Matematika

[10]   Robert C P and J Rousseau and Harold Jeffreyss 2009 Theory of Probability Revisited *Statistical Science* **24**(2) 141-72

[11]   Saputro D R S, P Widyaningsih, F Handayani, and N A Kurdhi 2016 Prior and Posterior Dirichlet Distribution on Bayesian Networks (BNs) *AIP Publishing* **1827**(36)

[12]   Sinay M S and J S J Hsu 2014 Bayesian Inferensi of a multivariate regression model (Hindawi Publishing Corporation)

[13]   Sun D and J O Berger 2006 *Objective priors for the multivariate normal model* (Benidorm)

[14]   Walsh B 2004 Markov Chain Monte Carlo and Gibbs sampling *Lecture Notes for EEB* 581

[15]   Yang R and J O Berger 1998 A *catalog of non-informative priors* Technical Report, Dake University, ISDS 97-42