

杨 飘, 黄介武. 部分线性变系数分位数模型的贝叶斯 P-样条估计[J]. 智能计算机与应用, 2025, 15(1): 88-94. DOI: 10.20169/j. issn. 2095-2163. 24051604

部分线性变系数分位数模型的贝叶斯 P-样条估计

杨 飘, 黄介武

(贵州民族大学 数据科学与信息工程学院, 贵阳 550025)

摘要: 部分线性变系数模型是一类重要的半参数回归模型, 针对该模型参数估计问题, 本文利用贝叶斯 P-样条方法近似非参数部分的未知光滑函数, 进而利用非对称拉普拉斯分布实现贝叶斯分位数回归, 推导出所有未知参数的条件后验分布, 通过 Gibbs 抽样和 Metropolis-Hastings 算法获得参数的估计值。通过数值模拟对贝叶斯 P-样条方法与 B-样条方法的估计效果进行比较分析, 结果显示在均方误差和标准差准则下, 贝叶斯 P-样条方法在不同分位点上的估计效果更优。

关键词: 部分线性变系数模型; 贝叶斯 P-样条; B-样条; Gibbs 抽样; 均方误差

中图分类号: O212.1

文献标志码: A

文章编号: 2095-2163(2025)01-0088-07

Bayesian P-spline estimation of partially linear variable coefficient quantile model

YANG Piao, HUANG Jiewu

(School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang 550025, China)

Abstract: Partial linear variable coefficient model is an important semi-parametric regression model. For the parameter estimation problem of this model, the Bayesian P-spline method is utilized to approximate the unknown smooth function of the non-parametric part, and then the Bayesian quantile regression is implemented using the asymmetric Laplace distribution to derive the posterior distributions of all the unknown parameters in order to obtain the estimates of the parameters, the parameter estimates were obtained by Gibbs and Metropolis-Hastings algorithm. Meanwhile, the estimation effect of Bayesian P-spline method is compared and analyzed with B-spline method through numerical simulation, and the results show that the Bayesian P-spline method has better estimation effect at different quartiles under the mean square error and standard deviation criterion.

Key words: partial linear variable coefficient model; Bayesian P-spline; B-spline; Gibbs sampling; Mean square error

0 引言

部分线性变系数模型 (Partial Linear Varying Coefficient Models, PLVCMs) 由 Zhang 等^[1] 首次提出。作为线性模型和变系数模型的推广, 部分线性变系数模型由参数和非参数两部分组成, 既具有非参数模型适应性强和回归稳健的特点, 又具备参数模型易于解释的优点。针对部分线性变系数模型参数和非参数的估计问题, 目前已有许多不同的推断方法, Zhang 等^[1] 基于局部多项式拟合参数部分得到相应的估计值; Zhou 等^[2] 采用最小二乘法和小波法分别估计参数和非参数, 同时验证了其渐近正态性和收敛性; Fan 等^[3] 提出了轮廓最小二乘法来估

计模型中的参数分量, 并证明了广义似然方法半参数模型可用于参数分量的检验; Ahmad 等^[4] 运用广义级数法来估计模型的非参数部分; Shang 等^[5] 通过最小化局部 Walsh-average 回归方法研究了当误差分布偏离正态分布时模型的稳健性估计问题; 安佰玲等^[6] 基于 Profile-Liu 估计研究了部分线性变系数模型中参数分量的有偏估计问题。

当随机误差服从正态分布时, 最小二乘法估计法在线性无偏下是最优的, 但是其估计效果容易被异常值所影响, 这也使得最小二乘法存在着较大的局限性, 对数据要求较为苛刻。Koenker 等^[7] 提出了分位数回归方法以弥补最小二乘法估计法的不足, 对模型的限制条件较少, 不受异常值和误差项分

基金项目: 贵州省教育厅自然科学研究项目(黔教技[2022]015号), 黔教技[2023]061号, 黔教技[2023]062号, 黔教技[2023]012号)。

作者简介: 杨 飘(1998—), 女, 硕士研究生, 主要研究方向: 统计模型与统计计算。

通信作者: 黄介武(1977—), 男, 博士, 教授, 主要研究方向: 统计模型与统计计算。Email: 846221886@qq.com。

收稿日期: 2024-05-16

布的影响,具有较强的稳健性。贝叶斯估计方法可以充分利用先验信息推断参数的后验分布, Yu 等^[8]基于非对称拉普拉斯分布实现贝叶斯分位数回归估计,并通过马尔可夫链蒙特卡洛 (MCMC) 算法模拟未知模型参数的后验分布; Wang 等^[9]将分位数应用到部分线性变系数模型中,用基函数近似估计未知光滑系数; 李灿等^[10]基于非对称拉普拉斯分布实现贝叶斯复合分位数回归估计,用 B-样条基函数近似估计未知光滑系数,在均方误差预测意义下,贝叶斯复合分位数方法的预测效果更好; 刘艳霞等^[11]利用函数型主成分分析与 B-样条逼近相结合的方法对变系数部分函数型线性模型的未知函数系数进行估计,得到各估计量的渐近性质; Lang 等^[12-13]首次针对可加模型提出贝叶斯 P-样条,并将贝叶斯 P-样条应用到广义结构加性回归模型中,随后此方法得到广泛应用; 王纯杰等^[14]将贝叶斯 P-样条应用于右删失数据,验证了该方法可降低节点选择对估计的影响; Chen 等^[15-16]利用贝叶斯 P-样条分位数回归方法,研究了部分线性可加空间自回归模型和部分线性变系数空间自回归模型的空间数据,并采用 MCMC 算法探索未知参数的联合后验分布。

