

# 基于切片 Gibbs 抽样算法的空间误差模型的贝叶斯参数估计

李澳归<sup>1</sup>, 赵远英<sup>2,3\*</sup>

(1. 贵州大学数学与统计学院, 贵州 贵阳 550025; 2. 贵阳学院理学院, 贵州 贵阳 550005; 3. 贵州交通职业大学轨道交通工程系, 贵州 贵阳 551400)

**摘要:**提出一种切片 Gibbs 抽样的马尔可夫链蒙特卡罗 (Markov chain Monte Carlo, MCMC) 算法来计算空间误差模型未知参数的联合贝叶斯估计, 通过 2 个模拟仿真说明提出的贝叶斯估计方法的有效性与切片 Gibbs 抽样算法的优势, 实证分析说明模型和提出的贝叶斯估计方法的有效性。

**关键词:**贝叶斯估计; Gibbs 抽样; 切片抽样; 空间误差模型

**中图分类号:** O212.8 **文献标志码:** A

**引用格式:** 李澳归, 赵远英. 基于切片 Gibbs 抽样算法的空间误差模型的贝叶斯参数估计[J]. 山东大学学报(理学版), 2025, 60(3): 116-126.

## Bayesian parametric estimations for spatial error models based on slice-Gibbs sampling

LI Aogui<sup>1</sup>, ZHAO Yuanying<sup>2,3\*</sup>

(1. College of Mathematics and Statistic, Guizhou University, Guiyang 550025, Guizhou, China; 2. College of Science, Guiyang University, Guiyang 550005, Guizhou, China; 3. Department of Railway Engineering, Guizhou Communications Polytechnic University, Guiyang 551400, Guizhou, China)

**Abstract:** A Markov chain Monte Carlo (MCMC) technique called the slice-Gibbs sampling algorithm is proposed to calculate joints Bayesian estimations of unknown parameters for spatial error models, the proposed algorithm and Bayesian approach are illustrated by two simulation studies, the model and methodology are demonstrated by a real example.

**Key words:** Bayesian estimation; Gibbs sampling; slice sampling; spatial error model

## 0 引言

经典的计量经济学重点分析时间序列数据或面板数据的统计计量方法与应用, 而忽视地理空间近邻等因素带来的数据空间相关性与异质性, 从而可能导致参数估计偏差过大或者假设检验结果失效。将空间因素引入计量经济学的统计分析框架之后, 空间经济学成为经济学研究的主流的一部分。空间计量经济模型<sup>[1-5]</sup>是分析处理空间结构数据的有效工具。特别地, 空间误差模型引起了众多研究者的关注。例如: Guo 等<sup>[6]</sup>探讨了空间误差模型的稳健性检验; Chiara 等<sup>[7]</sup>获得了空间误差模型的递推估计; 李坤明等<sup>[8]</sup>使用非参数方法研究了空间误差模型的截面最小二乘估计; Vural 等<sup>[9]</sup>提出了空间误差模型的稳健估计方法; Dai 等<sup>[10]</sup>讨论了变系数空间误差模型的分位数回归方法等等。在贝叶斯统计框架下, 切片抽样方法是一种比较流行的抽样方法。陈晓弩<sup>[11]</sup>指出基于切片抽样的马尔可夫链蒙特卡罗 (Markov chain Monte Carlo, MCMC) 算法可以降低计算成本; Sengupta 等<sup>[12]</sup>的研究指出, 切片抽样相较于自适应 MCMC 抽样产生独立样本

收稿日期: 2023-11-08; 网络出版时间: 2024-11-13 14:33:10

基金项目: 国家自然科学基金资助项目 (12161014); 贵州省省级科技计划项目 (黔科合支撑 [2023] 一般 139) 2018 年度贵州省高层次创新型人才项目阶段性成果

第一作者: 李澳归 (1999—), 男, 硕士研究生, 研究方向为概率论与数理统计. E-mail: lags123@163.com

\* 通信作者: 赵远英 (1981—), 男, 教授, 硕士生导师, 博士, 研究方向为应用统计. E-mail: zhaoyuanying\_@126.com

的数量更多;文献[13-15]通过研究发现:切片抽样能够避免 Metropolis-Hastings(MH)算法<sup>[16-17]</sup>中的回顾性调参,提高抽样的效率。关于切片抽样方法的更多讨论,请参考文献[18-19]。

虽然在空间误差模型及其统计推断方法已经取得了丰硕的研究成果;然而,鲜有使用切片抽样方法研究空间误差模型中贝叶斯参数估计问题,因此,本文的主要目标就是在贝叶斯统计和空间误差模型的框架背景下,讨论模型未知参数的估计问题。

## 1 空间误差模型及其假设

本文研究如下的空间误差模型:

$$\begin{cases} \mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \boldsymbol{\mu}, \\ \boldsymbol{\mu} = \boldsymbol{\rho}\boldsymbol{\omega}\boldsymbol{\mu} + \boldsymbol{\varepsilon}, \\ \boldsymbol{\varepsilon} \sim N(\mathbf{0}_n, \sigma^2\mathbf{I}_n), \end{cases} \quad (1)$$

其中:  $\mathbf{y}$  为  $n \times 1$  型的响应变量向量;  $\mathbf{x}$  为协变量构成的  $n \times p$  型设计矩阵;  $\boldsymbol{\beta}$  为  $p \times 1$  型回归系数;  $\boldsymbol{\mu}$  为  $n \times 1$  型的随机误差向量。本文考虑随机误差向量  $\boldsymbol{\mu}$  存在空间相关结构  $\boldsymbol{\mu} = \boldsymbol{\rho}\boldsymbol{\omega}\boldsymbol{\mu} + \boldsymbol{\varepsilon}$ , 其中  $\boldsymbol{\rho}$  是空间相关参数, 且  $-1 < \boldsymbol{\rho} < 1$ ;  $n \times n$  型空间权重矩阵  $\boldsymbol{\omega}$  必须满足:  $\forall i = 1, 2, \dots, n$ ,

$$(i) \omega_{ii} = 0; \quad (ii) \sum_{j=1}^n \omega_{ij} = 1. \quad (2)$$

在模型(1)中,  $\boldsymbol{\varepsilon}$  是最终的  $n \times 1$  型随机误差向量,  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}_n, \sigma^2\mathbf{I}_n)$  表示  $\boldsymbol{\varepsilon}$  服从均值为  $\mathbf{0}_n$ 、协方差矩阵为  $\sigma^2\mathbf{I}_n$  的  $n$  元正态分布, 其中  $\mathbf{0}_n$  表示  $n \times 1$  型的零向量,  $\mathbf{I}_n$  表示  $n$  阶单位矩阵。

## 2 贝叶斯分析

### 2.1 似然函数与未知参数的先验分布假设

用  $\mathbf{A}$  表示  $\mathbf{I}_n - \boldsymbol{\rho}\boldsymbol{\omega}$ , 由模型(1)容易得到  $\mathbf{y}$  的均值和协方差矩阵如下:

$$\begin{cases} E(\mathbf{y}) = \mathbf{x}\boldsymbol{\beta}, \\ \text{Var}(\mathbf{y}) = \sigma^2\mathbf{A}^{-1}(\mathbf{A}^{-1})^T, \end{cases}$$

从而有

$$\mathbf{y} \sim N(\mathbf{x}\boldsymbol{\beta}, \sigma^2\mathbf{A}^{-1}(\mathbf{A}^{-1})^T). \quad (3)$$

