Advances in Applied Mathematics 应用数学进展, 2020, 9(12), 2166-2175
Published Online December 2020 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2020.912252

Hans汉斯

变系数异方差模型的贝叶斯分析

许芳忠，徐登可

浙江农林大学统计系，浙江 杭州
Email: 175384319@qq.com

收稿日期：2020年11月26日；录用日期：2020年12月11日；发布日期：2020年12月18日

摘　要

基于方差建模研究了变系数异方差模型的贝叶斯估计和异常点识别，其中非参数部分采用B样条逼近。主要通过应用Gibbs抽样和Metropolis-Hastings算法相结合的混合算法获得模型的贝叶斯估计和通过K-L距离贝叶斯诊断统计量来识别数据异常点。模拟研究显示所提出的贝叶斯分析方法是可行有效的。

关键词

异方差模型，Metropolis-Hastings算法，贝叶斯估计，K-L距离，B样条

Bayesian Analysis of Varying Coefficient Heteroscedastic Models

Fangzhong Xu, Dengke Xu

Department of Statistics, Zhejiang Agriculture and Forestry University, Hangzhou Zhejiang
Email: 175384319@qq.com

Received: Nov. 26th, 2020; accepted: Dec. 11th, 2020; published: Dec. 18th, 2020

Abstract

Based on variance modeling, Bayesian estimation and outlier identification of varying coefficient heteroscedastic models are studied, where the nonparametric part is approximated by B-spline. By combining the Gibbs sampler and Metropolis-Hastings algorithm, Bayesian estimation and Bayesian diagnosis statistics based on the K-L distance are obtained to identify outliers. Simulation studies show that the proposed Bayesian methods are feasible and effective.

Keywords

Heteroscedastic Models, Metropolis-Hastings Algorithm, Bayesian Estimation, K-L Distance, B-Spline

文章引用: 许芳忠, 徐登可. 变系数异方差模型的贝叶斯分析[J]. 应用数学进展, 2020, 9(12): 2166-2175.
DOI: 10.12677/aam.2020.912252

1. 引言

异方差数据常常出现在经济学、生物学、环境科学等领域，是很多统计学家研究的热点方向之一。其中处理此类数据最常用的方法是方差建模法，即不仅对均值建立回归模型，同时也对方差建立回归模型进行分析，有些文献称之为联合均值与方差模型。这个模型主要体现了对方差的重视，它能更好地解释数据变化的原因和规律。特别最近这些年已经有很多学者对基于方差建模的异方差模型研究了模型的参数估计、变量选择以及异方差检验等统计推断。例如，吴刘仓等[1]对联合均值与方差模型提出一种同时对均值模型和方差模型的变量选择方法；李双双等[2]研究了联合均值与方差混合专家回归模型的参数估计问题；赵远英等[3]对响应变量带有不可忽略缺失数据的联合均值与方差模型的贝叶斯估计问题进行了研究；戴琳等[4]基于联合均值与方差模型研究了模型的参数估计与基于数据删除模型考虑了统计诊断问题。其它类似的相关研究还可以具体参见文献[5] [6] [7] [8]。但是这些文献大多数都是基于参数异方差模型展开统计分析的，很少有文献和学者研究基于方差建模的变系数异方差模型的贝叶斯参数估计、异常点识别等统计推断问题。

因此本文针对方差建模的变系数异方差模型，应用 Gibbs 抽样和 Metropolis-Hastings 算法相结合的混合算法研究模型的贝叶斯估计和基于 K-L 距离研究贝叶斯数据删除影响诊断方法，以识别模型的异常点。

2. 模型与符号

2.1. 基于方差建模的变系数异方差模型

针对异方差数据和基于方差建模的思想，提出了如下变系数异方差模型：

$$\begin{cases} y_i = x_i^{\mathrm{T}}\beta + z_i^{\mathrm{T}}\alpha(u_i) + \varepsilon_i, \varepsilon_i \sim N(0,\sigma_i^2), \\ \sigma_i^2 = g(h_i^{\mathrm{T}}\gamma), \\ i = 1,2,\cdots,n. \end{cases} \tag{1}$$

其中 y_i 为响应变量且令 $Y=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$；$x_i=(x_{i1},\cdots,x_{ip})^{\mathrm{T}}$，$z_i=(z_{i1},\cdots,z_{iq})^{\mathrm{T}}$ 是与 y_i 的均值相关的解释变量的观测值，且 $\beta=(\beta_1,\cdots,\beta_p)^{\mathrm{T}}$ 是均值模型中一个 $p\times 1$ 维未知回归参数向量，$\alpha(u_i)=(\alpha_1(u_i),\alpha_2(u_i),\cdots,\alpha_d(u_i))^{\mathrm{T}}$ 是一个 $d\times 1$ 维的未知函数参数；$h_i=(h_{i1},\cdots,h_{iq})^{\mathrm{T}}$ 是与 y_i 的方差相关的解释变量的观测值，且 $\gamma=(\gamma_1,\cdots,\gamma_q)^{\mathrm{T}}$ 是方差模型中一个 $q\times 1$ 维的未知回归参数向量。另外，$g(\cdot)$ 是一个已知函数，为了模型的可识别性，一般假设 $g(\cdot)$ 是一个单调函数且 $g(\cdot)>0$。在本文中 $g(x)=\exp(x)$。