本文从贝叶斯角度出发,研究部分线性变系数模型的分位数回归。通过贝叶斯分位数回归方法估计模型的参数部分,进而利用贝叶斯 P-样条方法近似模型中的非参数函数,并通过数值模拟探究该方法的有效性。

1 模型介绍

部分线性变系数模型的一般表现形式如下:

$$Y_i = \mathbf{X}_i^T \boldsymbol{\beta} + \mathbf{Z}_i^T \boldsymbol{\alpha}(U_i) + \varepsilon_i, i = 1, \dots, n \quad (1)$$

其中, $\mathbf{Y} = (Y_1, \dots, Y_n)$ 是 n 维响应变量向量; $\mathbf{X}_i = (X_{i1}, \dots, X_{iq})^T$, $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{ip})^T$ 和 $\mathbf{U} = (U_1, \dots, U_n)^T$ 是解释变量向量; $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)^T$ 是 q 维未知参数向量; $\boldsymbol{\alpha}(\cdot) = (\alpha_1(\cdot), \alpha_2(\cdot), \dots, \alpha_p(\cdot))^T$ 是 p 维未知光滑函数向量; 在给定 $(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{U})$ 的情况下, $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T$ 是均值为 0、方差有限的 n 维随机误差向量,每个 ε_i 是独立同分布的且独立于 $(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{U})$ 。

为模型(1)建立分位数回归模型:

$$(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}(\cdot)) = \underset{\boldsymbol{\beta}, \boldsymbol{\alpha}}{\operatorname{argmin}} \sum_{i=1}^n \rho_{\tau}(Y_i - \mathbf{X}_i^T \boldsymbol{\beta} - \mathbf{Z}_i^T \boldsymbol{\alpha}(U_i)), \quad i = 1, \dots, n \quad (2)$$

其中, $\rho_{\tau}(\cdot)$ 为损失函数, τ 是分位数。

Yu 和 Moyeed^[8]介绍了当误差服从非对称拉普

拉斯分布 (ALD) 时,即 $\varepsilon_i \sim \text{ALD}(0, \sigma, \tau)$, 则 $Y_i \sim \text{ALD}(\mu, \sigma, \tau)$, 其密度函数可以表示如下:

$$f(Y_i | \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\rho_{\tau}\left(\frac{Y_i - \mu}{\sigma}\right)\right\} \quad (3)$$

其中, $\mu = \mathbf{X}_i^T \boldsymbol{\beta} + \mathbf{Z}_i^T \boldsymbol{\alpha}(U_i)$ 是位置参数, $\sigma > 0$ 是尺度参数, $0 < \tau < 1$ 。

则 Y_i 的似然函数如下:

$$L(Y_i | \boldsymbol{\beta}, \boldsymbol{\alpha}(\cdot), \sigma, \tau) = \frac{\tau^n (1-\tau)^n}{\sigma^n} \exp\left\{-\sum_{i=1}^n \rho_{\tau}\left(\frac{Y_i - \mathbf{X}_i^T \boldsymbol{\beta} - \mathbf{Z}_i^T \boldsymbol{\alpha}(U_i)}{\sigma}\right)\right\}, \quad i = 1, 2, \dots, n \quad (4)$$

Kozumi 等^[17-18]指出, ALD 可以表示成正态-指数混合分布,模型(1)可写为:

$$Y_i = \mathbf{X}_i^T \boldsymbol{\beta} + \mathbf{Z}_i^T \boldsymbol{\alpha}(U_i) + \theta_1 e_i + \sqrt{\theta_2 \sigma e_i} v_i, i = 1, \dots, n \quad (5)$$

$$\text{其中, } \theta_1 = \frac{1-2\tau}{\tau(1-\tau)}, \theta_2 = \frac{2}{\tau(1-\tau)},$$

$$e \sim \text{Exp}\left(\frac{1}{\sigma}\right), v \sim N(0, 1).$$

上式可写为正态分布的形式:

$$Y_i \sim N(\mathbf{X}_i^T \boldsymbol{\beta} + \mathbf{Z}_i^T \boldsymbol{\alpha}(U_i) + \theta_1 e_i, \theta_2 \sigma e_i), i = 1, \dots, n \quad (6)$$

当误差服从 ALD 时,求解式(2)完全等价于通过误差服从非对称拉普拉斯分布而形成的似然函数的最大化。

2 贝叶斯 P-样条估计

2.1 贝叶斯 P-样条近似

针对部分线性变系数模型非参数部分的估计问题,出现了许多不同的估计方法,如局部多项式估计、光滑样条估计、核估计等。在这些方法中, B-样条估计较为成熟,对于 $j = 1, \dots, p$, 假设解释变量的定义域为 $[a, b]$, 在定义域的等分节点 $a < \xi_{j1} < \dots < \xi_{jm} < b$ 上,非参数部分的未知光滑函数 $\alpha(\cdot)$ 由 m 个内部节点的 k 次 B-样条基函数近似,可以写成 $L = k + m + 1$ 个 B-样条基函数的线性组合的形式。假设 $\mathbf{B}(U_i) = (B_1(U_i), \dots, B_L(U_i))^T$ 为样条基函数向量,那么 $\alpha_j(U_i)$ 可近似为 $\mathbf{B}^T(U_i) \boldsymbol{\gamma}_j, i = 1, \dots, n; j = 1, \dots, p$, 即:

$$\alpha_j(U_i) \approx \sum_{r=1}^L B_r(U_i) \gamma_{jr} = \mathbf{B}^T(U_i) \boldsymbol{\gamma}_j \quad (7)$$