记  $\mathbf{D} = \{\mathbf{y}, \mathbf{x}\}$ ,  $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \sigma^2, \boldsymbol{\rho}\}$ , 故数据  $\mathbf{D}$  的似然函数为

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\rho}) = (2\pi\sigma^2)^{-\frac{n}{2}} |\mathbf{A}| \exp\left\{-\frac{1}{2\sigma^2} [\mathbf{A}(\mathbf{y} - \mathbf{x}\boldsymbol{\beta})]^T [\mathbf{A}(\mathbf{y} - \mathbf{x}\boldsymbol{\beta})]\right\}, \quad (4)$$

其中  $|\mathbf{A}|$  表示矩阵  $\mathbf{A}$  的行列式。

本文假设  $\boldsymbol{\theta}$  的先验分布为

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\rho}) = \pi(\boldsymbol{\beta})\pi(\sigma^2)\pi(\boldsymbol{\rho}), \quad (5)$$

且

$$\pi(\boldsymbol{\beta}) \sim N(\boldsymbol{\beta}_0, \mathbf{H}_\beta), \pi(\sigma^2) \sim \text{IG}(a, b), \pi(\boldsymbol{\rho}) \sim U(\rho_{\min}, \rho_{\max}), \quad (6)$$

其中  $N(\boldsymbol{\beta}_0, \mathbf{H}_\beta)$  表示期望为  $\boldsymbol{\beta}_0$ 、协方差矩阵为  $\mathbf{H}_\beta$  的正态分布,  $\text{IG}(a, b)$  代表参数为  $a$  与  $b$  的逆  $\Gamma$  分布(且  $a > 0, b > 0$ ),  $U(\rho_{\min}, \rho_{\max})$  表示区间  $(\rho_{\min}, \rho_{\max})$  上的均匀分布,  $\boldsymbol{\beta}_0, \mathbf{H}_\beta, a, b, \rho_{\min}, \rho_{\max}$  是给定的超参数。式(6)中参数  $\boldsymbol{\beta}, \sigma^2$  和  $\boldsymbol{\rho}$  相应的概率密度函数分别为

$$\pi(\boldsymbol{\beta}) = (2\pi)^{-\frac{p}{2}} |\mathbf{H}_\beta|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)^T \mathbf{H}_\beta^{-1}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)\right\}, \quad (7)$$

$$\pi(\sigma^2) = \frac{b^a}{\Gamma(a)} (\sigma^2)^{-a-1} \exp\left\{-\frac{b}{\sigma^2}\right\}, \quad (8)$$

$$\pi(\boldsymbol{\rho}) = \frac{1}{\rho_{\max} - \rho_{\min}} I(\rho_{\min} < \boldsymbol{\rho} < \rho_{\max}), \quad (9)$$

其中  $I(\cdot)$  表示示性函数。

## 2.2 未知参数的后验分布与 Gibbs 抽样

根据式(4)、(6),数据  $D$  与参数  $\theta$  的联合后验分布为

$$p(D, \theta) \propto L(\theta) \pi(\theta) \propto (\sigma^2)^{-a-\frac{n}{2}-1} |A| \exp\left\{-\frac{1}{2\sigma^2} [A(y-x\beta)]^T [A(y-x\beta)]\right\} \\ -\frac{b}{\sigma^2} - \frac{1}{2} (\beta - \beta_0)^T H_\beta^{-1} (\beta - \beta_0) \} I\{\rho_{\min} < \rho < \rho_{\max}\}. \quad (10)$$

下面分别导出未知参数  $\beta$ 、 $\sigma^2$ 、 $\rho$  的后验分布,并给出获得未知参数  $\theta = \{\beta, \sigma^2, \rho\}$  贝叶斯估计的 Gibbs 抽样<sup>[20]</sup>算法。

**步骤 1** 设置参数  $\beta$ 、 $\sigma^2$  与  $\rho$  的初值为  $\beta^{(0)}$ 、 $\sigma^{2(0)}$  与  $\rho^{(0)}$ ,并令  $m=0$ 。

**步骤 2** 从条件分布  $p(\beta | y, x, \sigma^{2(m)}, \rho^{(m)})$  中抽取随机样本  $\beta^{(m+1)}$ ,

$$p(\beta | y, x, \sigma^2, \rho) \sim N(\beta^*, \Sigma_\beta^*), \quad (11)$$

其中  $\beta^* = \Sigma_\beta \left[ \frac{1}{\sigma^2} (Ax)^T Ay + H_\beta^{-1} \beta_0 \right]$ ,  $\Sigma_\beta = \left\{ \frac{1}{\sigma^2} (Ax)^T (Ax) + H_\beta^{-1} \right\}^{-1}$ 。

**步骤 3** 从条件分布  $p(\sigma^2 | y, x, \beta^{(m+1)}, \rho^{(m)})$  中随机抽取样本  $\sigma^{2(m+1)}$ ,

$$p(\sigma^2 | y, x, \beta, \rho) \sim \text{IG}(a^*, b^*), \quad (12)$$

其中  $a^* = a + \frac{n}{2}$ ,  $b^* = a + \frac{1}{2} [A(y-x\beta)]^T [A(y-x\beta)]$ 。

**步骤 4** 从条件分布  $p(\rho | y, x, \beta^{(m+1)}, \sigma^{2(m+1)})$  中抽取随机样本  $\rho^{(m+1)}$ ,

$$p(\rho | y, x, \beta, \sigma^2) \propto |A| \exp\left\{-\frac{1}{2\sigma^2} [A(y-x\beta)]^T [A(y-x\beta)]\right\} I\{\rho_{\min} < \rho < \rho_{\max}\}. \quad (13)$$

**步骤 5** 令  $m=m+1$ ,重复步骤(2)–(4),并在算法收敛后收集后验样本序列  $\{(\beta^{(t)}, \sigma^{2(t)}, \rho^{(t)}): t=1, 2, \dots, T\}$ 。

注 1 从条件分布(11)、(12)中可以直接抽样,获取未知参数  $\beta$  与  $\sigma^2$  后验分布的样本,但条件分布(13)是非正常的标准分布,不能直接抽样。

## 2.3 参数 $\rho$ 的抽样方法

根据 Lesage 等<sup>[1]</sup>,可以借助 MH 算法<sup>[16-17]</sup>从条件分布(13)中随机抽样。具体算法如下:假设  $\rho$  在当前的状态值为  $\rho^{(m)}$ ,从建议分布  $N(\rho^{(m)}, c^2)$  中产生一个待定的随机样本  $\rho^*$ ,同时,从均匀分布  $U(0, 1)$  中产生一个独立的随机数  $u$ ,若  $u \leq \min\left\{1, \frac{p(\rho^* | y, x, \beta, \sigma^2)}{p(\rho^{(m)} | y, x, \beta, \sigma^2)}\right\}$ ,则令  $\rho^{(m+1)} = \rho^*$ ;否则令  $\rho^{(m+1)} = \rho^{(m)}$ 。根据 Lesage 等<sup>[1]</sup>的建议,选择调优参数  $c^2$  使得接受待定转移样本  $\rho^*$  的接受概率为 40%~60%。称上述基于 MH 算法的 Gibbs 抽样为 MH-Gibbs 抽样算法。

基于 Neal<sup>[18]</sup>和 Liu<sup>[19]</sup>的研究,本文提出采用切片(slice)抽样算法<sup>[18]</sup>获得条件分布(13)的随机样本,因此本文称基于切片抽样算法的 Gibbs 抽样为切片 Gibbs 抽样算法。从条件分布(13)中获取随机样本的切片抽样算法如下。

