DOI:10.13718/j.cnki.xsxb.2023.06.009

基于指数平方损失的纵向多折点回归模型的稳健估计与统计推断①

唐铭，　李婷婷

西南大学 数学与统计学院，重庆 400715

摘要：本文基于指数平方损失函数研究纵向多折点回归模型的稳健估计与统计推断问题. 为提高参数估计方法的效率，基于局部线性平滑方法和修正的 Cholesky 分解方法提出纵向多折点回归模型参数估计的迭代算法，研究了参数估计的渐近正态性质，同时讨论了指数平方损失函数中关键调谐参数的选择、模型中折点个数的确定方法和折线效应的检验问题等，数值模拟展示了本文所提方法的有限样本表现.

关　键　词：纵向数据；多折点模型；Cholesky 分解；局部线性平滑；指数平方损失

中图分类号：O212.1　　**文献标志码：A**　　**文章编号：**1000-5471(2023)06-0059-11

Robust Estimation and Statistical Inference for Longitudinal Multi-Kink Regression Model Based on Exponential Square Loss

TANG Ming，　LI Tingting

School of Mathematics and Statistics，Southwest University，Chongqing 400715，China

Abstract: In this article, we investigate the robust parameter estimation and statistical inference of the longitudinal multi-kink regression model based on the exponential square loss function. A procedure based on local linear smoothing technique and modified Cholesky decomposition is proposed to improve the estimation efficiency of parameters, and the asymptotic normality are established under some mild conditions. Furthermore, we propose a data-driven procedure to automatically selecting the additional tuning parameter in exponential square loss function. In addition, a weighted cumulative sum type statistic for testing the existence of a kink-point, and a modified Bayesian information criterion for estimating the number of kink-points are developed. Finally, simulation studies show the finite sample performance of the proposed methods.

Key words: longitudinal data; multi-kink model; Cholesky decomposition; local linear smoothing; exponential square loss

折点回归模型主要研究响应变量与协变量间的分阶段连续变化特征，在金融、经济、工业、医学等领域有着重要应用. 文献[1]首次提出单折点回归模型，并基于似然函数提出折点参数估计的网格搜索法. 文

① 收稿日期：2022-09-13
作者简介：唐铭，硕士研究生，主要从事回归模型参数估计研究.
通信作者：李婷婷，副教授.

献[2]基于累积和统计量及似然比统计量提出折点效应的假设检验方法. 值得一提的是，文献[1]所提的网格搜索法虽然可以生成合理的估计，但计算成本却很高. 文献[3]基于泰勒展开提出了一种局部线性平滑的参数估计方法，在保证估计准确性的同时极大提高计算效率. 文献[4]首次将单折点回归模型拓展到多折点回归模型，基于局部线性平滑法提出了分位数损失函数下的参数估计、变点存在性检验统计量，及折点个数确定的贝叶斯信息准则，并研究了所提估计及统计量的大样本性质.

以上文献的研究大多讨论的是独立同分布数据的折点模型. 然而，随着应用领域的不断拓展，所处理的数据类型越来越复杂，纵向数据便是复杂的数据类型之一. 针对纵向折点回归模型，文献[5]在独立工作矩阵下考虑了纵向单折点分位数回归模型的估计与检验；文献[6]考虑了纵向数据的多折点分位数回归模型. 为融合纵向数据同一个个体内部的相关性，文献[6]基于文献[7]提出的二次推断函数方法(quadratic inference function，QIF)研究了相关结构下纵向多折点分位数回归模型的估计与统计推断，但其所能刻画的仍是等相关和AR(1)等特殊结构的矩阵. 文献[8]提出修正的Cholesky分解方法，该方法不局限于特殊相关结构，且能保证估计的协方差阵的正定性，具有更广泛的适用性.

