

Monte Carlo Simulation

10-Page Essay

Xiaolong Wang

Department of Statistics, University of Chicago

Markov chain Monte Carlo (MCMC) methods have become essential tools for Bayesian inference, allowing efficient sampling from complex posterior distributions with unknown normalizing constants. Traditional MCMC algorithms, such as Metropolis-Hastings and Gibbs Samplers, are assumed to sample from a fixed dimensional parameter space. However, in the real world we might want to choose a model that better fits the data from a set of models. For example, variable selection in Bayesian regression, or Bayesian selection from models with different parameter dimensions.

To address this challenge, Peter Green introduced the Reversible Jump MCMC (RJMCMC) algorithm in 1995 [3], which extends the MCMC framework to allow for trans-dimensional moves. Unlike standard MCMC, which samples within a fixed parameter space, RJMCMC enables the exploration of multiple models with different numbers of parameters by "jumping" between spaces of different dimensions. This flexibility makes RJMCMC a powerful tool for Bayesian model selection and structural learning problems where the number of parameters or model complexity needs to be inferred from data.

This paper will introduce RJMCMC with a focus on method and application. The remainder of the paper is structured as follows. Section 1 introduces the RJMCMC method and algorithm details. Section 2 performs a simulation study applying RJMCMC to Bayesian Linear Regression. Section 3 introduced a Bayesian model selection example adapted from the Poisson versus Negative-Binomial problem discussed in [4]. Section 4 concludes with Discussion and Bibliography.

1 Reversible Jump MCMC

Reversible Jump MCMC (RJMCMC) is an extension of standard MCMC methods that allows exploration of models with different parameter dimensions. In conventional MCMC, the Markov chain is constructed to sample from a fixed-dimensional parameter space while satisfying key properties - namely, aperiodicity, irreducibility, and detailed balance - to ensure convergence to the target posterior distribution. RJMCMC extends these ideas by allowing the chain to "jump" between models, each with its own parameter space, while preserving these essential properties necessary for MCMC convergence. The detailed proof can be found in [4], where the author views RJMCMC as a generalization of the MH algorithm.

A typical reversible jump algorithm procedure is summarized in Algorithm 1.1

In step 7 of the algorithm, taking a bayesian formulation, we can derive $f(\theta_{j^*} | Y) \propto f(Y | \theta_{j^*}, M = j^*)f(\theta_{j^*} | M = j^*)\pi(j^*)$. Where $f(Y | \theta_{j^*}, M = j^*)$ is the likelihood function, $f(\theta_{j^*} | M = j^*)$ is the prior distribution for θ in the model j^* and $\pi(j^*)$ is our prior belief about model j^* . In addition, note that if the model proposal step (step 5) is rejected, then the move remains within the same model, and the algorithm reduces to a standard Metropolis-Hastings update. In this case, since the dimension of the parameter space does not change, the acceptance probability simplifies to the Metropolis-Hastings ratio computed solely within the current model. This ensures that when no jump is accepted, the Markov chain still explores the parameter space of the current model correctly, preserving the desired convergence properties to the target posterior distribution.

Algorithm 1.1 Reversible Jump MCMC

- 1: **Input:** Number of iterations N , initial state $(j^{(0)}, \theta_j^{(0)})$ with $\dim(\theta_j^{(0)}) = n_j^{(0)}$, model proposal probabilities $h(j, j^*)$, proposal kernel $q(u \mid \theta_j, j, j^*)$, bijective dimension-matching function g , and observed data Y .
- 2: **for** $n = 0$ to $N - 1$ **do**
- 3: **Current state:** $(j^{(n)}, \theta_j^{(n)})$
- 4: **Propose a new model** j^* with probability $h(j^* \mid j^{(n)})$
- 5: **Proposal draw:** Generate u from $q(u \mid \theta_j^{(n)}, j^*, j^{(n)})$
- 6: **Dimension matching:** Set

$$(\theta_{j^*}, u^*) = g(\theta_j^{(n)}, u), \quad \dim(\theta_j^{(n)}) + \dim(u) = \dim(\theta_{j^*}) + \dim(u^*).$$