由模型(1)，可以获得如下似然函数

$$L(\beta,\gamma,\alpha(\cdot)\mid Y,X,Z,H,U)=(2\pi)^{-\frac{n}{2}}\prod_{i=1}^{n}(\sigma_i^2)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\frac{(y_i-x_i^{\mathrm{T}}\beta-z_i^{\mathrm{T}}\alpha(u_i))^2}{\sigma_i^2}\right) \tag{2}$$

其中 $X=(x_1^{\mathrm{T}},\cdots,x_n^{\mathrm{T}})^{\mathrm{T}}$，$Z=(z_1^{\mathrm{T}},\cdots,z_n^{\mathrm{T}})^{\mathrm{T}}$，$H=(h_1^{\mathrm{T}},\cdots,h_n^{\mathrm{T}})^{\mathrm{T}}$，$U=(u_1,\cdots,u_n)^{\mathrm{T}}$。因为 $\alpha(\cdot)$ 是非参数，(2)式还不能直接进行优化。因此，首先用 B 样条来逼近非参数函数 $\alpha(\cdot)$。具体如下，令 $0=s_0<s_1<\cdots<s_{k_n}<s_{k_n+1}=1$ 是[0,1]区间上的一个剖分。用 s_i 作为内节点，那么就有阶为 M 和维数为 $L=k_n+M$ 的正则化 B 样条基函数，这也形成了线性样条空间的一个基。节点选择一般是样条光滑估计中的一个重要方面。类似于文献

[9]，内节点的数目选取为 $n^{\frac{1}{5}}$ 的整数部分。这样，由 $\pi^{\mathrm{T}}(u)\lambda_k$ 逼近 $\alpha_k(u)$，其中 $\pi(u)=(\pi_1(u),\cdots,\pi_L(u))^{\mathrm{T}}$ 是基函数向量和 $\lambda_k\in R^L, k=1,2,\cdots,d$。利用这些符号，(1)式中的均值模型可以写成以下形式：

$$\mu_i=x_i^{\mathrm{T}}\beta+B_i^{\mathrm{T}}\lambda$$

其中 $B_i=I_d\otimes\pi(u_i)\cdot z_i, \lambda=(\lambda_1^{\mathrm{T}},\cdots,\lambda_d^{\mathrm{T}})^{\mathrm{T}}, K=dL$。

这样，似然函数(2)可以被重新写成以下形式：

$$L(\theta\mid Y,X,Z,H,U)=(2\pi)^{-\frac{n}{2}}\prod_{i=1}^{n}(\sigma_i^2)^{-\frac{1}{2}}\exp\left(\frac{1}{2}\sum_{i=1}^{n}\frac{(y_i-x_i^{\mathrm{T}}\beta-B_i^{\mathrm{T}}\lambda)^2}{\sigma_i^2}\right),\tag{3}$$

其中 $\theta=(\beta^{\mathrm{T}},\gamma^{\mathrm{T}},\lambda^{\mathrm{T}})^{\mathrm{T}}$。

2.2. K-L 距离

K-L 距离也称作 K-L 信息，具有距离和信息的某些性质，在统计上以反映两个模型或分布的差异而著称。根据韦博成[10]对 K-L 距离的介绍，密度函数为 $f(x)$ 与 $g(x)$ 的两个分布的 K-L 距离 $K(f,g)$ 被定义为：$K(f,g)=E_f\left\{\log\frac{f(x)}{g(x)}\right\}$，其中 E_f 表示对 $f(x)$ 求期望。

3. 贝叶斯分析

3.1. 先验分布

为了应用贝叶斯方法来估计模型(1)中的未知参数，需要给出未知参数的先验分布信息。为了简便，假设 β,γ 和 λ 相互独立且具有正态先验分布，分别为 $\beta\sim N(\beta_0,\Sigma_\beta)$，$\gamma\sim N(\gamma_0,\Sigma_\gamma)$，$\lambda\sim N(\lambda_0,\tau^2I_K)$，$\tau^2\sim IG(a_\tau,b_\tau)$，其中“IG”表示逆 gamma 分布，并且假设超参数 $\beta_0,\gamma_0,\lambda_0,\Sigma_\beta,\Sigma_\gamma$ 和 a_τ,b_τ 是已知的。

3.2. Gibbs 抽样和条件分布

基于式子(3)，按照以下过程用Gibbs抽样从后验分布 $p(\theta\mid Y,X,Z,H,U)$ 中进行抽样，其中 $\theta=(\beta^{\mathrm{T}},\gamma^{\mathrm{T}},\lambda^{\mathrm{T}})^{\mathrm{T}}$。

步骤 1. 令参数的初值 $\theta^{(0)}=(\beta^{(0)^{\mathrm{T}}},\gamma^{(0)^{\mathrm{T}}},\lambda^{(0)^{\mathrm{T}}})^{\mathrm{T}}$。

步骤 2. 基于 $\theta^{(l)}=(\beta^{(l)^{\mathrm{T}}},\gamma^{(l)^{\mathrm{T}}},\lambda^{(l)^{\mathrm{T}}})^{\mathrm{T}}$，计算 $\Sigma^{(l)}=diag\{\sigma_1^{2(l)},\cdots,\sigma_n^{2(l)}\}$。