众所周知，基于经典平方损失的估计对异常值非常敏感. 为处理包含大量异常值的数据，众多稳健估计的方法被相继提出，如Huber损失函数法[9]、秩回归[10]及分位数回归方法[11]等. 文献[12]提出一种新的稳健估计方法，即基于指数平方损失函数的参数估计方法. 该方法的显著特征是引入一个额外的调谐参数，可通过选择合适的调谐参数实现模型参数的自适应稳健估计. 文献[13-15]关于指数平方损失函数的相关研究均表明，基于该损失的参数估计相对于经典的稳健方法有着更好的表现，能够获得更好的鲁棒性和有效性.

本文基于指数平方损失和修正的Cholesky分解方法研究纵向多折点回归模型的参数估计及统计推断.

1 模型与估计算法

1.1 纵向多折点回归模型

考虑折点个数 K 及折点位置 $\boldsymbol{\tau}=(\tau_1,\cdots,\tau_K)^{\mathrm{T}}$ 均未知的纵向多折点回归模型：

$$Y_{ij}=a_0+a_1X_{ij}+\sum_{k=1}^{K}b_k(X_{ij}-\tau_k)_{+}+\boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\beta}+e_{ij},\ i=1,\cdots,n,\ j=1,\cdots,m_i \tag{1}$$

其中：n 为个体数，m_i 为第 i 个个体的重复观测次数，$(a)_{+}=a\cdot I(a\geqslant 0)$，$I$ 为示性函数，Y_{ij} 为响应变量观测值，$\boldsymbol{Z}_{ij}$ 为 p 维协变量观测值，X_{ij} 为有界门限变量，e_{ij} 为随机误差，记 $\boldsymbol{e}_i=(e_{i1},\cdots,e_{im_i})^{\mathrm{T}}$. 由模型(1)可知，变量 X_{ij} 的回归系数在折点 $\boldsymbol{\tau}=(\tau_1,\cdots,\tau_K)^{\mathrm{T}}$ 处会发生变化. 记 $\boldsymbol{b}=(b_1,\cdots,b_K)^{\mathrm{T}}$，$\boldsymbol{\vartheta}=(a_0,a_1,\boldsymbol{b}^{\mathrm{T}},\boldsymbol{\beta}^{\mathrm{T}})^{\mathrm{T}}$，以及参数向量 $\boldsymbol{\theta}=(\boldsymbol{\vartheta}^{\mathrm{T}},\boldsymbol{\tau}^{\mathrm{T}})^{\mathrm{T}}$，其中 $\boldsymbol{\vartheta}\in G\subset\mathbb{R}^{2+K+p}$，$\boldsymbol{\tau}\in T\subset D^K$. G，T 均为紧集，D 为 X_{ij} 的支撑集. 下面，首先对假设折点个数 K 已知后的参数估计算法进行介绍，随后给出用于确定折点个数 K 的模型选择方法.

1.2 基于指数平方损失的纵向多折点回归模型的参数估计算法

在模型(1)中，对于 $k=1,\cdots,K$，$b_k(X_{ij}-\tau_k)_{+}$ 关于 τ_k 不可导，使用文献[3]所提的基于局部线性平滑的快速迭代法，将 $b_k(X_{ij}-\tau_k)_{+}$ 在与折点真实位置相近的 $\tau_k^{(0)}$ 附近进行一阶泰勒展开

$$b_k(X_{ij}-\tau_k)_{+}\approx b_k(X_{ij}-\tau_k^{(0)})_{+}-b_k(\tau_k-\tau_k^{(0)})I(X_{ij}>\tau_k^{(0)}) \tag{2}$$

进而有如下的近似回归模型

$$Y_{ij}=a_0+a_1X_{ij}+\sum_{k=1}^{K}b_k(X_{ij}-\tau_k^{(0)})_{+}+\sum_{k=1}^{K}\mu_k\{-I(X_{ij}>\tau_k^{(0)}\}+\boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\beta}+\tilde{e}_{ij}^{(0)} \tag{3}$$