**算法** 空间误差模型中参数  $\rho$  的抽样算法

**输入:** 参数  $\rho$  当前的迭代值  $\rho^{(m)}$ ,正比于  $\rho$  的后验条件分布  $p(\rho | y, x, \beta^{(m+1)}, \sigma^{2(m+1)})$  的函数  $f(\rho | y, x, \beta^{(m+1)}, \sigma^{2(m+1)})$  以及抽样区间  $[L, R] = [\rho_{\min}, \rho_{\max}]$ 。

**Step 1:** 从均匀分布  $U(0, f(\rho | y, x, \beta^{(m+1)}, \sigma^{2(m+1)}))$  中随机抽样  $\eta$ ,并令  $z = \log(\eta)$ ;

**while** 迭代次数为  $m+1$  时:

**Step 2:** 从均匀分布  $U(L, R)$  中随机抽样  $\rho^*$ ;

**Step 3:**

如果  $f(\rho^* | y, x, \beta^{(m+1)}, \sigma^{2(m+1)}) < z$  则

如果  $\rho^* < \rho^{(m)}$  令  $L = \rho^*$ ;

否则令  $R = \rho^*$ ;

回到 Step 2;

否则 令  $\rho^{(m+1)} = \rho^*$ ;

End while

输出: 新的迭代值  $\rho^{(m+1)}$ 。

注2 (1) 对于 MH-Gibbs 抽样算法来说, MH 算法的调参是一个不可避免的问题。如果调优参数  $c^2$  太小会导致效率低下的随机游走; 而调优参数  $c^2$  太大将会导致频繁的拒绝。换言之, MH-Gibbs 抽样算法需要人工调参。更为糟糕的是在进行贝叶斯统计推断时, 不同的参数设置、不同的样本容量、不同的先验分布等条件下的数据分析都需要重新人工调参。(2) 对切片 Gibbs 抽样算法而言, 切片抽样可以实现自适应地调参, 且能保证算法更快地收敛为目标分布, 其有效性要比 MH 算法高得多。

### 2.4 贝叶斯估计与标准差

基于上述 Gibbs 抽样算法以及算法收敛后收集的后验样本序列, 获得参数  $\beta, \sigma^2$  与  $\rho$  的联合贝叶斯估计为

$$\hat{\beta} = \frac{1}{T} \sum_{t=1}^T \beta^{(t)}, \quad \hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T \sigma^{2(t)}, \quad \hat{\rho} = \frac{1}{T} \sum_{t=1}^T \rho^{(t)}, \tag{14}$$

同时获得参数  $\beta, \sigma^2$  与  $\rho$  的后验协方差阵的贝叶斯估计, 例如  $\beta$  的后验协方差矩阵  $\text{Var}(\beta)$  的估计为

$$\widehat{\text{Var}}(\beta) = \frac{1}{T-1} \sum_{t=1}^T (\beta^{(t)} - \hat{\beta})(\beta^{(t)} - \hat{\beta})^T,$$

其中  $\widehat{\text{Var}}(\beta)$  的对角线元素的算术平方根即为  $\hat{\beta}$  的标准差。

根据 Geyer<sup>[21]</sup> 的讨论, 后验样本序列  $\{(\beta^{(t)}, \sigma^{2(t)}, \rho^{(t)}) : t=1, 2, \dots, T\}$  的样本均值与样本协方差阵是模型参数  $\beta, \sigma^2$  与  $\rho$  的后验均值与后验协方差阵的相合估计。

## 3 模拟仿真研究

使用 2 个蒙特卡罗模拟仿真研究来展示文章中所提出来的贝叶斯估计方法的有限样本性能。

### 3.1 贝叶斯估计方法的性能

在第 1 个模拟仿真研究中  $\forall i=1, 2, \dots, n$ , 假设模拟数据  $y_i$  由模型 (1) 产生, 其中协变量  $x_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4})^T$  的每一分量都独立地服从一元标准正态分布  $N(0, 1)$ 。模型中回归系数的真值被设置为

$$\beta = (\beta_1, \beta_2, \beta_3, \beta_4)^T = \left(1, -\frac{\sqrt{2}}{2}, \frac{e}{3}, -1\right)^T, \text{ 其中 } e \text{ 表示自然常数。与 Wang}^{[22]} \text{ 相同, 模型中的空间权重矩阵 } \omega$$

被设置为: (i)  $\omega_{1,2} = \omega_{n,n-1} = 1$ ; (ii)  $\forall i=2, 3, \dots, n-1$ , 设置  $\omega$  的第  $i$  行中第  $i-1$  列与  $i+1$  列的元素均为  $\frac{1}{2}$ ,

即  $\omega_{i,i-1} = \omega_{i,i+1} = \frac{1}{2}$ , (iii) 除上述 2 种情形外,  $\omega$  的其余元素为 0。假设如下 3 种空间相关参数真值的情形:

(1)  $\rho=0.1$ ; (2)  $\rho=-0.5$ ; (3)  $\rho=0.9$ 。这 3 种情形分别代表较弱正相关关系、一般强度负相关关系和较强正相关关系。此外考虑如下 2 种  $\sigma^2$  真值的设置: (1)  $\sigma^2=0.25$ ; (2)  $\sigma^2=1$ 。在模拟研究过程中, 考虑样本容量分别为  $n=40$  和  $n=300$  情形对参数估计性能的影响, 同时考虑以下 2 种不同的先验信息(超参数)对贝叶斯估计性能的敏感性:

**类型 1** (良好的先验)  $\beta_0$  取为  $\beta$  的真值,  $H_\beta = 0.01I_4$ ,  $a=b=0.001$ ,  $\rho_{\min} = -1$ ,  $\rho_{\max} = 1$ , 其中  $I_k$  表示  $k$  阶单位矩阵。

**类型 2** (无信息先验)  $\beta_0 = \mathbf{0}_4$ ,  $H_\beta = 1000I_4$ ,  $a=b=0.001$ ,  $\rho_{\min} = -1$ ,  $\rho_{\max} = 1$ , 其中  $\mathbf{0}_m$  表示  $m \times 1$  型零向量。

基于上述设置随机产生 200 次重复抽样数据, 使用文章提出的切片 Gibbs 抽样算法进行拟合, 表 1—2 列出了相应的数值计算结果。对任意的  $j=1, 2, \dots, 200$ , 记  $\hat{\theta}_j$  为基于第  $j$  次重复抽样数据得到的  $\theta$  的参数估计,

$\theta^{\text{true}}$  表示  $\theta$  的真值,  $\text{MEAN} = \frac{1}{200} \sum_{j=1}^{200} \hat{\theta}_j$ ,  $\text{Bias} = \text{abs}(\theta^{\text{true}} - \text{MEAN})$ , (其中  $\text{abs}$  表示绝对值),  $\text{STD} =$