步骤 3. 基于 $\theta^{(l)}=(\beta^{(l)^{\mathrm{T}}},\gamma^{(l)^{\mathrm{T}}},\lambda^{(l)^{\mathrm{T}}})^{\mathrm{T}}$ 按照以下抽取 $\theta^{(l+1)}=(\beta^{(l+1)^{\mathrm{T}}},\gamma^{(l+1)^{\mathrm{T}}},\lambda^{(l+1)^{\mathrm{T}}})^{\mathrm{T}}$；

- 抽样 $\tau^{2(l+1)}$：

$$p(\tau^2\mid\lambda)\propto(\tau^2)^{-\frac{K}{2}-a_\tau-1}\exp\left\{\frac{-\frac{1}{2}(\lambda-\lambda_0)^{\mathrm{T}}(\lambda-\lambda_0)-b_\tau}{\tau^2}\right\}\tag{4}$$

- 抽样 $\lambda^{(l+1)}$：

$$p(\lambda\mid Y,X,Z,H,U,\beta,\gamma)\propto\exp\left\{-\frac{1}{2}(\lambda-\lambda^*)^{\mathrm{T}}\Sigma_\lambda^{*-1}(\lambda-\lambda^*)\right\},\tag{5}$$

其中 $\lambda^{*}=\Sigma_{\lambda}^{*}\left(\Omega^{\mathrm{T}}\Sigma^{-1}\left(Y-X\beta\right)+\tau^{-2}I_{K}\lambda_{0}\right)$，$\Sigma_{\lambda}^{*}=\left(\tau^{-2}I_{K}+\Omega^{\mathrm{T}}\Sigma^{-1}\Omega\right)^{-1}$，$\Omega=\left(B_{1}^{\mathrm{T}},\cdots,B_{n}^{\mathrm{T}}\right)^{\mathrm{T}}$。

- 抽样 $\beta^{(l+1)}$：

$$p\left(\beta\mid Y,X,Z,H,U,\gamma,\lambda\right)\propto\exp\left\{-\frac{1}{2}\left(\beta-b^{*}\right)^{\mathrm{T}}B^{*-1}\left(\beta-b^{*}\right)\right\},\tag{6}$$

其中 $b^{*}=B^{*}\left(X^{T}\Sigma^{-1}\left(Y-\Omega\lambda\right)+\Sigma_{\beta}^{-1}\beta_{0}\right)$，$B^{*}=\left(\Sigma_{\beta}^{-1}+X^{\mathrm{T}}\Sigma^{-1}X\right)^{-1}$。

- 抽样 $\gamma^{(l+1)}$：

$$p\left(\gamma\mid Y,X,Z,H,U,\beta,\lambda\right)\propto\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}h_{i}^{\mathrm{T}}\gamma-\frac{1}{2}\sum_{i=1}^{n}\frac{\left(y_{i}-x_{i}^{\mathrm{T}}\beta-B_{i}^{\mathrm{T}}\lambda\right)^{2}}{\exp\left(h_{i}^{\mathrm{T}}\gamma\right)}-\frac{1}{2}\left(\gamma-\gamma_{0}\right)^{\mathrm{T}}\Sigma_{\gamma}^{-1}\left(\gamma-\gamma_{0}\right)\right\}\tag{7}$$

步骤4。重复步骤2和3。

这样就通过以上算法产生了样本序列$\left(\beta^{(l)},\gamma^{(l)},\lambda^{(l)},\tau^{2(l)}\right),l=1,2,\cdots$。从(4)~(7)式中很容易发现，条件分布 $p\left(\tau^{2}\mid\lambda\right),p\left(\beta\mid Y,X,Z,H,U,\gamma,\lambda\right),p\left(\lambda\mid Y,X,Z,H,U,\beta,\gamma\right)$是熟悉的正态分布和逆 gamma 分布。从这两个分布抽取随机数是比较容易的。但是条件分布 $p\left(\gamma\mid Y,X,Z,H,U,\beta,\lambda\right)$是一不规则且相当复杂的分布，如何从这个分布中抽取随机数有点困难。在这里主要应用 MH 算法抽取随机数。选择正态分布 $N\left(\gamma^{(l)},\sigma_{\gamma}^{2}\Omega_{\gamma}^{-1}\right)$作为建议分布，其中通过选择 σ_{γ}^{2}，来使得接受概率在 0.25 与 0.45 之间，且取 $\Omega_{\gamma}=\Sigma_{\gamma}^{-1}+\frac{1}{2}\sum_{i=1}^{n}\frac{\left(y_{i}-x_{i}^{\mathrm{T}}\beta-B_{i}^{\mathrm{T}}\lambda\right)^{2}}{\exp\left(h_{i}^{\mathrm{T}}\gamma\right)}h_{i}h_{i}^{\mathrm{T}}$。

3.3. 贝叶斯估计

利用以上提出的计算过程来产生观测值来获得参数 β,γ 和 λ 的贝叶斯估计。令 $\left\{\theta^{(j)}=\left(\beta^{(j)},\gamma^{(j)},\lambda^{(j)}\right):j=1,2,\cdots,J\right\}$是通过上述混合算法从联合条件分布 $p\left(\beta,\gamma,\lambda\mid Y,X,Z,H,U\right)$中产生的观测值，那么 β,γ 和 λ 的贝叶斯估计为：

$$\hat{\beta}=\frac{1}{J}\sum_{j=1}^{J}\beta^{(j)},\hat{\gamma}=\frac{1}{J}\sum_{j=1}^{J}\gamma^{(j)},\hat{\lambda}=\frac{1}{J}\sum_{j=1}^{J}\lambda^{(j)}。$$