其中 $\mu_k=b_k(\tau_k-\tau_k^{(0)})$. 记 $\boldsymbol{\mu}=(\mu_1,\cdots,\mu_K)^{\mathrm{T}}$，$\boldsymbol{\eta}=(\boldsymbol{\vartheta}^{\mathrm{T}},\boldsymbol{\mu}^{\mathrm{T}})^{\mathrm{T}}$ 为未知系数向量. 为便于表达，记 $\tilde{\boldsymbol{\chi}}_{ij}=(1,X_{ij},(X_{ij}-\tau_1^{(0)})_{+},\cdots,(X_{ij}-\tau_k^{(0)})_{+},\boldsymbol{Z}_{ij}{}^{\mathrm{T}},-I(X_{ij}>\tau_1^{(0)}),\cdots,-I(X_{ij}>\tau_K^{(0)}))^{\mathrm{T}}$，$\tilde{\boldsymbol{\chi}}_i=(\tilde{\boldsymbol{\chi}}_{i1},\cdots,\tilde{\boldsymbol{\chi}}_{im_i})^{\mathrm{T}}$，$\tilde{\boldsymbol{e}}_i=(\tilde{e}_{i1}^{(0)},\cdots,\tilde{e}_{im_i}^{(0)})^{\mathrm{T}}$. 基于文献[12]，本文基于指数平方损失函数对模型(2)的未知参数进行估计，即关于 $\boldsymbol{\eta}$ 最小化如下目标函数

$$S_n(\boldsymbol{\eta})=\sum_{i=1}^{n}\sum_{j=1}^{m_i}\varphi_{\gamma}(Y_{ij}-\tilde{\boldsymbol{\chi}}_{ij}^{\mathrm{T}}\boldsymbol{\eta}) \tag{4}$$

其中 $\varphi_\gamma(x)=1-\exp\left(-\frac{x^2}{\gamma}\right)$ 为指数平方损失函数，γ 称为调谐参数. 不难发现，当 γ 取值较小时，可以为大的绝对值残差赋予较小的权重，从而降低异常值对参数估计的影响. 当 $\gamma\to\infty$ 时该损失函数与最小二乘的平方损失函数近似. 因此在指数平方损失函数中，可以通过选择合适的 γ 实现模型参数的自适应稳健估计. 考虑纵向数据个体重复观测之间的相关性，结合(4)式，为提高参数估计效率，本文提出求解如下广义估计方程

$$\sum_{i=1}^{n}\tilde{\boldsymbol{\chi}}_i^{\mathrm{T}}\boldsymbol{V}_i^{-1}\psi_\gamma(\boldsymbol{Y}_i-\tilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})=0 \tag{5}$$

以获得未知参数向量 $\boldsymbol{\eta}$ 的估计. 其中 $\boldsymbol{Y}_i=(Y_{i1},\cdots,Y_{im_i})^{\mathrm{T}}$，$\psi_\gamma(x)=\frac{2x}{\gamma}\exp\left(-\frac{x^2}{\gamma}\right)$. $\boldsymbol{V}_i$ 用于刻画个体残差之间的相关性，理论上最优的矩阵$\boldsymbol{V}_i=\mathrm{Cov}(\psi_\gamma(\boldsymbol{Y}_i-\tilde{\boldsymbol{\chi}}_i\boldsymbol{\eta}))$. 然而由于个体内部相关性无法观测，$\boldsymbol{V}_i$ 的具体形式亦难以确定. 文献[7] 提出 QIF 法，该方法使用基矩阵的线性形式替代$\boldsymbol{V}_i^{-1}$，但所能刻画的仍是等相关和 AR(1) 等一些具有特殊结构的矩阵. 本文使用文献[8] 提出的修正 Cholesky 分解方法，该方法不局限于特殊相关结构，具有更广泛的适用性. 具体来说，同文献[15] 类似，将$\boldsymbol{V}_i$ 分解为如下的矩阵乘积形式

$$\boldsymbol{\Phi}_i\boldsymbol{V}_i\boldsymbol{\Phi}_i^{\mathrm{T}}=\boldsymbol{D}_i \tag{6}$$