$\sqrt{\frac{1}{200-1} \sum_{j=1}^{200} (\hat{\theta}_j - \text{MEAN})^2}$ ,  $\text{RMSE} = \sqrt{\frac{1}{200-1} \sum_{j=1}^{200} (\hat{\theta}_j - \theta^{\text{true}})^2}$ , Median, 5%th, 95%th 分别表示基于 200 次重复抽样数据得到的估计的中位数、5%分位数和 95%分位数。综合表 1、2 的结果发现:1) 3 类不同空间参数设置下的参数估计的 Bias 都很小(不超过 0.09),说明文章提出的贝叶斯估计方法在较弱相关、一般强度相关和较强相关情形下的估计精度都很高;2) 基于切片 Gibbs 抽样方法在不同先验信息条件下的估计性能均较为稳健,且在先验类型 1 情况下的贝叶斯估计略优于先验类型 2 情况下的贝叶斯估计;3) 随着样本容量  $n$  的增大,贝叶斯估计方法的表现越来越稳定。

表 1 第一个模拟仿真研究的数值结果(基于  $\sigma^2 = 0.25$ )  
Table 1 The numerical result of the simulation 1 (based on  $\sigma^2 = 0.25$ )

先验信息	$\rho$	参数	$n=40$					$n=300$						
			Bias	Median	5%	95%	STD	RMSE	Bias	Median	5%	95%	STD	RMSE
0.1		$\beta_1$	0.005 9	1.002 4	0.923 4	1.086 5	0.049 9	0.050 1	0.001 7	0.998 8	0.949 7	1.040 1	0.027 3	0.027 3
		$\beta_2$	0.004 2	-0.705 1	-0.808 5	-0.624 5	0.053 2	0.053 3	0.002 4	-0.710 0	-0.753 0	-0.660 5	0.027 0	0.027 1
		$\beta_3$	0.001 0	0.905 8	0.831 7	0.981 4	0.048 8	0.048 7	0.001 7	0.901 8	0.855 6	0.952 7	0.028 8	0.028 8
		$\beta_4$	0.004 6	-1.008 1	-1.088 1	-0.923 1	0.052 4	0.052 5	0.005 1	-1.006 6	-1.051 0	-0.964 9	0.027 1	0.027 5
		$\sigma^2$	0.011 1	0.257 0	0.106 1	0.368 6	0.063 6	0.064 4	0.000 8	0.250 3	0.219 8	0.283 3	0.019 1	0.019 1
		$\rho$	0.004 7	0.105 2	-0.102 5	0.338 8	0.144 6	0.144 3	0.002 3	0.107 8	0.007 8	0.184 6	0.054 7	0.054 6
		SUM (RMSE)						0.413 3						0.184 4
类型 1 -0.5		$\beta_1$	0.004 1	0.994 1	0.922 6	1.0816	0.047 4	0.047 5	0.001 5	0.998 0	0.949 1	1.039 3	0.025 4	0.025 4
		$\beta_2$	0.004 3	-0.710 9	-0.804 5	-0.635 2	0.054 9	0.054 9	0.000 3	-0.706 6	-0.752 6	-0.668 7	0.025 6	0.025 5
		$\beta_3$	0.005 2	0.913 5	0.832 0	0.984 9	0.045 8	0.046 0	0.000 3	0.908 6	0.864 5	0.943 4	0.024 1	0.024 0
		$\beta_4$	0.002 4	-1.005 1	-1.080 8	-0.924 4	0.048 1	0.048 0	0.001 8	-0.999 3	-1.042 7	-0.955 8	0.025 6	0.025 6
		$\sigma^2$	0.003 1	0.245 3	0.159 7	0.365 8	0.060 4	0.060 3	0.002 5	0.251 6	0.220 3	0.286 1	0.021 3	0.021 4
		$\rho$	0.021 9	-0.481 0	-0.674 0	-0.249 2	0.129 3	0.130 8	0.001 0	-0.498 2	-0.563 7	-0.425 7	0.040 4	0.040 3
		SUM (RMSE)						0.387 5						0.162 2
0.9		$\beta_1$	0.000 7	1.001 4	0.918 4	1.078 5	0.049 5	0.049 4	0.001 6	0.999 0	0.960 3	1.036 7	0.022 7	0.022 7
		$\beta_2$	0.005 9	-0.711 8	-0.784 7	-0.640 6	0.044 4	0.044 7	0.000 7	-0.711 1	-0.743 8	-0.666 4	0.023 8	0.023 8
		$\beta_3$	0.000 3	0.907 1	0.828 2	0.975 8	0.046 1	0.046 0	0.001 2	0.908 2	0.868 1	0.943 2	0.023 1	0.023 1
		$\beta_4$	0.006 5	-0.993 2	-1.068 4	-0.913 9	0.046 2	0.046 6	0.001 2	-0.998 6	-1.035 4	-0.961 9	0.022 4	0.022 4
		$\sigma^2$	0.024 2	0.267 5	0.171 9	0.398 4	0.069 1	0.073 0	0.003 2	0.255 5	0.215 6	0.286 9	0.022 8	0.022 9
		$\rho$	0.020 7	0.887 9	0.779 5	0.916 0	0.048 7	0.052 9	0.001 7	0.900 2	0.876 1	0.915 4	0.013 2	0.013 3
		SUM (RMSE)						0.312 6						0.128 2
类型 2 0.1		$\beta_1$	0.009 3	0.997 6	0.833 6	1.116 1	0.087 1	0.087 4	0.000 6	0.996 1	0.946 8	1.048 3	0.030 9	0.030 8
		$\beta_2$	0.000 3	-0.707 2	-0.884 3	-0.560 1	0.097 5	0.097 2	0.003 0	-0.710 0	-0.758 5	-0.663 4	0.029 5	0.029 5
		$\beta_3$	0.013 5	0.917 6	0.797 4	1.048 4	0.076 3	0.077 3	0.000 3	0.906 1	0.861 4	0.948 8	0.027 3	0.027 3
		$\beta_4$	0.002 3	-1.003 6	-1.140 8	-0.854 6	0.085 0	0.084 8	0.001 0	-1.003 1	-1.050 2	-0.953 1	0.029 0	0.028 9
		$\sigma^2$	0.005 7	0.250 8	0.163 6	0.371 0	0.062 5	0.062 6	0.000 1	0.249 7	0.219 8	0.282 3	0.018 0	0.018 0
		$\rho$	0.010 5	0.113 0	-0.151 3	0.366 1	0.158 1	0.158 1	0.000 6	0.104 0	0.005 4	0.190 3	0.056 7	0.056 6
		SUM (RMSE)						0.567 4						0.191 1
类型 2 -0.5		$\beta_1$	0.001 2	0.994 2	0.869 9	1.142 7	0.089 0	0.088 8	0.001 8	0.998 8	0.948 1	1.037 3	0.025 1	0.025 1
		$\beta_2$	0.001 7	-0.712 9	-0.859 9	-0.569 7	0.089 8	0.089 6	0.003 1	-0.709 2	-0.756 2	-0.665 2	0.028 1	0.028 2
		$\beta_3$	0.013 3	0.903 0	0.736 2	1.012 7	0.083 1	0.083 9	0.001 0	0.907 6	0.857 6	0.949 8	0.026 9	0.026 9
		$\beta_4$	0.002 6	-1.005 8	-1.147 7	-0.876 8	0.078 2	0.078 1	0.000 1	-0.999 4	-1.047 1	-0.953 6	0.029 6	0.029 5
		$\sigma^2$	0.017 2	0.256 5	0.166 0	0.407 3	0.071 2	0.073 1	0.001 8	0.249 6	0.217 6	0.290 5	0.022 6	0.022 6
		$\rho$	0.049 4	-0.460 2	-0.653 3	-0.207 6	0.131 4	0.140 0	0.006 2	-0.495 5	-0.579 6	-0.404 5	0.049 3	0.049 6
		SUM (RMSE)						0.553 5						0.181 9