其中, $\boldsymbol{\gamma}_j = (\gamma_{j1}, \dots, \gamma_{jL})^T$ 为样条系数向量。

则模型(1)可写为:

$$Y_i = \mathbf{X}_i^T \boldsymbol{\beta} + \mathbf{Z}_i^T \mathbf{B}^T(U_i) \boldsymbol{\gamma}_j + \varepsilon_i, i = 1, \dots, n; j = 1, \dots, p \quad (8)$$

在B-样条中,当内部节点 m 过大或过小时,容易出现过拟合或欠拟合的情况。因此 Eilers 等^[19]针对这种情况提出了P-样条方法,即在每个样条系数上添加一个粗糙惩罚项,公式如下:

$$L = l(Y | X, Z, \beta, \gamma) - \lambda \sum_{k=t+1}^m (\Delta^k \gamma_j)^2 \quad (9)$$

其中, λ 为惩罚参数; $\Delta^t \gamma_j$ 为 γ_j 的 t 阶差分,通常情况下取 $t = 1$ 或 $t = 2$ 。

根据贝叶斯思想,未知参数可以当作未知的随机变量,给其赋予先验分布。Lang 和 Brezger^[12]在P-样条基础上加入了贝叶斯思想,得到贝叶斯P-样条方法,对样条系数 γ_j 使用二阶随机游走模型代替,即样条系数 γ_j 取二阶差分,公式如下:

$$\gamma_j = 2\gamma_{j-1} - \gamma_{j-2} + \omega_j \quad (10)$$

其中, ω_j 是高斯误差, $\omega_j \sim N(0, \delta_j^2)$ 。

参数 δ_j^2 用于控制曲线的光滑度。惩罚参数一般由交叉验证获得,计算量较大,本文给样条系数 γ 赋予先验,可以避免惩罚参数的选取。对于样条系数 γ ,赋予随机游走先验:

$$\pi(\gamma_j | \delta_j) \propto \left(\frac{1}{\sqrt{2\pi\delta_j^2}} \right)^{L-t} \exp \left\{ -\frac{1}{2\delta_j^2} \gamma_j^T H_{\gamma_j} \gamma_j \right\} \quad (11)$$

其中, t 是随机游走的阶数, H_{γ_j} 是 t 阶随机游走先验的惩罚矩阵,等于 $(Q_{t-1} \times \dots \times Q_0)^T (Q_{t-1} \times \dots \times Q_0)$, 其中 Q_t 是 $(L-l-1) \times (L-1)$ 维矩阵,公式如下:

$$Q_l = \begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -1 & 1 \end{pmatrix}, l = 0, \dots, t-1 \quad (12)$$

超参数 δ_j^2 的先验分布为 $\delta_j^2 \sim IG(a_\delta, b_\delta)$, 其中,超参数 a_δ, b_δ 是已知的,通常取 $a_\delta = 1, b_\delta = 0.005$ 或 $b_\delta = 0.0005$ 。由于对模型(1)的非参数部分进行了贝叶斯P-样条近似,那么部分线性变系数模型的贝叶斯分位数回归模型可以写成下式:

$$Y_i = X_i^T \beta + Z_i^T B^T(U_i) \gamma_j + \theta_1 e_i + \sqrt{\theta_2 \sigma e_i} v_i, \quad i = 1, \dots, n; j = 1, \dots, p \quad (13)$$

2.2 参数的条件后验分布

首先为参数 β, e, σ, v 分别选择先验分布 $\beta \sim N_q(\mu_\beta, \sigma_\beta I)$, $e \sim \exp(\frac{1}{\sigma})$, $\sigma \sim IG(a_\sigma, b_\sigma)$, $v \sim N(0, I)$ 。

其中,超参数 $\mu_\beta, \sigma_\beta, a_\sigma, b_\sigma$ 是已知的。由式

(5), Y_i 的条件概率分布为 $Y_i \sim N(X_i^T \beta + Z_i^T B^T(U_i) \gamma_j + \theta_1 e_i, \theta_2 \sigma e_i)$, 其联合后验密度函数可表示为 $\pi(\beta, e_i, \sigma, \gamma_j, \delta_j^2 | Y_i) \propto \pi(Y_i | \beta, e_i, \sigma, \gamma_j, \delta_j^2) \pi(\beta) \pi(\sigma) \cdot \pi(e_i) \prod_{j=1}^p \pi(\gamma_j | \delta_j^2) \pi(\delta_j^2)$, 得到参数 $\beta, e_i, \sigma, \gamma_j, \delta_j^2$ 的条件后验密度分别为:

(1) $\pi(\beta | Y_i, e_i, \sigma, \gamma_j, \delta_j^2) \propto \pi(Y_i | \beta, e_i, \sigma, \gamma_j, \delta_j^2) \pi(\beta) \propto \prod_{i=1}^n \prod_{j=1}^p \left(\frac{1}{\sqrt{\theta_2 \sigma e_i}} \exp \left(-\frac{1}{2\theta_2 \sigma e_i} (Y_i - X_i^T \beta - Z_i^T B^T(U_i) \gamma_j - \theta_1 e_i)^2 \right) \exp \left(-\frac{1}{2} (\beta - \mu'_\beta)^T (\sigma'_\beta I_q)^{-1} (\beta - \mu'_\beta) \right) \right) \propto \exp \left(-\frac{1}{2} (\beta - \mu'_\beta)^T (\sigma'_\beta)^{-1} (\beta - \mu'_\beta) \right)$, 其条件后验分布为 $\beta | Y_i, e_i, \sigma, \gamma_j, \delta_j^2 \sim N(\mu'_\beta, \sigma'_\beta)$, 其中 $\mu'_\beta = \sigma'_\beta \sum_{i=1}^n \sum_{j=1}^p \frac{1}{\theta_2 \sigma e_i} (X_i (Y_i - Z_i^T B^T(U_i) \gamma_j - \theta_1 e_i)) + (\sigma'_\beta I_q)^{-1} \mu_\beta, \sigma'_\beta = \left(\sum_{i=1}^n \frac{X_i^T X_i}{\theta_2 \sigma e_i} (\sigma'_\beta I_q)^{-1} \right)^{-1}$ 。