其中 $\boldsymbol{\Phi}_i$ 是主对角线元素全为 1 的下三角矩阵，第(j,k) 个元素是如下自回归方程系数 $\phi_{ijk,\gamma}$ 的相反数

$$\psi_\gamma(Y_{ij}-\tilde{\boldsymbol{\chi}}_{ij}\boldsymbol{\eta})=\sum_{k=1}^{j-1}\phi_{ijk}\psi_\gamma(Y_{ik}-\tilde{\boldsymbol{\chi}}_{ik}\boldsymbol{\eta})+\varepsilon_{ij}$$

特别地，当 $j=1$ 时，$\varphi_\gamma(Y_{ij}-\tilde{\boldsymbol{\chi}}_{ij}\boldsymbol{\eta})=\varepsilon_{ij}$. $\boldsymbol{D}_i=\mathrm{diag}(d_{i1}^2,\cdots,d_{im_i}^2)$，其中 $d_{ij}^2=\mathrm{Var}(\varepsilon_{ij})$ 为新息方差. 对 ϕ_{ijk} 及 d_{ij}^2 建立广义线性方程

$$\phi_{ijk}=\boldsymbol{\omega}_{ijk}^{\mathrm{T}}\boldsymbol{\rho},\ \log(d_{ij}^2)=\boldsymbol{r}_{ij}^{\mathrm{T}}\boldsymbol{\varrho} \tag{7}$$

其中：$\boldsymbol{\varrho}$与 r_{ij} 为 q 维向量，$\boldsymbol{\rho}$ 与 $\boldsymbol{\omega}_{ijk}$ 为 p 维向量，r_{ij} 及 $\boldsymbol{\omega}_{ijk}$ 通常与观测时间相关. 参考文献[15]，在本文的模拟与实证中均设定 $\boldsymbol{r}_{ij}=(1,t_{ij},\cdots,t_{ij}^{q-1})^{\mathrm{T}}$，$\boldsymbol{\omega}_{ijk}=(1,t_{ij}-t_{ik},\cdots,(t_{ij}-t_{ik})^{p-1})^{\mathrm{T}}$，其中 t_{ij} 为时间变量，记录第 i 名受试者第 j 次观测的时间. 记 $\boldsymbol{\xi}=(\boldsymbol{\rho}^{\mathrm{T}},\boldsymbol{\varrho}^{\mathrm{T}},\boldsymbol{\eta}^{\mathrm{T}})^{\mathrm{T}}$，关于 $\boldsymbol{\xi}$ 求解如下广义估计方程组

$$\boldsymbol{U}(\boldsymbol{\xi})=(\boldsymbol{U}_1^{\mathrm{T}}(\boldsymbol{\xi}),\boldsymbol{U}_2^{\mathrm{T}}(\boldsymbol{\xi}),\boldsymbol{U}_3^{\mathrm{T}}(\boldsymbol{\xi}))^{\mathrm{T}}=\boldsymbol{0} \tag{8}$$

其中

$$\boldsymbol{U}_1(\boldsymbol{\xi})=\sum_{i=1}^{n}\boldsymbol{T}_i^{\mathrm{T}}\boldsymbol{D}_i^{-1}\varepsilon_i$$

$$\boldsymbol{U}_2(\boldsymbol{\xi})=\sum_{i=1}^{n}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{D}_i\boldsymbol{W}_i^{-1}(\boldsymbol{\varepsilon}_i^2-\boldsymbol{d}_i^2)$$

$$\boldsymbol{U}_3(\boldsymbol{\xi})=\sum_{i=1}^{n}\tilde{\boldsymbol{\chi}}_i^{\mathrm{T}}\boldsymbol{V}_i^{-1}\psi_\gamma(\boldsymbol{Y}_i-\tilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})$$