- 7: **Acceptance probability:**

$$\alpha_{j^{(n)} \rightarrow j^*} = \min \left\{ 1, \frac{f(\theta_{j^*} \mid Y) h(j^{(n)} \mid j^*) q(u^* \mid \theta_j^{(n)}, j^*, j^{(n)}) \left| \frac{\partial(\theta_{j^*}, u^*)}{\partial(\theta_j^{(n)}, u)} \right|}{f(\theta_j^{(n)} \mid Y) h(j^* \mid j^{(n)}) q(u \mid \theta_j^{(n)}, j^{(n)}, j^*)} \right\}.$$

- 8: **Accept/reject step:**

- With probability $\alpha_{j^{(n)} \rightarrow j^*}$, set

$$(j^{(n+1)}, \theta_j^{(n+1)}) = \begin{cases} (j^*, \theta_{j^*}), & \text{with probability } \alpha_{j^{(n)} \rightarrow j^*} \\ (j^{(n)}, \theta_j^{(n)}), & \text{with probability } 1 - \alpha_{j^{(n)} \rightarrow j^*} \end{cases}$$

- 9: **end for**

- 10: **Output:** The chain of states

$$\{(j^{(n)}, \theta_j^{(n)}) : n = 0, 1, \dots, N\}.$$

2 Variable Selection using RJMCMC

2.1 Bayesian Linear Regression set up

Assume that the true model is

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{1i}^2 + \epsilon_i$$

where

$$\epsilon_i \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma^2)$$

We choose the following priors for $\beta_0, \beta_1, \beta_2$:

$$\beta_0 \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\beta^2) \quad \beta_1 \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\beta^2)$$

$$\beta_2 | m = 0 \equiv 0$$

$$\beta_2 | m = 1 \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\beta^2)$$

Here m is the model indicator and there are two models, model 0 assumes $\beta_2 = 0$ and model 1 assumes $\beta_2 \neq 0$.

2.2 Model 0 to Model 0 or Model 1 to Model 1

The move will be a standard Metropolis-Hastings move because the dimension does not change after moving. Here we use a Random Walk Metropolis Hastings kernel to propose β^* :

$$\beta^* = \beta_{curr} + \xi^*, \quad \xi^* \stackrel{\text{i.i.d.}}{\sim} N(0, \tau^2 I_n)$$

Since the proposal Markov Kernel is symmetric about 0, we simplify the acceptance probability to

$$\alpha = \min\left(1, \frac{f(\beta^*|Y)}{f(\beta_{curr}|Y)}\right)$$

Taking a Bayesian formulation, we obtain:

$$\alpha = \min\left(1, \frac{f(\beta^*)f(Y|\beta^*)}{f(\beta_{curr})f(Y|\beta_{curr})}\right)$$

Metropolis-Hastings Acceptance Ratio for Model 0 to Model 0

Model 0 is: $Y_i = \beta_0 + \beta_1 X_{1i} + \epsilon_i$, $\epsilon_i \sim N(0, \sigma^2)$. The prior distribution is: $\beta_0 \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\beta^2)$ $\beta_1 \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\beta^2)$. The proposed Markov kernel is: $\beta^* = \beta_{curr} + \xi^*$, $\xi^* \stackrel{\text{i.i.d.}}{\sim} N(0, \tau^2 I_2)$. Simplifying the likelihood and prior ratio, we obtain the Acceptance Ratio:

$$\alpha = \min\left(1, \frac{\prod_{i=1}^n \exp\left(-\frac{(Y_i - \beta_0^* - \beta_1^* X_{1i})^2}{2\sigma^2}\right) \times \exp\left(-\frac{(\beta_0^*)^2}{2\sigma_\beta^2}\right) \times \exp\left(-\frac{(\beta_1^*)^2}{2\sigma_\beta^2}\right)}{\prod_{i=1}^n \exp\left(-\frac{(Y_i - \beta_0 - \beta_1 X_{1i})^2}{2\sigma^2}\right) \times \exp\left(-\frac{\beta_0^2}{2\sigma_\beta^2}\right) \times \exp\left(-\frac{\beta_1^2}{2\sigma_\beta^2}\right)}\right)$$