表 1(续)

先验信息	$\rho$	参数	$n=40$					$n=300$						
			Bias	Median	5%	95%	STD	RMSE	Bias	Median	5%	95%	STD	RMSE
类型 2	0.9	$\beta_1$	0.001 3	-1.002 2	0.866 8	1.144 9	0.079 5	0.079 3	0.000 5	1.001 8	0.956 9	1.035 6	0.024 1	0.024 0
		$\beta_2$	0.005 6	-0.698 6	-0.816 0	-0.582 4	0.074 9	0.075 0	0.000 5	-0.707 2	-0.744 4	-0.665 0	0.025 7	0.025 6
		$\beta_3$	0.008 4	0.898 3	0.766 4	1.037 3	0.082 0	0.082 3	0.001 2	0.905 9	0.866 5	0.949 6	0.025 3	0.025 2
		$\beta_4$	0.000 1	-1.001 6	-1.134 6	-0.890 7	0.074 1	0.073 9	0.003 2	-1.003 7	-1.041 8	-0.964 6	0.024 6	0.024 7
		$\sigma^2$	0.032 2	0.277 9	0.167 3	0.411 9	0.075 0	0.081 4	0.003 7	0.251 8	0.216 5	0.297 9	0.024 0	0.024 3
		$\rho$	0.027 9	0.881 9	0.774 4	0.930 1	0.048 8	0.056 1	0.001 8	0.898 3	0.872 8	0.919 4	0.014 2	0.014 3
		SUM (RMSE)							0.448 0					

表 2 第一个模拟仿真研究的数值结果(基于  $\sigma^2 = 1$ )  
Table 2 The numerical result of the simulation 1 (based on  $\sigma^2 = 1$ )

先验信息	$\rho$	参数	$n=40$					$n=300$						
			Bias	Median	5%	95%	STD	RMSE	Bias	Median	5%	95%	STD	RMSE
类型 1	-0.5	$\beta_1$	0.000 3	0.999 3	0.911 8	1.071 9	0.046 5	0.046 3	0.001 0	0.997 1	0.922 0	1.069 8	0.044 9	0.044 8
		$\beta_2$	0.004 3	-0.716 2	-0.788 0	-0.629 5	0.048 0	0.048 1	0.000 0	-0.705 6	-0.787 6	-0.639 1	0.044 3	0.044 2
		$\beta_3$	0.007 0	0.911 8	0.821 3	0.990 8	0.047 6	0.048 0	0.005 2	0.912 4	0.841 4	0.975 8	0.043 4	0.043 6
		$\beta_4$	0.002 7	-0.955 3	-1.082 0	-0.923 4	0.049 0	0.048 9	0.001 0	-1.003 5	-1.075 8	-0.930 3	0.043 3	0.043 2
		$\sigma^2$	0.036 2	1.003 4	0.660 0	1.452 8	0.252 8	0.254 7	0.005 0	0.996 2	0.853 5	1.122 7	0.077 9	0.077 9
		$\rho$	0.009 3	0.090 0	-0.184 7	0.319 3	0.154 6	0.154 5	0.005 4	0.096 4	-0.000 5	0.178 1	0.057 3	0.057 4
		SUM (RMSE)							0.600 5					
类型 1	0.9	$\beta_1$	0.001 3	0.998 8	0.927 3	1.076 9	0.045 3	0.045 2	0.000 8	1.000 6	0.928 0	1.074 4	0.043 9	0.043 8
		$\beta_2$	0.000 5	-0.714 0	-0.775 3	-0.625 1	0.046 9	0.046 8	0.004 0	-0.702 3	-0.779 6	-0.635 0	0.043 1	0.043 2
		$\beta_3$	0.001 2	0.907 8	0.832 3	0.988 9	0.049 8	0.049 6	0.001 6	0.905 1	0.836 9	0.974 2	0.043 0	0.042 9
		$\beta_4$	0.002 5	-1.000 5	-1.094 4	-0.918 1	0.051 3	0.051 2	0.001 8	-0.998 8	-1.059 6	-0.933 4	0.040 9	0.040 8
		$\sigma^2$	0.041 5	1.010 4	0.639 0	0.531 3	0.266 9	0.269 4	0.005 9	1.015 3	0.844 8	1.1595	0.091 8	0.091 7
		$\rho$	0.040 7	-0.481 4	-0.650 6	-0.193 6	0.136 5	0.142 1	0.000 5	-0.500 7	-0.581 0	-0.434 9	0.044 5	0.044 4
		SUM (RMSE)							0.604 3					
类型 2	0.1	$\beta_1$	0.015 0	0.989 7	0.729 8	1.246 5	0.166 0	0.166 2	0.001 1	0.996 5	0.903 8	1.088 7	0.057 6	0.057 4
		$\beta_2$	0.005 5	-0.706 3	-1.010 0	-0.457 6	0.163 4	0.163 1	0.003 4	-0.712 6	-0.794 3	-0.610 0	0.055 5	0.055 5
		$\beta_3$	0.023 6	0.928 6	0.669 7	1.186 4	0.161 7	0.163 0	0.000 7	0.908 7	0.803 2	0.999 8	0.057 0	0.056 9
		$\beta_4$	0.021 5	-1.016 9	-1.301 6	-0.772 0	0.169 4	0.170 4	0.011 2	-1.006 2	-1.121 7	-0.924 6	0.060 0	0.060 9
		$\sigma^2$	0.040 6	1.031 4	0.645 0	1.460 1	0.259 5	0.262 1	0.002 6	1.001 3	0.871 3	1.144 9	0.082 7	0.082 6
		$\rho$	0.010 6	0.078 2	-0.176 3	0.367 8	0.160 2	0.160 2	0.003 5	0.103 6	0.003 7	0.201 0	0.059 7	0.059 7
		SUM (RMSE)							1.085 0					

表2(续)

先验信息	$\rho$	参数	$n=40$					$n=300$						
			Bias	Median	5%	95%	STD	RMSE	Bias	Median	5%	95%	STD	RMSE
-0.5		$\beta_1$	0.000 9	0.997 4	0.732 0	1.262 6	0.165 0	0.164 6	0.001 3	1.002 1	0.895 8	1.087 8	0.056 7	0.056 6
		$\beta_2$	0.004 0	-0.725 4	-0.968 3	-0.407 4	0.159 9	0.159 5	0.000 7	-0.709 3	-0.808 4	-0.613 1	0.058 8	0.058 7
		$\beta_3$	0.009 8	0.915 4	0.654 0	1.202 3	0.169 0	0.168 8	0.004 0	0.919 1	0.811 2	0.993 1	0.051 1	0.051 1
		$\beta_4$	0.002 7	-0.994 0	-1.281 6	-0.759 8	0.160 8	0.160 4	0.001 4	-0.995 4	-1.083 5	-0.917 2	0.052 1	0.052 0
		$\sigma^2$	0.069 9	1.057 1	0.655 6	0.507 2	0.262 6	0.271 1	0.000 4	1.003 6	0.850 6	1.156 8	0.091 2	0.091 0
		$\rho$	0.032 5	-0.478 2	-0.664 6	-0.228 9	0.130 2	0.133 9	0.000 0	-0.502 4	-0.538 9	-0.411 6	0.046 7	0.046 6
			SUM (RMSE)					1.058 3						0.356 0
0.9		$\beta_1$	0.010 6	-1.011 9	0.761 3	1.249 7	0.151 8	0.151 8	0.001 1	0.999 7	0.915 6	1.079 1	0.049 2	0.049 1
		$\beta_2$	0.001 6	-0.711 5	-0.952 0	-0.464 9	0.145 3	0.145 0	0.000 7	-0.705 4	-0.804 2	-0.627 9	0.053 8	0.053 6
		$\beta_3$	0.001 1	0.904 5	0.669 6	1.166 1	0.150 4	0.150 0	0.002 5	0.905 9	0.830 9	0.980 2	0.047 0	0.046 9
		$\beta_4$	0.013 1	-0.999 2	-1.199 1	-0.775 6	0.138 8	0.139 0	0.000 4	-0.998 7	-1.087 2	-0.907 9	0.053 9	0.053 7
		$\sigma^2$	0.089 7	1.059 6	0.652 7	0.653 0	0.310 9	0.322 9	0.009 7	0.999 4	0.877 7	1.164 7	0.088 8	0.089 1
		$\rho$	0.025 6	0.889 0	0.750 2	0.932 5	0.052 6	0.058 4	0.002 0	0.899 2	0.869 5	0.916 9	0.014 2	0.014 3
			SUM (RMSE)					0.967 1						0.306 7