$\boldsymbol{T}_i^{\mathrm{T}}=-\frac{\partial\boldsymbol{\varepsilon}_i^{\mathrm{T}}}{\partial\boldsymbol{\rho}}$，$\boldsymbol{\varepsilon}_i=(\varepsilon_{i1},\cdots,\varepsilon_{im_i})^{\mathrm{T}}$，$\boldsymbol{d}_i^2=(d_{i1}^2,\cdots,d_{im_i}^2)^{\mathrm{T}}$，$\boldsymbol{R}_i=(r_{i1},\cdots,r_{im_i})^{\mathrm{T}}$，$\boldsymbol{W}_i=\mathrm{Cov}(\boldsymbol{\varepsilon}_i^2)$. 参考文献[15]，在本文的模拟与实证中均设定$\boldsymbol{W}_i=2\mathrm{diag}(d_{i1}^4,\cdots,d_{im_i}^4)$.

本文使用 Newton-Raphson 迭代算法求解方程组(8) 以保证估计的精度. 下面给出在指定调谐参数 γ 下，基于指数平方损失的纵向多折点回归模型的参数估计算法：

步骤 1　给定初始折点位置$\boldsymbol{\tau}^{(0)}$，计算模型(3) 的普通最小二乘估计作为 $\boldsymbol{\eta}$ 的初始值 $\boldsymbol{\eta}^{(0)}$.

步骤 2　基于$\boldsymbol{\tau}^{(s)}$ 及 $\boldsymbol{\eta}^{(s)}$，应用修正的Cholesky 分解法估计协方差阵 $\boldsymbol{V}_i$，使用 Newton-Raphson 算法迭代求解方程组(8) 获得$\boldsymbol{\eta}^{(s+1)}$，具体步骤如下：

步骤 2.1　设定 $\boldsymbol{V}_i^{(s+1,0)}$ 为$m_i\times m_i$ 的单位阵，$\boldsymbol{\rho}^{(s+1,0)}=\boldsymbol{0}$，$\boldsymbol{\varrho}^{(s+1,0)}=\boldsymbol{0}$，给定折点位置 $\boldsymbol{\tau}^{(s)}$，以模型(3) 的普通最小二乘估计作为 $\boldsymbol{\eta}^{(s+1,0)}$.

步骤 2.2　基于 $\boldsymbol{\eta}^{(s+1,r)}$，$\boldsymbol{\varrho}^{(s+1,r)}$ 以及 $\boldsymbol{\rho}^{(s+1,r)}$，求解 $\boldsymbol{\eta}^{(s+1,r+1)}$，具体步骤如下：

步骤 2.2.1　基于 $\boldsymbol{\eta}^{(s+1,r)}$ 及$\boldsymbol{\varrho}^{(s+1,r)}$，通过下式迭代至收敛得到 $\boldsymbol{\rho}^{(s+1,r+1)}$：

$$\boldsymbol{\rho}^{(s+1,\ r+1,\ l_1+1)}=\boldsymbol{\rho}^{(s+1,\ r+1,\ l_1)}+\left\{\left(\sum_{i=1}^{n}\boldsymbol{T}_i^{\mathrm{T}}\boldsymbol{D}_i^{-1}\boldsymbol{T}_i\right)^{-1}\times\sum_{i=1}^{n}\boldsymbol{T}_i^{\mathrm{T}}\boldsymbol{D}_i^{-1}\boldsymbol{\varepsilon}_i\right\}\bigg|_{\boldsymbol{\eta}=\boldsymbol{\eta}^{(s+1,\ r)},\ \boldsymbol{\rho}=\boldsymbol{\rho}^{(s+1,\ r+1,\ l_1)},\ \boldsymbol{\varrho}=\boldsymbol{\varrho}^{(s+1,\ r)}}$$