(2) $\pi(\sigma | Y_i, \beta, e_i, \gamma_j, \delta_j^2) \propto \pi(Y_i | \beta, e_i, \sigma, \gamma_j, \delta_j^2) \pi(\sigma) \pi(e_i) \propto \prod_{i=1}^n \prod_{j=1}^p \left(\frac{1}{\sqrt{\theta_2 \sigma e_i}} \exp \left(-\frac{1}{2\theta_2 \sigma e_i} (Y_i - X_i^T \beta - Z_i^T B^T(U_i) \gamma_j - \theta_1 e_i)^2 \right) \frac{1}{\sigma} \exp \left(-\frac{e_i}{\sigma} \right) \left(\frac{1}{\sigma} \right)^{a_\sigma+1} \exp \left(-\frac{b_\sigma}{\sigma} \right) \right) \propto \left(\frac{1}{\sigma} \right)^{\left(\frac{3n}{2} + a_\sigma + 1 \right)} \exp \left(-\frac{1}{\sigma} \left(\sum_{i=1}^n \sum_{j=1}^p \left(\frac{1}{2\theta_2 e_i} (Y_i - X_i^T \beta - Z_i^T B^T(U_i) \gamma_j - \theta_1 e_i)^2 + e_i \right) + b_\sigma \right) \right)$, 其条件后验分布为 $\sigma | Y_i, e_i, \beta, \gamma_j, \delta_j^2 \sim IG(a'_\sigma, b'_\sigma)$, 其中 $a'_\sigma = \frac{3n}{2} + a_\sigma, b'_\sigma = \sum_{i=1}^n \sum_{j=1}^p \left(\frac{(Y_i - X_i^T \beta - Z_i^T B^T(U_i) \gamma_j - \theta_1 e_i)^2}{2\theta_2 e_i} + e_i \right) + b_\sigma$ 。

(3) $\pi(e_i | Y_i, \beta, \sigma, \gamma_j, \delta_j^2) \propto \pi(Y_i | \beta, e_i, \sigma, \gamma_j, \delta_j^2) \pi(e_i) \propto \prod_{i=1}^n \prod_{j=1}^p \left(\frac{1}{\sqrt{\theta_2 \sigma e_i}} \exp \left(-\frac{1}{2\theta_2 \sigma e_i} (Y_i - X_i^T \beta - Z_i^T B^T(U_i) \gamma_j - \theta_1 e_i)^2 \right) \frac{1}{\sigma} \exp \left(-\frac{e_i}{\sigma} \right) \right) \propto \left(\frac{1}{e_i} \right)^{\frac{1}{2}} \exp \left(-\frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^p \frac{(Y_i - X_i^T \beta - Z_i^T B^T(U_i) \gamma_j)^2}{\theta_2 \sigma} e_i^{-1} + \left(\frac{\theta_1^2}{\theta_2 \sigma} + \frac{2}{\sigma} \right) e_i \right) \right) \propto \left(\frac{1}{e_i} \right)^{\frac{1}{2}} \exp \left(-\frac{1}{2} (a_e e_i^{-1} + b_e e_i) \right)$, 其条件

后验分布为 $e_i | Y_i, \boldsymbol{\beta}, \sigma, \boldsymbol{\gamma}_j, \delta_j^2 \sim GIG(\lambda, a_e, b_e)$, 其中 $\lambda = \frac{1}{2}, a_e = \frac{1}{2\sigma^2} (Y_i - \mathbf{X}_i^T \boldsymbol{\beta} - \mathbf{Z}_i^T \mathbf{B}^T(U_i) \boldsymbol{\gamma}_j)^2, b_e = \frac{\theta_1^2}{2\sigma^2} + \frac{2}{\sigma}, GIG(\lambda, a, b)$ 表示广义逆高斯分布, 其概率密度函数为: $f(x | \lambda, a, b) = \frac{x^{-\lambda} (\sqrt{ab})^\lambda}{2K_\lambda(\sqrt{ab})} x^{\lambda-1} \exp\left(-\frac{1}{2}(ax^{-1} + bx)\right), x > 0, -\infty < \lambda < \infty, a, b \geq 0$.

$$(4) \pi(\boldsymbol{\gamma}_j | Y_i, \boldsymbol{\beta}, \sigma, e_i, \delta_j^2) \propto \pi(Y_i | \boldsymbol{\beta}, e_i, \sigma, \boldsymbol{\gamma}_j, \delta_j^2) \pi(\boldsymbol{\gamma}_j) \propto \prod_{i=1}^n \prod_{j=1}^p \left(\exp\left(-\frac{1}{2\theta_2 \sigma e_i} (Y_i - \mathbf{X}_i^T \boldsymbol{\beta} - \mathbf{Z}_i^T \mathbf{B}^T(U_i) \boldsymbol{\gamma}_j - \theta_1 e_i)^2 \exp\left(-\frac{1}{2\delta_j^2} \boldsymbol{\gamma}_j^T \mathbf{H}_{\boldsymbol{\gamma}_j} \boldsymbol{\gamma}_j\right)\right) \right) \propto \exp\left(-\sum_{i=1}^n \sum_{j=1}^p \left(-\frac{1}{2\theta_2 \sigma e_i} (Y_i - \mathbf{X}_i^T \boldsymbol{\beta} - \mathbf{Z}_i^T \mathbf{B}^T(U_i) \boldsymbol{\gamma}_j - \theta_1 e_i)^2 - \frac{1}{2\delta_j^2} \boldsymbol{\gamma}_j^T \mathbf{H}_{\boldsymbol{\gamma}_j} \boldsymbol{\gamma}_j\right)\right).$$