根据文献[11]有，当 J 趋于无穷时，$\hat{\theta}=\left(\hat{\beta},\hat{\gamma},\hat{\lambda}\right)$是对应后验均值向量的相合估计。类似地，后验协方差矩阵 $Var\left(\theta\mid Y,X,Z,H,U\right)$的相合估计可以通过观测$\left\{\theta^{(j)}:j=1,2,\cdots,J\right\}$的样本协方差矩阵来获得，即

$$Var\left(\theta\mid Y,X,Z,H,U\right)=\left(J-1\right)^{-1}\sum_{j=1}^{J}\left(\theta^{(j)}-\hat{\theta}\right)\left(\theta^{(j)}-\hat{\theta}\right)^{\mathrm{T}}。$$

这样，后验标准误就可以通过该矩阵的对角元素来获得。

3.4. 贝叶斯诊断

在贝叶斯统计诊断分析中已存在许多诊断统计量用来评价个体观测对参数后验分布的影响，在这主要基于 K-L 距离研究贝叶斯数据删除影响的统计诊断方法。对任意的 $i=1,\cdots,n$，记$\left\{y_{i},x_{i},z_{i},h_{i},u_{i}\right\}$是第 i 个个体观测数据点，$D=\left\{Y,X,Z,H,U\right\}$为完全数据集，$D_{-i}$为完全数据集 D 删除第 i 个个体观测数据点得到的数据集，$L\left(\theta\mid D\right)$与$L\left(\theta\mid D_{-i}\right)$分别表示基于数据 D 与 D_{-i} 的似然函数，则 θ 于数据 D 与 D_{-i} 的后

验分布分别为 $p(\theta|D)\propto L(\theta|D)p(\theta)$， $p(\theta|D_{-i})\propto L(\theta|D_{-i})p(\theta)$。根据 Cho 等[7]的讨论，定义 K-L 距离为：

$$K(P,P_{-i})=\int p(\theta|D)\log\left\{\frac{p(\theta|D)}{p(\theta|D_{-i})}\right\}\mathrm{d}\theta, \tag{8}$$

其中 P 与 P_{-i} 分别表示 θ 基于数据 D 与 D_{-i} 的后验分布，注意到 K-L 距离 $K(P,P_{-i})$ 是完全数据集 D 删除第 i 个数据点前后对参数 θ 后验分布影响的一种很好的度量。经过简单的计算(8)式变为：

$$K(P,P_{-i})=\log E_{\theta}\left[\frac{L(\theta|D_{-i})}{L(\theta|D)}|D\right]+E_{\theta}\left[\log\frac{L(\theta|D)}{L(\theta|D_{-i})}|D\right], \tag{9}$$

其中 $E_{\theta}[.|D]$ 表示 θ 基于数据 D 的后验期望。由 Gibbs 抽样算法抽取的随机观测序 $\left\{\theta^{(j)}:j=1,2,\cdots,J\right\}$，可以得到 K-L 距离 $K(P,P_{-i})$ 的估计为：

$$K(P,P_{-i})=\log\left[\frac{1}{J}\sum_{j=1}^{J}\frac{L(\theta^{(j)}|D_{-i})}{L(\theta^{(j)}|D)}\right]+\frac{1}{J}\sum_{j=1}^{J}\log\left[\frac{L(\theta^{(j)}|D)}{L(\theta^{(j)}|D_{-i})}\right]。 \tag{10}$$

对任意的 $i=1,\cdots,n$，当 $K(P,P_{-i})$ 很大时，可以诊断第 i 个个体观测数据点为异常点。

4. 模拟研究

在这通过模拟研究来说明本文所提出的贝叶斯分析方法的有效性。选择均值模型为 $\mu_i=x_i^{\mathrm{T}}\beta+z_i^{\mathrm{T}}\alpha(u_i)$ 和方差模型为 $\sigma_i^2=\exp(h_i^{\mathrm{T}}\gamma)$，其中均值参数和方差参数的真实值分别为 $\beta=(1,-0.5,0.5)^{\mathrm{T}}$ 和 $\gamma=(1,-0.5,0.5)^{\mathrm{T}}$，另外 $d=2$，$\alpha_1(u_i)=0.5\sin(2\pi u_i)$，$\alpha_2(u_i)=8u_i(1-u_i^2)$；$x_i,z_i,h_i$ 分别是 $3\times1,2\times1$ 和 3×1 的协变量向量，其中的元素产生于均值为零，协方差矩阵为 $\Sigma=\left(\sigma_{ij}=0.5^{|i-j|}\right)$ 的正态分布。这样响应变量 Y_i 就可以从多元正态分布 $N(\mu_i,\Sigma_i)(i=1,\cdots,n)$ 中产生。

为了调查贝叶斯分析方法对先验分布的敏感程度，考虑以下三种有关未知参数 β,γ,λ 的先验分布中超参数值的设置的情形：

TypeI：

$\beta_0=(1,-0.5,0.5)^{\mathrm{T}},\Sigma_{\beta}=0.25\times\mathrm{I}_3,\gamma_0=(1,-0.5,0.5)^{\mathrm{T}},\Sigma_{\gamma}=0.25\times\mathrm{I}_3,\lambda_0=(0,\cdots,0)^{\mathrm{T}},a_{\tau}=b_{\tau}=1$。这种设置表示具有很好的先验信息。