步骤 2.2.2 基于 $\boldsymbol{\eta}^{(s+1,\ r)}$ 及 $\boldsymbol{\rho}^{(s+1,\ r+1)}$，通过下式迭代至收敛得到$\boldsymbol{\varrho}^{(s+1,\ r+1)}$：

$$\boldsymbol{\varrho}^{(s+1,\ r+1,\ l_2+1)}=\boldsymbol{\varrho}^{(s+1,\ r+1,\ l_2)}+\left\{\left(\sum_{i=1}^{n}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{D}_i\boldsymbol{W}_i^{-1}\boldsymbol{D}_i\boldsymbol{R}_i\right)^{-1}\times\sum_{i=1}^{n}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{D}_i\boldsymbol{W}_i^{-1}(\boldsymbol{\varepsilon}_i^2-\boldsymbol{d}_i^2)\right\}\bigg|_{\boldsymbol{\eta}=\boldsymbol{\eta}^{(s+1,\ r)},\ \boldsymbol{\rho}=\boldsymbol{\rho}^{(s+1,\ r+1)},\ \boldsymbol{\varrho}=\boldsymbol{\varrho}^{(s+1,\ r+1,\ l_2)}}$$

步骤 2.2.3 根据(6)式及(7)式，可得 $\boldsymbol{V}_{i,\ r+1}$. 通过下式迭代至收敛得到 $\boldsymbol{\eta}^{(s+1,\ r+1)}$：

$$\boldsymbol{\eta}^{(s+1,\ r+1,\ l_3+1)}=\boldsymbol{\eta}^{(s+1,\ r+1,\ l_3)}+$$
$$\left\{\left(\sum_{i=1}^{n}\widetilde{\boldsymbol{\chi}}_i^{\mathrm{T}}\boldsymbol{V}_{i,\ r+1}^{-1}\boldsymbol{\Lambda}_i(\boldsymbol{\eta})\widetilde{\boldsymbol{\chi}}_i\right)^{-1}\times\sum_{i=1}^{n}\widetilde{\boldsymbol{\chi}}_i^{\mathrm{T}}\boldsymbol{V}_{i,\ r+1}^{-1}\psi_\gamma(\boldsymbol{Y}_i-\widetilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})\right\}\bigg|_{\boldsymbol{\eta}=\boldsymbol{\eta}^{(s+1,\ r+1,\ l_3)},\ \boldsymbol{\rho}=\boldsymbol{\rho}^{(s+1,\ r+1)},\ \boldsymbol{\varrho}=\boldsymbol{\varrho}^{(s+1,\ r+1)}}$$

其中 $\boldsymbol{\Lambda}_i(\boldsymbol{\eta})=\mathrm{diag}\{\Lambda_{i1}(\boldsymbol{\eta}),\ \cdots,\ \Lambda_{im_i}(\boldsymbol{\eta})\}$，$\Lambda_{ij}(\boldsymbol{\eta})=\psi'_\gamma(Y_{ij}-\widetilde{\boldsymbol{\chi}}_{ij}\boldsymbol{\eta})$，$\psi'_\gamma(x)=\exp\left(-\dfrac{x^2}{\gamma}\right)\left(\dfrac{2}{\gamma}-\dfrac{4x^2}{\gamma^2}\right)$.

步骤 2.3 重复步骤 2.2 直至参数收敛，可得 $\boldsymbol{\eta}^{(s+1)}$.

步骤 3 更新折点位置$\boldsymbol{\tau}^{(s+1)}$. 通过

$$\tau_k^{(s+1)}=\tau_k^{(s)}+\frac{\mu_k^{(s+1)}}{b_k^{(s+1)}},\ k=1,2,\cdots,K \tag{9}$$

步骤 4 重复步骤 2－步骤 3 直至参数收敛.

注 1 当 $\boldsymbol{V}_i$ 为单位阵时，方程(5)和最小化目标函数(4)的解等价. 因此，在上述步骤中省略步骤 2.2.1 及步骤 2.2.2，设置步骤 2.2.3 中 $\boldsymbol{V}_{i,\ r+1}$ 为单位阵，即可获得独立工作矩阵情况的参数估计.

1.3 最佳调谐参数 γ_{opt} 的选择

在指数平方损失函数 $\varphi_\gamma(x)=1-\exp\left(-\dfrac{x^2}{\gamma}\right)$ 中，如何选择调谐参数 γ 实现回归参数的最佳稳健估计是一个重要问题. 参考文献[15]，使用网格搜索法，选择使得回归参数估计值的渐近协方差阵的行列式最小的 γ 作为最优调谐参数 γ_{opt}，具体实现方法见第 2 节中注 2.