Metropolis-Hastings Acceptance Ratio for Model 1 to Model 1

Similarly, the proposal Markov Kernel is:

$$\beta^* = \beta_{curr} + \xi^*, \quad \xi^* \stackrel{\text{i.i.d.}}{\sim} N(0, \tau^2 I_3)$$

Expanding the likelihood and prior terms:

$$\alpha = \min \left(1, \frac{\prod_{i=1}^n \exp\left(-\frac{(Y_i - \beta_0^* - \beta_1^* X_{1i} - \beta_2^* X_{1i}^2)^2}{2\sigma^2}\right) \times \exp\left(-\frac{(\beta_0^*)^2}{2\sigma_\beta^2}\right) \times \exp\left(-\frac{(\beta_1^*)^2}{2\sigma_\beta^2}\right) \times \exp\left(-\frac{(\beta_2^*)^2}{2\sigma_\beta^2}\right)}{\prod_{i=1}^n \exp\left(-\frac{(Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{1i}^2)^2}{2\sigma^2}\right) \times \exp\left(-\frac{\beta_0^2}{2\sigma_\beta^2}\right) \times \exp\left(-\frac{\beta_1^2}{2\sigma_\beta^2}\right) \times \exp\left(-\frac{\beta_2^2}{2\sigma_\beta^2}\right)} \right)$$

2.3 Model 0 to Model 1

Let us first consider a move from Model 0 to Model 1 (i.e. $j = 2$ and $j^* = 3$). That is, let us consider a move from $x = (2, \theta) = (2, \beta_0, \beta_1)$ to $x^* = (3, \beta_0^*, \beta_1^*, \beta_2^*)$. For simplicity, we will choose the proposed transition to keep the β_0 and β_1 for x^* as they were in x (i.e. set $\beta_0^* = \beta_0$, $\beta_1^* = \beta_1$), and to propose an independent β_2^* . We can generate $u \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_u^2)$ and set $\beta_2^* = u$. Then

$$(\beta_0^*, \beta_1^*, \beta_2^*) = g(\beta_0, \beta_1, u) = (\beta_0, \beta_1, u)$$

Since the dimension matching function g is the identity function, the Jacobian is equal to 1. Thus if we set transition probability $h(1, 1) = h(1, 2) = h(2, 1) = h(2, 2) = 1/2$, we have

$$\alpha_{0 \rightarrow 1} = \min \left(1, \frac{f(Y|\beta_0^*, \beta_1^*, \beta_2^*, M=1)f(\beta_0^*, \beta_1^*, \beta_2^*|M=1)p(M=1)}{f(Y|\beta_0, \beta_1, M=0)f(\beta_0, \beta_1|M=0)p(M=0)q(u)} \right)$$

Since $\beta_0^* = \beta_0$, $\beta_1^* = \beta_1$, the prior ratio can be simplified to:

$$\frac{f(\beta_0^*, \beta_1^*, \beta_2^*|M=1)}{f(\beta_0, \beta_1|M=0)} = \frac{1}{\sqrt{2\pi\sigma_\beta^2}} \exp\left(-\frac{(\beta_2^*)^2}{2\sigma_\beta^2}\right)$$

Since $\beta_2^* = u$ is sampled from $N(0, \sigma_u^2)$, the proposal density is:

$$q(u) = \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left(-\frac{u^2}{2\sigma_u^2}\right).$$

Suppose we have equal prior belief on Model 0 and Model 1, that is, $p(M=0) = p(M=1)$. By expanding the likelihood ratio, we obtain the final acceptance probability:

$$\alpha_{0 \rightarrow 1} = \min \left(1, \frac{\prod_{i=1}^n \exp\left(-\frac{(Y_i - \beta_0^* - \beta_1^* X_{1i} - \beta_2^* X_{1i}^2)^2}{2\sigma^2}\right) \times \frac{1}{\sqrt{2\pi\sigma_\beta^2}} \exp\left(-\frac{\beta_2^{*2}}{2\sigma_\beta^2}\right)}{\prod_{i=1}^n \exp\left(-\frac{(Y_i - \beta_0 - \beta_1 X_{1i})^2}{2\sigma^2}\right) \times \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left(-\frac{u^2}{2\sigma_u^2}\right)} \right)$$