TypeII：

$\beta_0=3\times(1,-0.5,0.5)^{\mathrm{T}},\Sigma_{\beta}=\mathrm{I}_3,\gamma_0=3\times(1,-0.5,0.5)^{\mathrm{T}},\Sigma_{\gamma}=\mathrm{I}_3,\lambda_0=(0,\cdots,0)^{\mathrm{T}},a_{\tau}=b_{\tau}=1$。这种设置表示具有较差的先验信息。

TypeIII：

$\beta_0=(0,0,0)^{\mathrm{T}},\Sigma_{\beta}=10\times\mathrm{I}_3,\gamma_0=(0,0,0)^{\mathrm{T}},\Sigma_{\gamma}=10\times\mathrm{I}_3,\lambda_0=(0,\cdots,0)^{\mathrm{T}},a_{\tau}=b_{\tau}=1$。这些超参数值的设置代表的是没有先验信息的情况。

在上面的各种情形下，应用联合 Gibbs 抽样和 Metropolis-Hastings 算法的混合算法来计算未知参数和光滑函数的贝叶斯估计。在模拟中分别令样本量 n = 80 和 n = 150。对于每一种情形，重复计算 100 次。对于每次重复产生的每一次数据集，MCMC 算法的收敛性可以通过 EPSR 值来检验，并且在每次运行中观测得到在 3000 次迭代以后 EPSR 值都小于 1.2。因此在每次重复计算中丢掉前 3000 次迭代以后再收集 J = 5000 个样本来产生贝叶斯估计。参数贝叶斯估计的模拟结果概括在表 1 中。为了调查估计函数 $\alpha_1(u)$

和 $\alpha_2(u)$ 的精确度，画出了在不同样本量和三种先验分布情形下非参数函数的平均估计曲线和真实曲线，并且展示在图 1~6 中。

在表 1 中，“Bias”表示基于 100 次重复计算未知参数的贝叶斯估计与真值的偏差的绝对值，“SD”表示前面给出的后验标准误的平均估计，和“RMS”表示的是基于 100 次重复计算的贝叶斯估计的均方误差的算术平方根。从表 1 中可以获得 i) 在估计的偏差、RMS 和 SD 值方面，不管何种先验信息贝叶斯估计都相当精确；ii) 当样本量逐渐变大时，估计也变得越来越好。从图 1~6 中展示了不管何种先验信息，估计出来的非参数函数的曲线与相应的真实函数的曲线逼近得都比较好。总之，从以上结果可以看出本文所提出的贝叶斯估计方法能很好地恢复变系数异方差模型中的真实信息。

Table 1. Bayesian estimation of model parameters under different sample sizes and prior distributions
表 1. 不同的样本量和先验分布下模型参数的贝叶斯估计结果

Type	参数	n = 80			n = 150		
		Bias	SD	RMS	Bias	SD	RMS
I	β_1	0.0027	0.1184	0.1160	0.0020	0.0803	0.0759
	β_2	0.0028	0.1310	0.1224	0.0061	0.0912	0.0880
	β_3	0.0019	0.1181	0.1060	0.0051	0.0804	0.0776
	γ_1	0.0208	0.2075	0.1884	0.0105	0.1453	0.1483
	γ_2	0.0301	0.2193	0.1964	0.0194	0.1569	0.1517
	γ_3	0.0026	0.2025	0.1799	0.0194	0.1418	0.1365
II	β_1	0.0466	0.1144	0.1265	0.0143	0.0797	0.0790
	β_2	0.0341	0.1238	0.1301	0.0107	0.0913	0.0910
	β_3	0.0182	0.1144	0.1075	0.0047	0.0804	0.0793
	γ_1	0.1612	0.2350	0.2685	0.0700	0.1530	0.1722
	γ_2	0.1906	0.2586	0.3032	0.0806	0.1671	0.1858
	γ_3	0.0919	0.2349	0.2537	0.0119	0.1489	0.1454
III	β_1	0.0125	0.1196	0.1179	0.0025	0.0814	0.0784
	β_2	0.0024	0.1303	0.1300	0.0084	0.0935	0.0914
	β_3	0.0016	0.1195	0.1116	0.0059	0.0823	0.0800
	γ_1	0.0219	0.2372	0.2303	0.0090	0.1534	0.1611
	γ_2	0.0363	0.2668	0.2516	0.0161	0.1686	0.1774
	γ_3	0.0014	0.2401	0.2410	0.0220	0.1499	0.1490

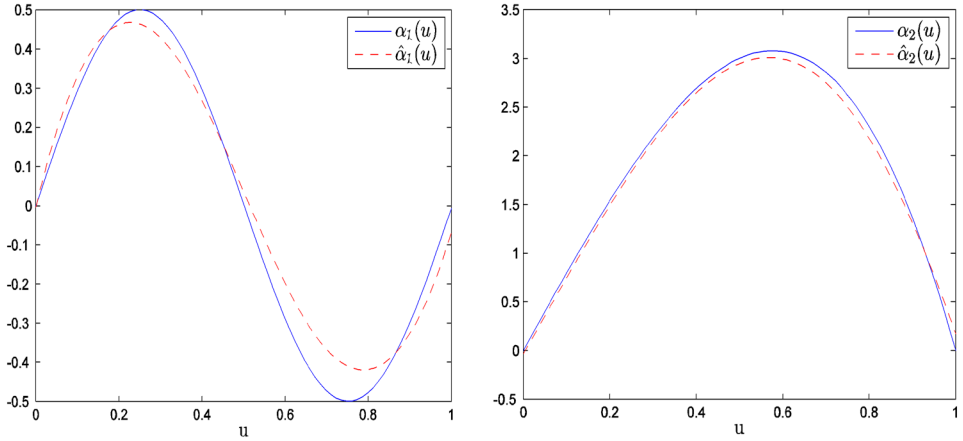

Figure 1. The average estimated curves of nonparametric parts $\alpha_1(u)$ and $\alpha_2(u)$ when $n = 80$ and prior information of Type I