$$(5) \pi(\delta_j^2 | Y_i, \boldsymbol{\beta}, e_i, \sigma, \boldsymbol{\gamma}_j) \propto \pi(Y_i | \boldsymbol{\beta}, e_i, \sigma, \boldsymbol{\gamma}_j, \delta_j^2) \pi(\delta_j^2) \propto \prod_{j=1}^p \left(\left(-\frac{1}{\sqrt{2\pi\delta_j^2}}\right)^{L-t} \exp\left(-\frac{1}{2\delta_j^2} \boldsymbol{\gamma}_j^T \mathbf{H}_{\boldsymbol{\gamma}_j} \boldsymbol{\gamma}_j\right) \left(\frac{1}{\delta_j^2}\right)^{a_\delta'+1} \exp\left(-\frac{b_\delta'}{\delta_j^2}\right) \right) \propto \left(\frac{1}{\delta_j^2}\right)^{\frac{1}{2}(L-t)+a_\delta'+1} \exp\left(\sum_{j=1}^p -\frac{1}{2\delta_j^2} (\boldsymbol{\gamma}_j^T \mathbf{H}_{\boldsymbol{\gamma}_j} \boldsymbol{\gamma}_j + 2b_\delta')\right),$$

其条件后验分布为 $\delta_j^2 | Y_i, \boldsymbol{\beta}, \sigma, \boldsymbol{\gamma}_j, e_i \sim IG(a'_\delta, b'_\delta)$, 其中 $a'_\delta = \frac{1}{2}(L-t) + a_\delta, b'_\delta = \sum_{j=1}^p \frac{1}{2} (\boldsymbol{\gamma}_j^T \mathbf{H}_{\boldsymbol{\gamma}_j} \boldsymbol{\gamma}_j + 2b_\delta)$.

2.3 MCMC 算法

在贝叶斯算法中, Gibbs 抽样和 Metropolis-Hastings (MH) 算法是最简单且应用最广泛的两种抽样方法^[20]. 本文主要是基于 Gibbs 和 MH 算法对模型的参数进行估计.

未知参数 $\boldsymbol{\beta}, e, \sigma, \delta^2$ 的条件后验分布是标准分布, 可直接通过 Gibbs 抽样从后验条件分布抽取样本观测值. 未知参数 $\boldsymbol{\gamma}$ 的条件后验分布为非标准分布, 无法从特定的分布中抽取, 所以采用 MH 算法对其进行抽样. 结合 Gibbs 抽样和 MH 算法的混合算法的具体步骤如下:

步骤 1 给定目标平稳分布 $\pi(\boldsymbol{\beta}, e, \sigma, \boldsymbol{\gamma}, \delta^2, Y, X, Z, U)$, 假设迭代次数为 N , 在第 A 次迭代时开始收敛;

步骤 2 给定初始值 $\boldsymbol{\beta}^{(0)}, \sigma^{(0)}, e^{(0)}, (\delta^2)^{(0)}, \boldsymbol{\gamma}^{(0)}$;

步骤 3 对于 $t = 1, 2, \dots, A, \dots, N$, 假设第 $t-1$ 次迭代的估计值为 $(\boldsymbol{\beta}^{(t-1)}, e^{(t-1)}, \sigma^{(t-1)}, (\delta^2)^{(t-1)}, \boldsymbol{\gamma}^{(t-1)})$, 利用 Gibbs 抽样, 执行以下步骤:

(1) 从条件后验分布 $\pi(\boldsymbol{\beta} | Y, e^{(t-1)}, \sigma^{(t-1)}, (\delta^2)^{(t-1)}, \boldsymbol{\gamma}^{(t-1)})$ 中抽取观测值 $\boldsymbol{\beta}^{(t)}$, 用 $\boldsymbol{\beta}^{(t)}$ 更新 $\boldsymbol{\beta}$;

(2) 从条件后验分布 $\pi(e | Y, \boldsymbol{\beta}^{(t-1)}, \sigma^{(t-1)}, (\delta^2)^{(t-1)}, \boldsymbol{\gamma}^{(t-1)})$ 中抽取观测值 $e^{(t)}$, 用 $e^{(t)}$ 更新 e ;

(3) 从条件后验分布 $\pi(\sigma | Y, \boldsymbol{\beta}^{(t-1)}, e^{(t-1)}, (\delta^2)^{(t-1)}, \boldsymbol{\gamma}^{(t-1)})$ 中抽取观测值 $\sigma^{(t)}$, 用 $\sigma^{(t)}$ 更新 σ ;

(4) 从条件后验分布 $\pi(\delta^2 | Y, \boldsymbol{\beta}^{(t-1)}, e^{(t-1)}, \sigma^{(t-1)}, \boldsymbol{\gamma}^{(t-1)})$ 中抽取观测值 $(\delta^2)^{(t)}$, 用 $(\delta^2)^{(t)}$ 更新 δ^2 ;