2.4 Model 1 to Model 0

Now we consider the move from Model 1 to model 0 (i.e. $k = 3$ and $k' = 2$). That is, let us consider a move from $x = (3, \beta_0, \beta_1, \beta_2)$ to $x^* = (2, \beta_0^*, \beta_1^*)$. We will choose the reverse move to keep β_0 and β_1 . The reverse move would to be set:

$$(\beta_0^*, \beta_1^*, u^*) = g'(\beta_0, \beta_1, \beta_2) = (\beta_0, \beta_1, \beta_2)$$

Again, the dimension matching function is the identity function, our Jacobian is still 1. The acceptance probability from model 1 to model 0 will be:

$$\alpha_{1 \rightarrow 0} = \min \left(1, \frac{f(Y|\beta_0^*, \beta_1^*, M=0)f(\beta_0^*, \beta_1^*, M=0)p(M=0)q(u^*)}{f(Y|\beta_0, \beta_1, \beta_2, M=1)f(\beta_0, \beta_1, \beta_2, M=1)p(M=1)} \right).$$

Expanding and simplifying the likelihood and prior ratio, we get the final acceptance probability for the move from model 1 to model 0:

$$\alpha_{1 \rightarrow 0} = \min \left(1, \frac{\prod_{i=1}^n \exp \left(-\frac{(Y_i - \beta_0^* - \beta_1^* X_{1i})^2}{2\sigma^2} \right) \times \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp \left(-\frac{u^2}{2\sigma_u^2} \right)}{\prod_{i=1}^n \exp \left(-\frac{(Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{1i}^2)^2}{2\sigma^2} \right) \times \frac{1}{\sqrt{2\pi\sigma_\beta^2}} \exp \left(-\frac{\beta_2^2}{2\sigma_\beta^2} \right)} \right)$$

which is the reciprocal of $\alpha_{0 \rightarrow 1}$.

2.5 Simulation Results

Now let us doing a simulation test to verify our derivation. Let the ture data generating process to be:

$$Y_i = 1 + 2X_{1i} + 0.5X_{1i}^2 + \epsilon_i, \quad \epsilon_i \stackrel{\text{i.i.d.}}{\sim} N(0, 0.5)$$

Let the prior parameters be $\sigma_\beta^2 = 10$, $\sigma_u^2 = 1$, $\tau^2 = 0.05$ And let the transition probabilities between models to be $h = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$. Supoose our prior beliefs on models $p(M=0) = p(M=1) = 1/2$.

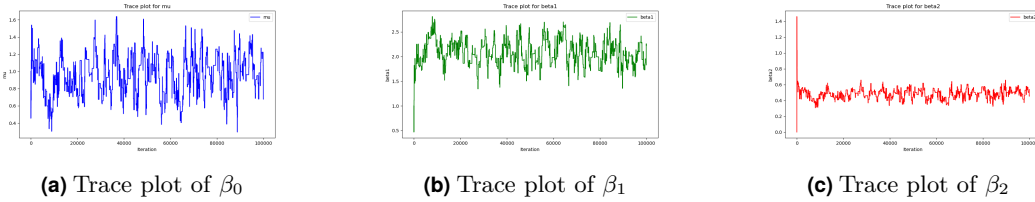


Figure 1 Trace plots for β_0 , β_1 , and β_2 .

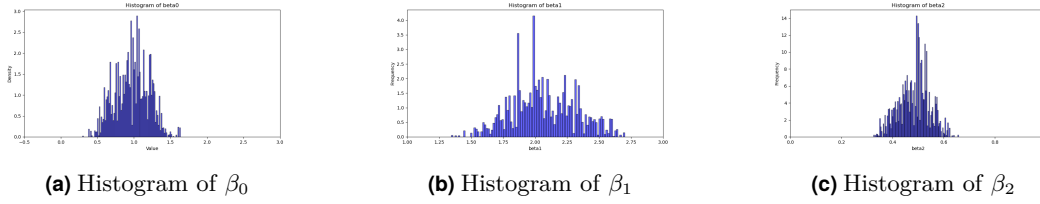


Figure 2 Histograms for β_0 , β_1 , and β_2 .