图 1. 当 $n = 80$ 和 Type I 先验信息下非参数部分 $\alpha_1(u)$ 和 $\alpha_2(u)$ 的平均估计曲线

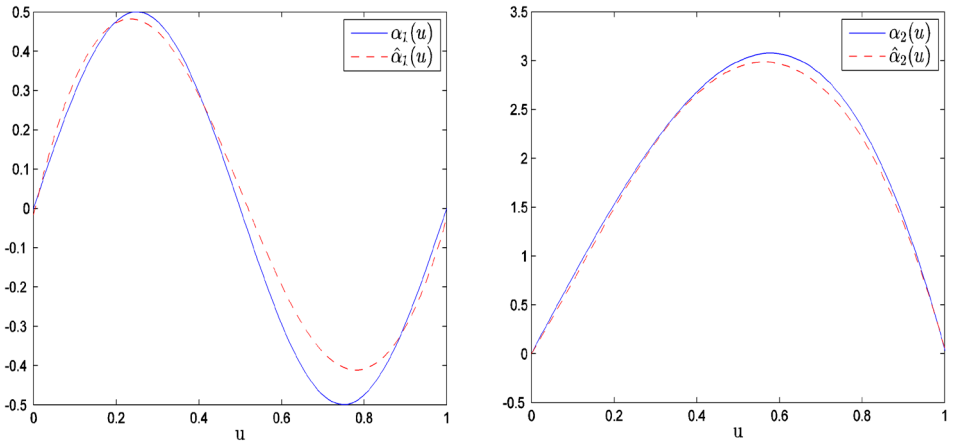

Figure 2. The average estimated curves of nonparametric parts $\alpha_1(u)$ and $\alpha_2(u)$ when $n = 80$ and prior information of Type II

图 2. 当 $n = 80$ 和 Type II 先验信息下非参数部分 $\alpha_1(u)$ 和 $\alpha_2(u)$ 的平均估计曲线

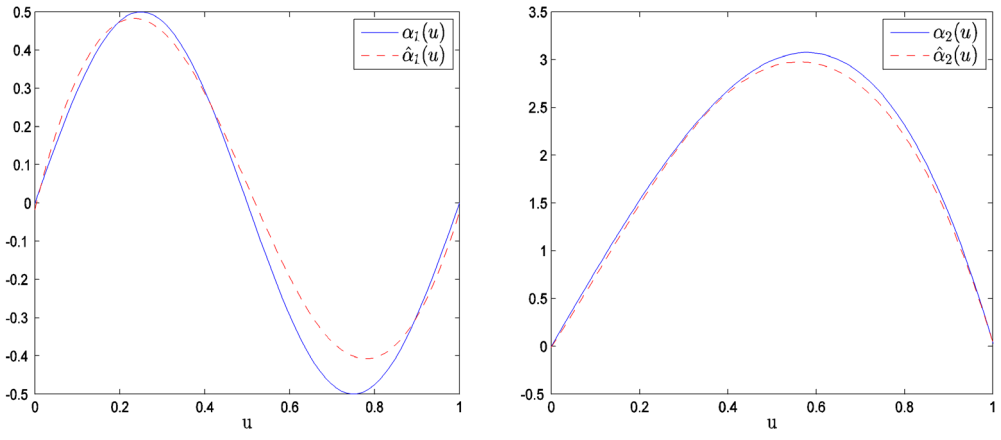

Figure 3. The average estimated curves of nonparametric parts $\alpha_1(u)$ and $\alpha_2(u)$ when $n = 80$ and prior information of Type III

图 3. 当 $n = 80$ 和 Type III 先验信息下非参数部分 $\alpha_1(u)$ 和 $\alpha_2(u)$ 的平均估计曲线

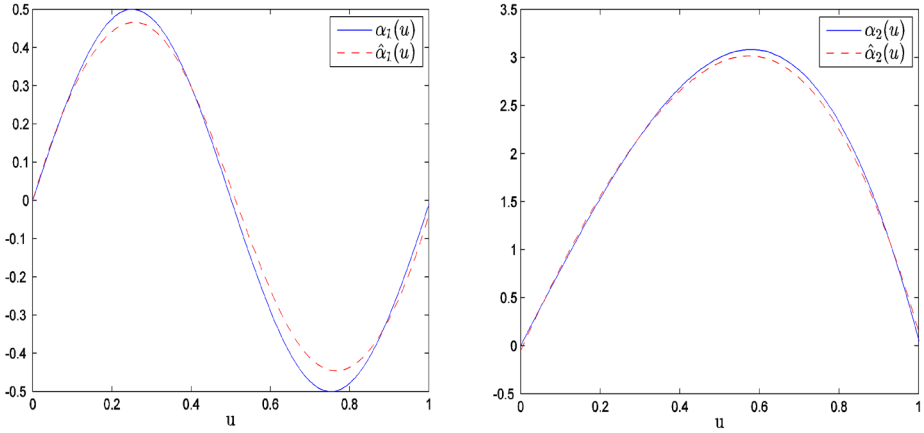

Figure 4. The average estimated curves of nonparametric parts $\alpha_1(u)$ and $\alpha_2(u)$ when n = 150 and prior information of Type I