步骤 4 利用 MH 算法, 选取合适的建议分布 $N(\boldsymbol{\gamma}, \sigma_1^2 I)$, 其中调节参数 σ_1 是事先给定的. 从建议分布中随机产生一个观测值 $\boldsymbol{\gamma}^*$, 从 $U(0, 1)$ 随机抽取一个数 u , 如果

$$u \leq \min\left(1, \frac{\pi(\boldsymbol{\gamma}^* | Y, \boldsymbol{\beta}^{(t-1)}, \sigma^{(t-1)}, e^{(t-1)}, (\delta^2)^{(t-1)})}{\pi(\boldsymbol{\gamma}^{(t-1)} | Y, \boldsymbol{\beta}^{(t-1)}, \sigma^{(t-1)}, e^{(t-1)}, (\delta^2)^{(t-1)})}\right)$$

成立, 则令 $\boldsymbol{\gamma}' = \boldsymbol{\gamma}^*$, 否则令 $\boldsymbol{\gamma}' = \boldsymbol{\gamma}^{t-1}$;

步骤 5 重复步骤 3 和步骤 4, 迭代 N 次, 返回样本序列 $(\boldsymbol{\beta}^{(1)}, e^{(1)}, \sigma^{(1)}, (\delta^2)^{(1)}, \boldsymbol{\gamma}^{(1)}), (\boldsymbol{\beta}^{(2)}, e^{(2)}, \sigma^{(2)}, (\delta^2)^{(2)}, \boldsymbol{\gamma}^{(2)}), \dots, (\boldsymbol{\beta}^{(A)}, e^{(A)}, \sigma^{(A)}, (\delta^2)^{(A)}, \boldsymbol{\gamma}^{(A)}), \dots, (\boldsymbol{\beta}^{(N)}, e^{(N)}, \sigma^{(N)}, (\delta^2)^{(N)}, \boldsymbol{\gamma}^{(N)})$;

步骤 6 输出样本 $(\boldsymbol{\beta}^{(t)}, e^{(t)}, \sigma^{(t)}, (\delta^2)^{(t)}, \boldsymbol{\gamma}^{(t)}), t = 1, 2, \dots, A, \dots, N$.

将抽样获得的样本取均值即为参数 $\boldsymbol{\beta}, e, \sigma, \delta^2, \boldsymbol{\gamma}$ 的估计值.

3 数值模拟

基于上述理论研究, 假定模型为 $Y = X_1 + 2X_2 + 2\sin(2\pi U)Z_1 + (1 - 2U)^2 Z_2 + \varepsilon$, 利用 R 软件进行随机模拟, 分别取样本量 $n = 100, n = 200, X = (X_1, X_2)^T, Z = (Z_1, Z_2)^T, X_i \sim N(0, I_2), Z_i \sim N(0, I_2), U \sim N(0, I), \varepsilon \sim N(0, 1)$, 其中 $\alpha_1(U) = 2\sin(2\pi U), \alpha_2(U) = (1 - 2U)^2, \boldsymbol{\beta} = (1, 2)^T$. 假设 $m = 15, k = 2, L = 18$, 根据各未知参数的条件后验分布, 利用 Gibbs 抽样和 MH 算法分别在分位数 $\tau = 0.1, 0.3, 0.5, 0.7, 0.9$ 处进行参数估计, 把得到的样本均值作为未知参数的最终估计值.

在样本量 n 分别为 100 和 200 时, 参数 β_1 和 β_2 在不同分位数下的均值、中位数、最小值以及最大值的估计效果见表 1, 可以看出参数的估计值较为接近真实值, 效果较好.

考虑 3 种类型的误差分布: 标准正态分布 $N(0, 1)$ 、自由度为 3 的 t 分布 $t(3)$ 、自由度为 2 的 χ^2 分

布 $\chi^2(2)$ 。在分位数 $\tau = 0.3, 0.5, 0.7$ 下,对每一种误差分布生成 100 个数据集,并分别使用贝叶斯 P-样条和 B-样条方法进行估计。两种估计方法的比

较结果见表 2,可以看出在不同的误差分布下,贝叶斯 P-样条方法拥有较小的均方误差 (MSE) 和标准差 (SD),说明其估计效果要优于 B-样条方法。

表 1 不同分位点上参数估计结果

Table 1 Parameter estimation results at different loci

τ	参数	真值	样本 ($n = 100$)				样本 ($n = 200$)			
			均值	中位数	最小值	最大值	均值	中位数	最小值	最大值
0.1	β_1	1	1.032 1	1.041 3	0.992 3	1.210 5	1.023 3	1.022 5	0.954 5	1.115 6
	β_2	2	2.104 0	2.119 4	1.987 2	2.156 9	1.971 6	1.975 4	1.934 5	2.007 2
0.3	β_1	1	0.965 8	0.995 1	0.893 9	1.154 4	1.004 1	1.007 9	0.963 1	1.036 6
	β_2	2	2.023 9	2.025 9	1.937 9	2.208 1	2.011 3	2.012 9	1.978 6	2.117 4
0.5	β_1	1	0.997 8	1.007 5	0.955 4	1.035 6	0.982 9	0.981 9	0.956 2	1.092 2
	β_2	2	1.994 9	1.996 8	1.952 7	2.026 4	1.992 2	1.992 6	1.954 1	2.025 3
0.7	β_1	1	1.045 5	1.034 1	0.972 4	1.109 8	0.988 2	0.995 3	0.899 7	1.018 7
	β_2	2	1.993 4	1.983 8	1.943 8	2.076 6	2.106 7	2.121 2	1.986 3	2.244 5
0.9	β_1	1	0.994 1	0.994 7	0.971 3	1.133 1	1.029 4	1.033 0	0.985 7	1.053 8
	β_2	2	2.056 0	2.044 7	1.938 7	2.248 8	2.015 4	2.015 4	1.930 1	2.067 7