The trace plots reveal that β_2 starts at 0, but the chain quickly jumps away from 0, and thereafter the values of β_2 remain consistently nonzero. This behavior suggests that the chain escapes the initial value and begins effectively exploring the parameter space, indicating convergence toward the high-probability region of the posterior distribution. Additionally, the trace plots for β_0 and β_1 exhibit stable fluctuations around their central values, which further indicates that the MCMC has reached a steady state.

The histograms show that the estimated distributions of the β parameters are centered around their true values. The concentration of estimates in the histograms implies that the MCMC method is accurately capturing the posterior distributions and that the variability observed is in line with the expected uncertainty in the parameter estimates.

3 Bayesian Model Selection

RJMCMC also plays a crucial role in Bayesian model selection. Bayesian model selection focuses on evaluating and comparing the posterior probabilities of competing models given the observed data. The model with the highest posterior probability is often considered to be the one best supported by the data. However, rather than selecting a single model, Bayesian methods often employ Bayesian Model Averaging (BMA). BMA combines predictions from multiple models, weighting each model by its posterior probability, to produce a more robust estimate.

RJMCMC facilitates this process by allowing transitions between models with different dimensional parameter spaces during a single MCMC run. This allows researchers not only to identify the model that best explains the data, but also to account for model uncertainty through model averaging, leading to more robust and reliable inference. For example, if a parameter of interest is estimated by each model, the overall estimate can be calculated as a weighted average of the individual model estimates, where the weights are given by the posterior probabilities of the models.

3.1 Poisson versus Negative Binomial

Suppose we have count data X_1, X_2, \dots, X_n that we believe is either from a Poisson distribution or a Negative Binomial based on the dispersion of the data. For data X of length n , the likelihood under *i.i.d* Poisson distribution (Model 0) with parameter λ is

$$f(X | \lambda) = \prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}.$$

Whereas under Negative Binomial (Model 1) with parameter p and r is

$$f(X | p, r) = \prod_{i=1}^n \frac{\Gamma(x_i + r)}{\Gamma(r) x_i!} (1 - p)^r p^{x_i}.$$

Note that in order to give a model averaging result, we let $\lambda = r(1 - p)/p$, which gives us $p = r/(\lambda + r)$, so that both model has mean λ . Then the likelihood function under Negative Binomial with parameter λ and r becomes

$$f(X | \lambda, r) = f(X | \lambda, r) = \prod_{i=1}^n \frac{\Gamma(x_i + r)}{\Gamma(r) x_i!} \left(\frac{r}{\lambda + r}\right)^r \left(\frac{\lambda}{\lambda + r}\right)^{x_i}.$$

3.2 Model 0 to Model 0 or Model 1 to Model 1

Since the dimension does not change, the acceptance probability will not involve a Jacobian term. This will be a Standard Metropolis-Hastings move. And the acceptance ratio for Model 0 is:

$$\alpha = \min \left\{ 1, \frac{\pi(\lambda^* | X) q(\lambda | \lambda^*)}{\pi(\lambda | X) q(\lambda^* | \lambda)} \right\}.$$

And the acceptance ratio for Model 1 is:

$$\alpha = \min \left\{ 1, \frac{\pi(\lambda^*, r^* | X) q(\lambda, r | \lambda^*, r^*)}{\pi(\lambda, r | X) q(\lambda^*, r^* | \lambda, r)} \right\}.$$

3.3 Model 0 to Model 1

Let us consider a move from Model 0 k to Model 1 k^* . That is, let us consider a move from $x = (1, \lambda)$ to $x^* = (2, \lambda^*, r^*)$. We will choose a proposed transition to keep the λ as it was in x , and propose $r^* \sim \exp(1)$. Since the inverse c.d.f. for r^* is $F^{-1}(r') = -\ln(1 - u)$ where $u \sim \text{unif}(0, 1)$ and $u \stackrel{d}{=} 1 - u$. Using the Inverse Transformation Method in Chapter 2, we can generate $r' = -\ln(u)$, $u \sim \text{unif}(0, 1)$. Now, the transition move becomes

$$(\lambda^*, r^*) = g(\lambda, u) = (\lambda, -\ln(u))$$

The Jacobian matrix of the transformation is given by

$$J = \begin{pmatrix} \frac{\partial \lambda^*}{\partial \lambda} & \frac{\partial \lambda^*}{\partial u} \\ \frac{\partial r^*}{\partial \lambda} & \frac{\partial r^*}{\partial u} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{u} \end{pmatrix}.$$