3.2 对比分析研究

有效样本尺度<sup>[23]</sup>(effective sample size, ESS)是反映贝叶斯抽样算法有效性的一个常用指标,常被用来比较不同 MCMC 算法的性能<sup>[24-25]</sup>。根据文献[23-24]的讨论,较小的 ESS 意味着马尔可夫链中后验样本较高的相关性和较差的混合。更多关于 ESS 的细节,请参阅文献[23-25]等。

为了说明本文提出贝叶斯抽样方法的有效性,下面是第 2 个模拟仿真研究。第 2 个模拟研究的模型和参数设置完全与第 1 个模拟研究相同,本文在不同空间参数设置与不同  $\sigma^2$  真值设置情形下,并基于样本容量分别为  $n=200,500$  这 2 种情形随机产生 200 次重复抽样数据,在不同的先验信息条件下分别使用文章第 2 章介绍的 MH-Gibbs 抽样算法和本文提出的切片 Gibbs 抽样算法 2 种方法进行拟合。

为了节约空间,表 3—4 只报道了上述 2 种算法的部分数值计算结果,其中 AVESS 表示基于 200 次重复数据得到的有效样本尺度估计的平均值。由表 3—4(其中, MH 抽样算法的总平均接受率为 0.510 5)可得:(1)就指标 Bias、STD 和 RMSE 而言,基于 MH-Gibbs 抽样算法和基于切片 Gibbs 抽样算法 2 种方法得到的参数估计结果非常接近;同时也能得到类似于第 1 个模拟研究的结论;(2)就指标 AVESS 而言,基于切片 Gibbs 抽样算法的结果明显优于基于 MH-Gibbs 抽样算法的结果,说明文章提出来的贝叶斯抽样算法更为有效。

表 3 第二个模拟仿真研究的数值结果(基于  $n=200, \sigma^2=1$  和先验类型 2)  
Table 3 The numerical result of the simulation 2(based on  $n=200, \sigma^2=1$  and prior Type 2)

$\rho$	参数	MH-Gibbs 抽样算法				切片 Gibbs 抽样算法			
		Bias	STD	RMSE	AVESS	Bias	STD	RMSE	AVESS
0.1	$\beta_1$	0.002 4	0.073 4	0.073 3	8 438.8	0.002 2	0.073 2	0.073 0	8 657.6
	$\beta_2$	0.000 4	0.071 9	0.071 8	8 383.2	0.000 4	0.071 7	0.071 5	8 596.0
	$\beta_3$	0.001 3	0.071 5	0.071 4	8 440.1	0.001 2	0.071 8	0.071 6	8 650.4
	$\beta_4$	0.012 4	0.070 2	0.071 1	8 420.8	0.012 5	0.070 1	0.071 1	8 602.9
	$\sigma^2$	0.009 1	0.107 8	0.107 9	8 106.3	0.009 3	0.107 9	0.108 1	8 290.3
	$\rho$	0.003 8	0.071 3	0.071 2	1 825.6	0.004 0	0.071 1	0.071 1	8 061.8
		SUM (RMSE&AVESS)			0.466 7	43 614.8			0.466 4

表 3(续)

$\rho$	参数	MH-Gibbs 抽样算法				切片 Gibbs 抽样算法			
		Bias	STD	RMSE	AVESS	Bias	STD	RMSE	AVESS
-0.5	$\beta_1$	0.004 8	0.066 3	0.066 4	8 629.1	0.004 7	0.066 6	0.066 6	8 703.4
	$\beta_2$	0.008 8	0.067 9	0.068 3	8 596.7	0.008 8	0.067 9	0.068 3	8 761.5
	$\beta_3$	0.003 8	0.064 7	0.064 7	8 581.9	0.003 6	0.064 6	0.064 6	8 685.4
	$\beta_4$	0.002 7	0.071 8	0.071 7	8 543.9	0.002 8	0.072 0	0.071 8	8 694.2
	$\sigma^2$	0.012 1	0.107 3	0.107 7	4 983.5	0.012 3	0.107 7	0.108 1	6 741.6
	$\rho$	0.008 4	0.051 6	0.051 2	1 597.0	0.008 6	0.051 4	0.052 0	6 717.1
	SUM (RMSE&AVESS)			0.430 0	40 932.1			0.431 4	48 273.2
0.9	$\beta_1$	0.001 1	0.057 5	0.057 3	8 797.3	0.001 0	0.057 5	0.057 4	8 769.7
	$\beta_2$	0.000 3	0.064 6	0.064 4	8 755.1	0.000 3	0.064 6	0.064 4	8 805.5
	$\beta_3$	0.012 7	0.057 9	0.059 1	8813.3	0.012 8	0.058 0	0.059 3	8 785.8
	$\beta_4$	0.001 2	0.057 7	0.057 6	8 810.5	0.001 4	0.057 8	0.057 7	8 850.2
	$\sigma^2$	0.014 6	0.115 6	0.116 2	3 465.3	0.014 5	0.115 4	0.116 0	5 521.7
	$\rho$	0.004 7	0.017 8	0.018 4	1 414.4	0.004 7	0.017 9	0.018 4	5 683.9
	SUM (RMSE&AVESS)			0.373 0	40 056.9			0.383 2	46 416.8

表 4 第二个模拟仿真研究的数值结果(基于  $n=500$ ,  $\sigma^2=0.25$  和先验类型 1)

Table 4 The numerical result of the simulation 2(based on  $n=500$ ,  $\sigma^2=0.25$  and prior Type 1)