表 2 两种估计方法的均方误差和标准差

Table 2 Mean square error and standard deviation of the two estimation methods

误差分布	τ	参数	贝叶斯 P-样条		B-样条	
			MSE	SD	MSE	SD
$\varepsilon \sim N(0,1)$	0.3	β_1	0.008 7	0.047 3	0.019 1	0.059 8
		β_2	0.004 6	0.039 7	0.009 3	0.061 9
	0.5	β_1	0.005 7	0.039 9	0.023 7	0.065 8
		β_2	0.003 0	0.017 3	0.006 9	0.036 1
	0.7	β_1	0.010 2	0.033 6	0.017 4	0.065 0
		β_2	0.004 1	0.044 1	0.006 7	0.047 9
$\varepsilon \sim t(3)$	0.3	β_1	0.002 7	0.035 0	0.029 3	0.045 4
		β_2	0.000 7	0.032 7	0.005 4	0.072 4
	0.5	β_1	0.011 6	0.028 4	0.036 5	0.039 9
		β_2	0.005 7	0.030 0	0.007 1	0.046 4
	0.7	β_1	0.004 0	0.048 1	0.017 8	0.065 8
		β_2	0.002 9	0.039 1	0.003 3	0.040 5
$\varepsilon \sim \chi^2(2)$	0.3	β_1	0.001 8	0.039 6	0.003 9	0.043 7
		β_2	0.002 2	0.033 2	0.013 9	0.056 9
	0.5	β_1	0.000 4	0.064 6	0.007 4	0.064 1
		β_2	0.006 2	0.028 0	0.028 1	0.057 3
	0.7	β_1	0.003 5	0.015 6	0.023 4	0.046 5
		β_2	0.003 9	0.024 8	0.008 9	0.057 1

为检验贝叶斯 P-样条方法的收敛性,利用图像法进行判断,参数 β_1 和 β_2 的迭代轨迹图如图 1 所示

设定 MCMC 算法抽样的迭代次数为 6 000,燃烧次数为 1 000。

由于在不同的样本量及另外 4 种分位数水平下的结果相似,图 1 只展示了当 $n = 100, \tau = 0.5$ 时,3 种误差分布下参数 β_1 和 β_2 的迭代轨迹图。由图 1 可知,得到后验抽样值是快速收敛的,说明贝叶斯

P-样条方法有较好的收敛性。

$n = 100$ 和 $n = 200$ 下,样条函数 $\alpha_1(U)$ 和 $\alpha_2(U)$ 的拟合图如图 2 和图 3 所示。可以看出贝叶斯 P-样条的拟合图像具有较好的拟合效果。