Thus, the determinant of the Jacobian is

$$|J| = \left| 1 \cdot \left(-\frac{1}{u}\right) - 0 \cdot 0 \right| = \frac{1}{u}.$$

The acceptance probability from $x = (1, \lambda)$ to $x^* = (2, \lambda^*, r^*)$ is:

$$\alpha((1, \lambda), (2, \lambda^*, r^*)) = \min \left\{ 1, \frac{\pi(2, \lambda^*, r^* | X) h(\text{Model} = 0 | \text{Model} = 1)}{\pi(1, \lambda | X) g(u) h(\text{Model} = 1 | \text{Model} = 0)} \times \left| \frac{\partial(\lambda^*, r^*)}{\partial(\lambda, u)} \right| \right\}.$$

where h is the transition probability.

3.4 Model 1 to Model 0

Let us now consider the move from Model 1 to Model 0. That is, let us consider a move from $x = (2, \lambda, r)$ to $x^* = (1, \lambda^*)$. Since $r^* = -\ln(u)$ in the forward move, the reverse move should be $u^* = \exp(-r)$. Now the reverse move becomes

$$(\lambda^*, u^*) = g^{-1}(\lambda, r) = (\lambda, \exp(-r))$$

Compute the Jacobian:

$$J = \begin{pmatrix} \frac{\partial \lambda^*}{\partial \lambda} & \frac{\partial \lambda^*}{\partial r} \\ \frac{\partial u^*}{\partial \lambda} & \frac{\partial u^*}{\partial r} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -e^{-r} \end{pmatrix}, \quad |J| = e^{-r}.$$

The acceptance probability from $x = (2, \lambda, r)$ to $x^* = (1, \lambda^*)$ is:

$$\alpha((2, \lambda, r), (1, \lambda^*)) = \min \left\{ 1, \frac{\pi(1, \lambda^* | X) h(\text{Model} = 1 | \text{Model} = 0) g(u^*)}{\pi(2, \lambda, r | X) h(\text{Model} = 0 | \text{Model} = 1)} \times \left| \frac{\partial(\lambda^*, u^*)}{\partial(\lambda, r)} \right| \right\}.$$

where h is the transition probability.

3.5 Posterior Model Probabilities

The posterior probability of each model can then be approximated from the RJMCMC output. Let N be the total number of iterations in the RJMCMC chain after burn-in, and let N_0 and N_1 be the number of iterations spent in Model 0 and Model 1, where $N = N_0 + N_1$. The posterior probability of model 0 and model 1 is approximated by:

$$P(\text{Model 0} | X) \approx \frac{N_0}{N}, \quad P(\text{Model 1} | X) \approx \frac{N_1}{N}.$$

These probabilities reflect how often the RJMCMC algorithm visits each model, which is proportional to their posterior support given the data X . If the data exhibit overdispersion (variance greater than the mean), the Negative Binomial model may receive higher posterior probability, while a Poisson model may dominate for equidispersed data.

3.6 Averaged Posterior Mean

Suppose our goal is to estimate the posterior mean of λ , the expected count under either of the two models. For each iteration of the RJMCMC chain, we record the value of λ for model 0 and model 1. (for Model 0) or $\lambda = r(1 - p)/p$ (for Model 1, though expressed as λ directly in the reparameterized likelihood).