图 4. 当 n = 150 和 Type I 先验信息下非参数部分 $\alpha_1(u)$ 和 $\alpha_2(u)$ 的平均估计曲线

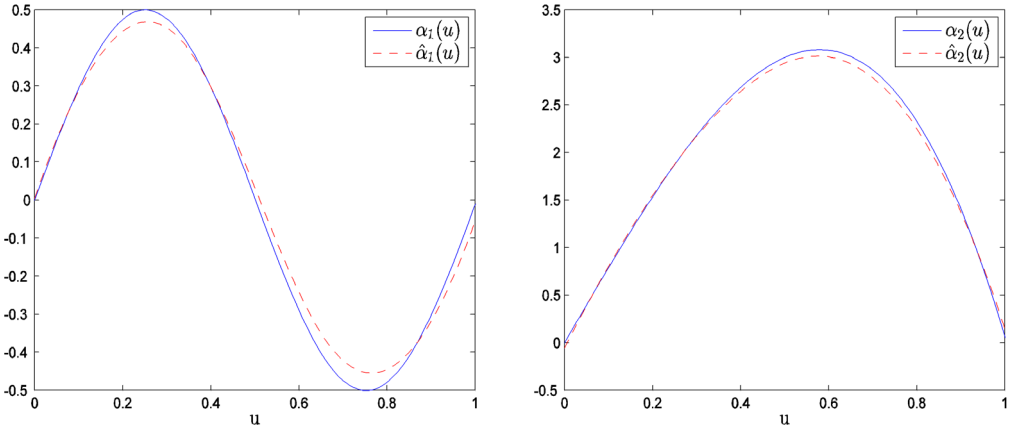

Figure 5. The average estimated curves of nonparametric parts $\alpha_1(u)$ and $\alpha_2(u)$ when n = 150 and prior information of Type II

图 5. 当 n = 150 和 Type II 先验信息下非参数部分 $\alpha_1(u)$ 和 $\alpha_2(u)$ 的平均估计曲线

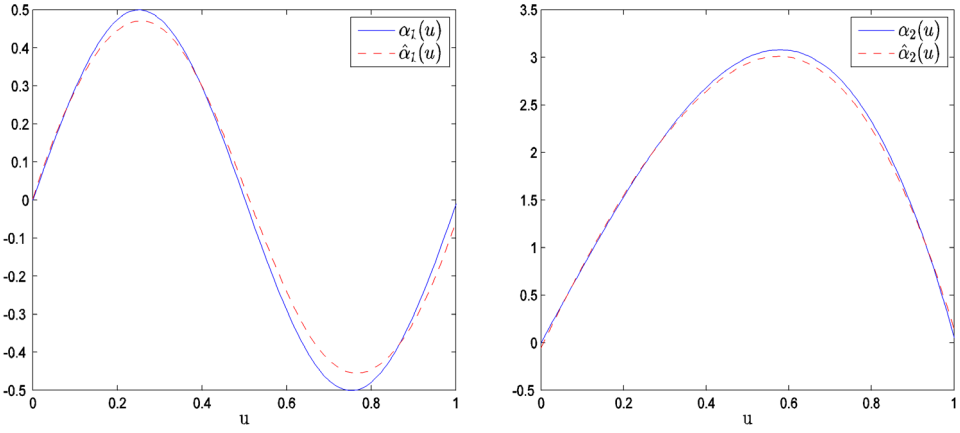

Figure 6. The average estimated curves of nonparametric parts $\alpha_1(u)$ and $\alpha_2(u)$ when n = 150 and prior information of Type III

图 6. 当 n = 150 和 Type III 先验信息下非参数部分 $\alpha_1(u)$ 和 $\alpha_2(u)$ 的平均估计曲线

最后为了检测根据 K-L 距离来识别异常点的效果，在第 10 个个体观测数据点的响应变量加 10 构成人工数据集 D。然后对人工数据集 D 应用本文介绍的贝叶斯影响诊断方法来检测影响观测。其中 MCMC 算法的收敛性可以通过 EPSR 值来检验，并且发现在 3000 次迭代以后 EPSR 值都小于 1.2。因此在计算中丢掉前 3000 次迭代以后再收集 J = 5000 个随机样本来通过(10)式计算 K-L 距离 $K\left(P,P_{-i}\right)$。图 7 和图 8 报道了相应的诊断结果。正如预期的一样，通过图 7 和图 8 很容易就发现，第 10 个个体观测数据点被诊断为异常点，且诊断方法对先验分布超参数取值的选取不是很敏感。