$\rho$	参数	MH-Gibbs 抽样算法				切片 Gibbs 抽样算法			
		Bias	STD	RMSE	AVESS	Bias	STD	RMSE	AVESS
0.1	$\beta_1$	0.000 4	0.021 0	0.021 0	8 665.3	0.000 4	0.021 1	0.021 1	8 737.6
	$\beta_2$	0.001 0	0.021 0	0.021 0	8 654.6	0.000 9	0.021 1	0.021 1	8 721.1
	$\beta_3$	0.000 9	0.023 9	0.023 9	8 674.2	0.000 9	0.023 9	0.023 9	8 747.5
	$\beta_4$	0.000 2	0.021 4	0.021 4	8 678.8	0.000 1	0.021 4	0.021 3	8 726.3
	$\sigma^2$	0.000 1	0.015 4	0.015 4	8 429.5	0.000 1	0.015 4	0.015 4	8 529.1
	$\rho$	0.001 8	0.044 9	0.044 8	1 897.1	0.001 7	0.044 6	0.044 5	8 422.2
	SUM (RMSE&AVESS)			0.147 5	44 999.5			0.147 1	51 883.8
-0.5	$\beta_1$	0.003 1	0.020 3	0.020 4	8 745.8	0.003 0	0.020 2	0.020 4	8 782.0
	$\beta_2$	0.000 5	0.021 7	0.021 6	8 711.8	0.000 6	0.021 6	0.021 6	8 781.6
	$\beta_3$	0.001 2	0.021 2	0.021 2	8 742.4	0.001 2	0.021 3	0.021 2	8 759.6
	$\beta_4$	0.001 1	0.022 0	0.021 9	8 763.8	0.001 2	0.022 0	0.021 9	8 733.6
	$\sigma^2$	0.002 3	0.017 4	0.017 5	5 278.0	0.002 3	0.017 4	0.017 5	6 996.2
	$\rho$	0.006 0	0.034 8	0.035 2	1 663.6	0.005 9	0.034 8	0.035 2	6 977.0
	SUM (RMSE&AVESS)			0.137 8	41 905.4			0.137 8	49 030.0
0.9	$\beta_1$	0.000 9	0.018 6	0.018 6	8 825.9	0.000 9	0.018 5	0.018 5	8 800.1
	$\beta_2$	0.000 5	0.016 6	0.016 6	8 819.7	0.000 5	0.016 6	0.016 5	8 816.2
	$\beta_3$	0.000 7	0.018 4	0.018 4	8 793.7	0.000 7	0.018 4	0.018 3	8 791.7
	$\beta_4$	0.000 9	0.018 3	0.018 3	8 797.3	0.000 9	0.018 2	0.018 2	8 777.1
	$\sigma^2$	0.002 1	0.017 7	0.017 8	3 662.8	0.002 0	0.017 8	0.017 8	5 774.7
	$\rho$	0.001 7	0.012 2	0.012 3	1 469.5	0.001 7	0.012 2	0.012 3	5 826.5
	SUM (RMSE&AVESS)			0.102 0	40 368.9			0.101 6	46 786.3

### 4 实证分析

实证分析采用的犯罪数据<sup>[2]</sup>来自美国俄亥俄州哥伦布市的 49 个社区,数据主要包括 CRIME(社区每千户家庭的住宅盗窃案和车辆盗窃案的总和)、INC(家庭收入)和 HOUSE(房屋价值)等,数据的详细描述请参考文献[2]。在空间误差模型的背景下,本文使用模型(1)研究变量 CRIME 与变量 INC、HOUSE 之间的关系。在具体的空间误差模型中,选择  $y^*$  表示 CRIME 作为响应变量,任意的  $i=1,2,\dots,49$ ,对  $y_i^*$  进行了如下的预处理:

$y_i = \frac{y_i^* - y_{\min}}{y_{\max} - y_{\min}}$ , 其中  $y_{\max}$  与  $y_{\min}$  分别为  $\{y_i^* : i=1,2,\dots,49\}$  的最大值与最小值; $x_1=1, x_2$  表示 INC,  $x_3$  表示 HOUSE 作为协变量。与文献[1]相类似,初始空间权重矩阵  $\omega^* = (\omega_{ij}^*)_{49 \times 49}$  被定义为

$$\omega_{ij}^* = \begin{cases} 1, & \text{如果社区 } i \text{ 与社区 } j \text{ 相邻,} \\ 0, & \text{如果社区 } i \text{ 与社区 } j \text{ 不相邻,} \end{cases}$$

然后再对  $\omega^* = (\omega_{ij}^*)_{n \times n}$  进行行正则化处理,即对任意的  $i=1,2,\dots,49$ ,使得  $\sum_{j=1}^{49} \omega_{ij} = 1$ , 最终获得空间权重矩阵  $\omega$ 。

基于上述描述和处理,本文在先验分布式(6)中取如下的超参数: $\beta_0 = \mathbf{0}_3, H_\beta = 1\ 000I_3, a = b = 0.001, \rho_{\min} = -1, \rho_{\max} = 1$ , 即在无信息先验的条件下使用模型(1)并基于切片 Gibbs 抽样算法拟合该数据。在实施切片 Gibbs 抽样算法过程中,使用参数不同的 3 组初始值产生的马尔可夫链来计算模型未知参数的 EPSR 值<sup>[23-26]</sup>。图 1、2 综合展示了每个未知参数 EPSR 值的计算结果。

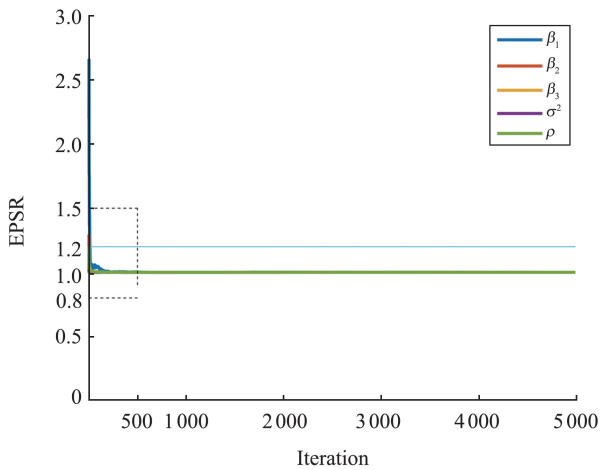


图 1 美国俄亥俄州哥伦布市犯罪数据分析中所有未知参数的 EPSR 值

Fig.1 EPSR values of all parameters for crime data analysis from Columbus, Ohio, USA

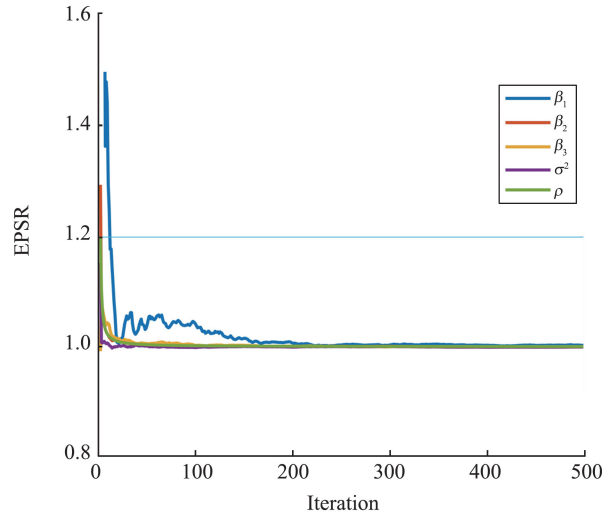


图 2 图 1 的子图(虚线框内)

Fig.2 The subgraph of Fig.1 (inside the dotted box)