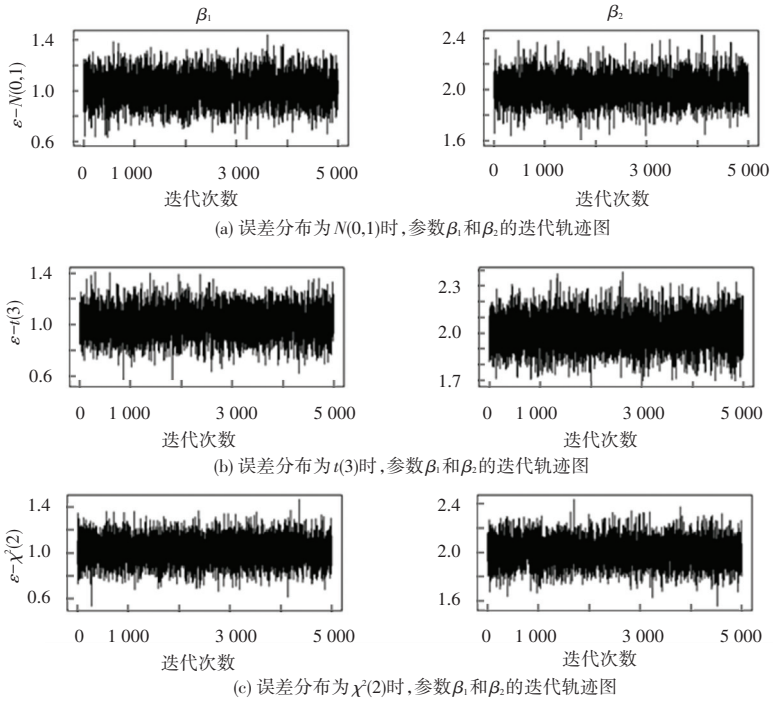


图 1 参数迭代轨迹图

Fig. 1 Parameter iteration trajectory diagram

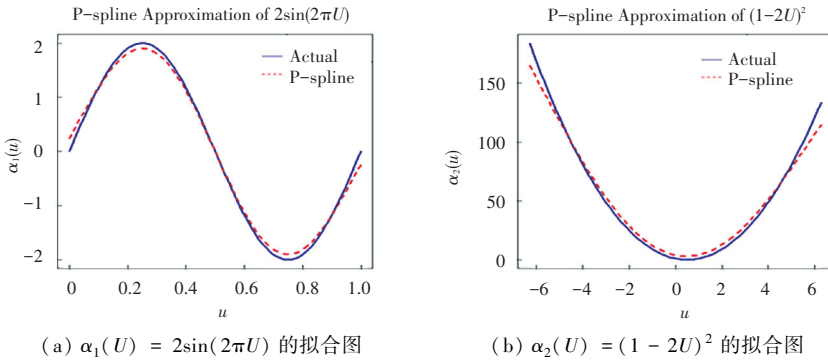


图 2 $n = 100$ 时,贝叶斯 P-样条下函数的拟合图

Fig. 2 Plots of the fit of the function under Bayesian P-splines for $n = 100$

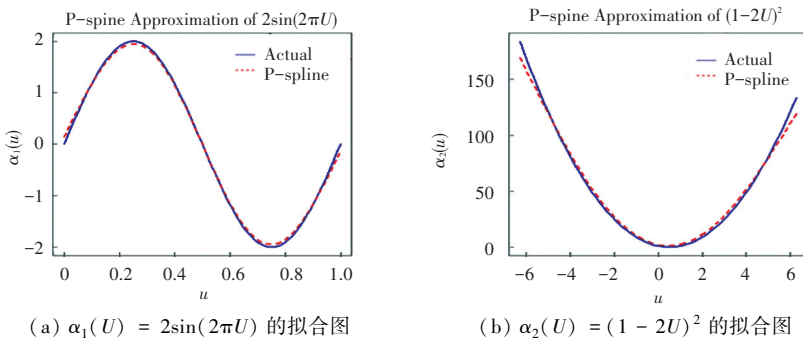


图 3 $n = 200$ 时,贝叶斯 P-样条下函数的拟合图

Fig. 3 Plots of the fit of the function under Bayesian P-splines for $n = 200$

4 结束语

本文利用贝叶斯 P-样条估计方法研究了部分线性变系数模型的分位数回归参数估计问题,并与李灿等^[21]所用的 B-样条估计方法进行比较。随机模拟结果表明,在均方误差和标准差准则下,贝叶斯 P-样条估计方法估计效果更稳定,更接近真实值。

参考文献

- [1] ZHANG W, LEE S, SONG X. Local polynomial fitting in semivarying coefficient model [J]. *Journal of Multivariate Analysis*, 2002, 82(1): 166-188.
- [2] ZHOU X, YOU J. Wavelet estimation in varying-coefficient partially linear regression models [J]. *Statistics & Probability Letters*, 2004, 68(1): 91-104.
- [3] FAN J, HUANG T. Profile likelihood inferences on semiparametric varying-coefficient partially linear models [J]. *Bernoulli*, 2005, 11(6): 1031-1057.
- [4] AHMAD I, LEELAHANON S, LI Q. Efficient estimation of a semiparametric partially linear varying coefficient model [J]. *The Annals of Statistics*, 2005, 33(1): 258-283.
- [5] SHANG S, ZUO C, WANG Z. Local walsh-average regression for semiparametric varying-coefficient models [J]. *Statistics & Probability Letters*, 2012, 82(10): 1815-1822.
- [6] 安佰玲, 卢琦, 马宁. 部分线性变系数模型的 Liu 估计 [J]. *吉首大学学报(自然科学版)*, 2022, 43(4): 1-6.
- [7] KOENKER R, BASSETT G. Regression quantiles [J]. *Econometrica*, 1978, 46(1): 33-50.
- [8] YU K, MOYEED R. Bayesian quantile regression [J]. *Statistics & Probability Letters*, 2001, 54(4): 437-447.
- [9] WANG H J, ZHU Z, ZHOU J. Quantile regression in partially linear varying coefficient models [J]. *The Annals of Statistics*, 2009, 37(6B): 3841-3866.
- [10] 李灿, 杨建波, 李荣. 部分线性变系数模型的贝叶斯复合分位数回归 [J]. *广西师范大学学报(自然科学版)*, 2024, 42(5): 117-129.
- [11] 刘艳霞, 王芝皓, 田茂再. 变系数部分函数型线性模型的复合分位数回归估计 [J/OL]. *数理统计与管理*, 1-12 [2024-10-31]. DOI:10.13860/j.cnki.slj.20230629-001.
- [12] LANG S, BREZGER A. Bayesian P-splines [J]. *Journal of Computational and Graphical Statistics*, 2004, 13(1): 183-212.
- [13] BREZGER A, LANG S. Generalized structured additive regression based on Bayesian P-splines [J]. *Computational Statistics & Data Analysis*, 2006, 50(4): 967-991.
- [14] 王纯杰, 罗琳琳, 李纯净, 等. 删失数据下部分线性模型的贝叶斯 P-样条估计 [J]. *东北师大学报(自然科学版)*, 2020, 52(4): 25-32.
- [15] CHEN Z, CHEN M, XING G. Bayesian estimation of partially linear additive spatial autoregressive models with P-splines [J]. *Mathematical Problems in Engineering*, 2021, 2021(1): 1777469.
- [16] CHEN Z, CHEN M, JU F. Bayesian P-splines quantile regression of partially linear varying coefficient spatial autoregressive models [J]. *Symmetry*, 2022, 14(6): 1175.
- [17] KOZUMI H, KOBAYASHI G. Gibbs sampling methods for Bayesian quantile regression [J]. *Journal of Statistical Computation and Simulation*, 2011, 81(11): 1565-1578.
- [18] YANG F, SHAN A, YUAN H. Gibbs sampling for mixture quantile regression based on asymmetric Laplace distribution [J]. *Communications in Statistics-Simulation and Computation*, 2019, 48(5): 1560-1573.
- [19] EILERS P, MARX B. Flexible smoothing using B-splines and penalized likelihood (with comments and rejoinder) [J]. *Statistical Science*, 1996, 11(2): 89-121.
- [20] 肖枝洪, 朱强. 统计模拟及其 R 实现 [M]. 武汉: 武汉大学出版社, 2010.
- [21] 李灿, 杨建波, 李荣. 部分线性变系数模型的贝叶斯分位数回归 [J]. *湖南文理学院学报(自然科学版)*, 2024, 36(1): 7-13.