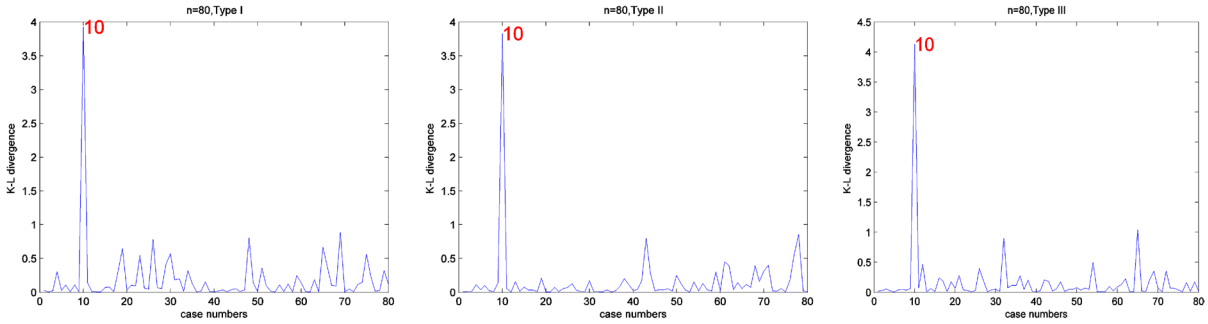

Figure 7. Bayesian case deletion diagnosis results based on different prior information and n = 80
图 7. 当 n = 80 时，基于不同的先验信息下贝叶斯数据删除影响诊断的数值结果

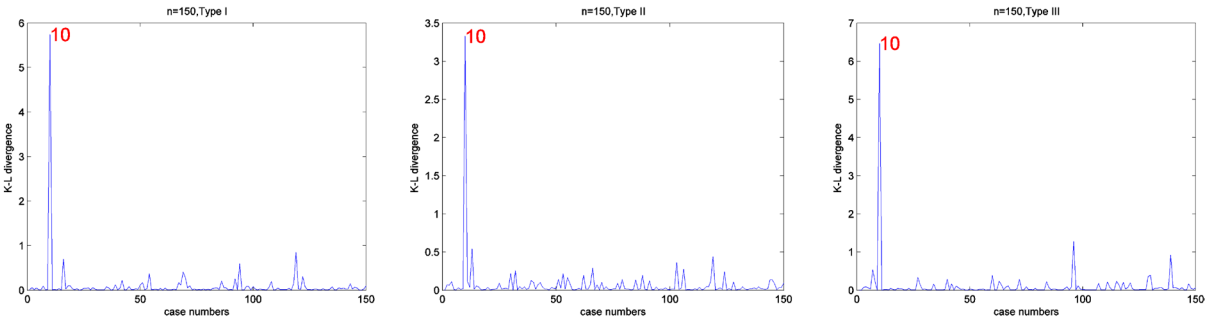

Figure 8. Bayesian case deletion diagnosis results based on different prior information and n = 150
图 8. 当 n = 150 时，基于不同的先验信息下贝叶斯数据删除影响诊断的数值结果

5. 结论

本文针对变系数异方差模型，基于 Gibbs 抽样和 MH 算法相结合的混合算法，以及根据 K-L 距离研究模型的贝叶斯估计和贝叶斯统计诊断方法。模拟研究显示了模型与贝叶斯方法的可行性和有效性。

基金项目

浙江省高校重大人文社科攻关计划项目资助(2018QN037)。

参考文献

[1] 吴刘仓，张忠占，徐登可. 联合均值与方差模型的变量选择[J]. 系统工程理论与实践, 2012(8): 1754-1760.

[2] 李双双，吴刘仓，戴琳. 联合均值与方差混合专家回归模型的参数估计[J]. 应用数学, 2019, 32(1): 134-140.

[3] 赵远英，吴刘仓，徐登可. 带有不可忽略缺失数据的联合均值与方差模型的贝叶斯估计[J]. 昆明理工大学学报(自然科学版), 2020, 45(1): 125-132.

[4] 戴琳，陶治，吴刘仓. 联合均值与方差模型的统计诊断[J]. 统计与信息论坛, 2017, 32(1): 14-19.

[5] 赵远英，徐登可，庞一成. 联合均值与方差模型的 Bayes 分析[J]. 高校应用数学学报, 2018, 33(2): 241-252.

[6] Cook, R.D. (1977) Detection of Influential Observations in Linear Regression. *Technometrics*, **19**, 15-18. https://doi.org/10.1080/00401706.1977.10489493

[7] Cho, H., Ibrahim, J.G., Sinha, D. and Zhu, H.T. (2009) Bayesian Case Influence Diagnostics for Survival Models. *Biometrics*, **65**, 116-124. https://doi.org/10.1111/j.1541-0420.2008.01037.x

[8] Tang, N.S. and Duan, X.D. (2012) A Semiparametric Bayesian Approach to Generalized Partial Linear Mixed Models for Longitudinal Data. *Computational Statistics & Data Analysis*, **56**, 4348-4365. https://doi.org/10.1016/j.csda.2012.03.018

[9] He, X.M. and Shi, P. (1994) Convergence Rate of B-Spline Estimators of Nonparametric Conditional Quantile Function. *Journal of Nonparametric Statistics*, **3**, 299-308. https://doi.org/10.1080/10485259408832589

[10] 韦博成. 参数统计教程[M]. 北京: 高等教育出版社, 2006.

[11] Geyer, C.J. (1992) Practical Markov Chain Monte Carlo. *Statistical Science*, **7**, 473-511. https://doi.org/10.1214/ss/1177011137