由图 1、2 易知,当迭代次数为 1 000 时,每个参数的 EPSR 值的计算结果均未超过 1.2,为了保证算法的收敛性,丢弃 Gibbs 抽样参数前 2 000 次迭代的迭代样本,使用 2 000 次迭代之后的 3 000 个随机样本,按照式(14)计算模型未知参数的贝叶斯估计。最后的估计结果见表 5(其中 MH 算法的接受率为 0.501 8),为了便于比较,同时在表 5 中展示了使用 MH-Gibbs 抽样算法拟合该数据的结果。在表 5 中,EST 表示参数的贝叶斯估计,SD 表示参数贝叶斯估计的标准差,ESS 表示参数有效样本尺度的估计。从表 5 不难看出:一方面使用切片 Gibbs 抽样算法和 MH-Gibbs 抽样算法得到的贝叶斯估计结果几乎是一致的;另一方面,基于切片 Gibbs 抽样算法得到的 ESS 明显大于基于 MH-Gibbs 抽样算法的 ESS,再一次展示了本文提出来的方法(切片 Gibbs 抽样算法)比 Lesage 和 Pace 提出的方法(MH-Gibbs 抽样算法)更为有效。

表5 美国俄亥俄州哥伦布市犯罪数据分析的贝叶斯估计  
Table 5 Bayesian estimation of crime data analysis from Columbus, Ohio, USA

参数	MH-Gibbs 抽样算法			切片 Gibbs 抽样算法		
	EST	SD	ESS	EST	SD	ESS
$\beta_1$	1.164 6	0.147 8	2 802.3	1.169 9	0.151 2	4 646.8
$\beta_2$	-1.851 6	0.505 9	5 884.2	-1.858 9	0.491 0	6 833.2
$\beta_3$	-0.629 5	0.165 3	3 436.2	-0.634 5	0.166 8	4 724.4
$\sigma^2$	0.079 9	0.017 7	6 642.0	0.080 0	0.017 9	7 651.9
$\rho$	0.314 3	0.231 3	1 166.9	0.307 8	0.229 8	3 620.0
SUM(ESS)			19 931.6			27 476.3

## 5 结语

基于空间误差模型,本文提出并使用切片 Gibbs 抽样算法对模型进行贝叶斯参数估计研究,第1个模拟仿真说明了提出来的切片 Gibbs 抽样算法对模型未知参数贝叶斯估计的可行性,第2个模拟仿真说明了切片 Gibbs 抽样算法在 ESS 上的性能表现优于 MH-Gibbs 抽样算法,实证分析说明了提出来的切片 Gibbs 抽样算法对于实际数据的拟合是非常有效的。

文章仅研究了空间误差模型未知参数的贝叶斯估计问题,而空间误差模型的贝叶斯分位数估计、贝叶斯半参数分析以及贝叶斯变量选择等问题,是值得进一步研究的课题。

### 参考文献:

[1] LESAGE J, PACE K R. Introduction to spatial econometrics[M]. Boca Raton: CRC Press, 2010.

[2] ANSELIN L. Spatial econometrics: methods and models[M]. New York: Springer, 1988.

[3] TIAN Ruiqin, XU Dengke, DU Jiang, et al. Bayesian estimation for partially linear varying coefficient spatial autoregressive models[J]. Statistics and Its Interface, 2022, 15(1):105-113.

[4] TANG Yangbing, DU Jiang, ZHANG Zhongzhan. A parametric specification test for linear spatial autoregressive models[J]. Spatial Statistics, 2023, 57(1):100767.

[5] JU Yuanyuan, YANG Han, HU Mingxing, et al. Bayesian influence analysis of the skew-normal spatial autoregression models [J]. Mathematics, 2022, 10(8):1306.

[6] GUO Penghui, LIU Lihu, QIAN Zhengming. Robust test for spatial error model: considering changes of spatial layouts and distribution misspecification[J]. Communications in Statistics: Simulation and Computation, 2015, 44(2):402-416.

[7] CHIARA G, GIANFRANCO P, GIUSEPPE A, et al. Recursive estimation of the spatial error model[J]. Geographical Analysis, 2022, 55(1):90-106.

[8] 李坤明,陈建宝. 非参数空间误差模型的截面最小二乘估计[J]. 数理统计与管理,2020,39(5):824-837.  
LI Kunming, CHEN Jianbao. Profile least square estimation of nonparametric spatial error model[J]. Journal of Applied Statistics and Management, 2020, 39(5):824-837.

[9] VURAL Y, YELIZ K M. Robust estimation approach for spatial error model[J]. Journal of Statistical Computation and Simulation, 2020, 90(9):1618-1638.

[10] DAI Xiaowen, LI Erqian, TIAN Maozai. Quantile regression for varying coefficient spatial error models[J]. Communications in Statistics: Theory and Methods, 2021, 50(10):2382-2397.

[11] 陈小弩. 基于切片抽样 MCMC 方法的比较分析[D]. 成都:四川大学,2007.  
CHEN Xiaonu. Comparative analysis based on slice sampling with MCMC[D]. Chengdu: Sichuan University, 2007.

[12] SENGUPTA B, FRISTON K J, PENNY W D. Gradient-free MCMC methods for dynamic causal modelling [J]. NeuroImage, 2015, 112:375-381.

[13] ZHANG Jiwei, ZHANG Zhaoyuan, TAO Jian. A Bayesian algorithm based on auxiliary variables for estimating GRM with non-ignorable missing data[J]. Computational Statistics, 2021, 36(4):2643-2669.

[14] LU Jing, ZHANG Jiwei, TAO Jian. Slice-Gibbs sampling algorithm for estimating the parameters of a multilevel item

- response model[J]. *Journal of Mathematical Psychology*, 2018, 82:12-25.
- [15] LU Jing, ZHANG Jiwei, ZHANG Zhaoyuan, et al. A novel and highly effective Bayesian sampling algorithm based on the auxiliary variables to estimate the testlet effect models[J]. *Frontiers in Psychology*, 2021, 12:509575.
- [16] METROPOLIS N, ROSENBLUTH A W, ROSENBLUTH M N, et al. Equation of state calculations by fast computing machines[J]. *Journal of Chemical Physics*, 1953, 21(6):1087-1092.
- [17] HASTINGS W K. Monte Carlo sampling methods using Markov chains and their applications[J]. *Biometrika*, 1970, 57(1): 97-109.
- [18] NEAL R M. Slice sampling[J]. *The Annals of Statistics*, 2003, 31(3):705-741.
- [19] LIU J S. Monte Carlo strategies in scientific computing[M]. Berlin: Springer, 2001
- [20] GEMAN S, GEMAN D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984, 6(6):721-741.
- [21] GEYER C J. Practical Markov chain Monte Carlo[J]. *Statistical Science*, 1992, 7(4):473-483.
- [22] CHEN Jiaqing, WANG Renfu, HUANG Yangxin. Semiparametric spatial autoregressive model: a two step Bayesian approach [J]. *Annals of Public Health and Research*, 2015, 2(1):1012.
- [23] GELMAN A, CARLIN B J, STERN S H, et al. Bayesian data analysis[M]. 3rd ed. Oxford: Taylor & Francis, 2013.
- [24] POLSON N G, SCOTT J G, WINDLE J. Bayesian inference for logistic models using Pólya-Gamma latent variables[J]. *Journal of the American Statistical Association*, 2013, 108(504):1339-1349.
- [25] ZHAO Yuanying, XU Dengke, DUAN Xingde, et al. A semiparametric Bayesian approach to binomial distribution logistic mixed-effects models for longitudinal data[J]. *Journal of Statistical Computation and Simulation*, 2022, 92(7):1438-1456.
- [26] 唐年胜, 韦博成. 非线性再生散度模型[M]. 北京: 科学出版社, 2007.
- TIANG Niansheng, WEI Bocheng. Nonlinear reproductive dispersion models[M]. Beijing: Science Press, 2007.

(编辑: 李艺)