Let $\lambda^{(t)}$ denote the value of λ at iteration t of the chain, where $t = 1, 2, \dots, N$. The Bayesian model-averaged posterior mean of λ is then:

$$E[\lambda | X] = \sum_{t=1}^N \frac{1}{N} \lambda^{(t)}.$$

Which can also be expressed in terms of the individual model contributions:

$$E[\lambda | X] = P(\text{Model 0} | X) \cdot E[\lambda | X, \text{Model 0}] + P(\text{Model 1} | X) \cdot E[\lambda | X, \text{Model 1}],$$

Where:

- $E[\lambda \mid X, \text{Model } 0]$ is the posterior mean of λ under the Poisson model, computed as the average of λ values sampled when the chain is in Model 0:

$$E[\lambda \mid X, \text{Model } 0] = \frac{1}{N_0} \sum_{t:k^{(t)}=0} \lambda^{(t)},$$

- $E[\lambda \mid X, \text{Model } 1]$ is the posterior mean of λ under the Negative Binomial model, computed as the average of λ values sampled when the chain is in Model 1:

$$E[\lambda \mid X, \text{Model } 1] = \frac{1}{N_1} \sum_{t:k^{(t)}=1} \lambda^{(t)}.$$

Here, $k^{(t)}$ denotes the model index at iteration t (0 for Poisson, 1 for Negative Binomial).

4 Discussion and Bibliography

Reversible Jump Markov Chain Monte Carlo is an extension of MCMC designed to handle problems involving variable-dimensional parameter spaces, such as Bayesian model selection and trans-dimensional inference.

Although RJMCMC provides a robust framework for Bayesian model selection, implementation has many challenges. Green[4] pointed out that constructing across-model proposals that lead to efficient parameter space exploration while retaining efficient convergence to the stationary distribution might be challenging to achieve. Designing proposal distributions that effectively transition between models of differing dimensions—such as from the Poisson (Model 0) to the Negative Binomial (Model 1)—requires careful balancing. If the proposals are too narrow, the algorithm may fail to explore the full range of possible models and parameters adequately. This leads to poor mixing and biased estimates of posterior model probabilities. Conversely, if the proposals are too broad, the acceptance rate of trans-dimensional moves may drop significantly, slowing convergence and rendering the algorithm computationally inefficient. Green then emphasized that the key idea in designing efficient across-model proposals is to generate new states that maintain posterior support comparable to that of the current state. By doing so, the probability of accepting both the forward move and its reverse is increased.

Recent research has focused on improving RJMCMC proposal distributions to enhance model transition efficiency and acceptance rates. Ma et al. (2016)[5] present a framework for designing efficient and irreversible MCMC samplers, emphasizing the importance of structuring transition kernels to enhance exploration across models. They discuss how irreversible jump processes can lead to more efficient model transitions compared to conventional reversible proposals. Hastie and Green (2011)[4] highlight the potential of adaptive proposals that learn from previous samples to significantly improve the efficiency of model transitions. Norets (2021)[6] discusses auxiliary priors that optimize proposal distributions in RJMCMC to ensure efficient transitions between variable dimensional models. The study provides theoretical insights into the design of diagonal transition probabilities that improve acceptance rates. Fan et al. (2009)

[1] explore the use of automated likelihood-based proposals to reduce the need for manual tuning of transitions between models. They evaluate the performance of continuous proposal sampling (CPS) strategies that improve the efficiency of RJMCMC. Finally, Gagnon (2021)[2] proposes informed reversible jump algorithms that adaptively adjust proposals based on previous samples, significantly improving acceptance rates. This approach has been shown to significantly improve acceptance rates in complex hierarchical Bayesian models.

References

- [1] Y. Fan, G. W. Peters, and S. A. Sisson. Automating and Evaluating Reversible Jump MCMC Proposal Distributions. *Statistics and Computing*, 19, 2008.
- [2] P. Gagnon. Informed Reversible Jump Algorithms. *Electronic Journal of Statistics*, 15, 2021.
- [3] P. J. Green. Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika*, 82, 1995.
- [4] D. I. Hastie and P. J. Green. Model Choice Using Reversible Jump Markov Chain Monte Carlo. *Statistica Neerlandica*, 66, 2012.
- [5] Y.-A. Ma, E. B. Fox, T. Chen, and L. Wu. Irreversible Samplers from Jump and Continuous Markov Processes. *Statistics and Computing*, 29, 2018.
- [6] A. Norets. OPTIMAL AUXILIARY PRIORS and REVERSIBLE JUMP PROPOSALS for a CLASS of VARIABLE DIMENSION MODELS. *Econometric Theory*, 37, 2020.