1.4 折点个数 K 的确定

1.2 节给出了当折点个数给定的时候模型参数的估计算法. 然而实际问题中，折点个数真值 K_0 通常是未知的. 参考文献[4]，提出如下基于指数平方损失的贝叶斯准则以确定折点个数：

$$\mathrm{BIC}(k)=\log\left(\frac{S_n(\hat{\boldsymbol{\eta}}_k)}{n}\right)+P_k\,\frac{\log n}{2n}C_n \tag{10}$$

其中$\hat{\boldsymbol{\eta}}_k$ 为指定折点个数为 k 时获得的参数估计，P_k 为此时模型中未知参数个数，S_n 的表达式见(4)式，$C_n>0$ 是与样本量有关的常数. 给定折点个数的最大值 K^*，依次设定折点个数 $k=0,1,\cdots,K^*$，选择使得(10)式最小的 k 为实际折点个数的估计值，记作$\hat{K}$.

2 大样本性质

记参数 $\boldsymbol{\xi}=(\boldsymbol{\rho}^{\mathrm{T}},\ \boldsymbol{\varrho}^{\mathrm{T}},\ \boldsymbol{\vartheta}^{\mathrm{T}},\boldsymbol{\mu}^{\mathrm{T}})^{\mathrm{T}}$ 的真值为 $\boldsymbol{\xi}_0=(\boldsymbol{\rho}_0^{\mathrm{T}},\ \boldsymbol{\varrho}_0^{\mathrm{T}},\ \boldsymbol{\vartheta}_0^{\mathrm{T}},\boldsymbol{\mu}_0^{\mathrm{T}})^{\mathrm{T}}$，参数的估计值记作$\hat{\boldsymbol{\xi}}=(\hat{\boldsymbol{\rho}}^{\mathrm{T}},\ \hat{\boldsymbol{\varrho}}^{\mathrm{T}},\ \hat{\boldsymbol{\vartheta}}^{\mathrm{T}},\ \hat{\boldsymbol{\mu}}^{\mathrm{T}})^{\mathrm{T}}$. 定义 $\widetilde{\boldsymbol{\chi}}_i=(\widetilde{\boldsymbol{\chi}}_{1,\ i},\ \widetilde{\boldsymbol{\chi}}_{2,\ i})$，$\widetilde{\boldsymbol{\chi}}_{1,\ i}=(\widetilde{\boldsymbol{\chi}}_{1,\ i1},\ \cdots,\ \widetilde{\boldsymbol{\chi}}_{1,\ im_i})^{\mathrm{T}}$，$\widetilde{\boldsymbol{\chi}}_{1,\ ij}=(1,\ X_{ij},\ (X_{ij}-\tau_1^{(0)})_+,\ \cdots,\ (X_{ij}-\tau_K^{(0)})_+,\ \boldsymbol{Z}_{ij}^{\mathrm{T}})$，$\widetilde{\boldsymbol{\chi}}_{2,\ i}=(\widetilde{\boldsymbol{\chi}}_{2,\ i1},\ \cdots,\ \widetilde{\boldsymbol{\chi}}_{2,\ im_i})^{\mathrm{T}}$，$\widetilde{\boldsymbol{\chi}}_{2,\ ij}=(-I(X_{ij}>\tau_1^{(0)}),\ \cdots,\ -I(X_{ij}>\tau_K^{(0)}))$. 因此，

$$\boldsymbol{U}_3(\boldsymbol{\xi})=\begin{pmatrix}\boldsymbol{U}_{1,\ 3}(\boldsymbol{\xi})\\ \boldsymbol{U}_{2,\ 3}(\boldsymbol{\xi})\end{pmatrix}=\begin{pmatrix}\displaystyle\sum_{i=1}^{n}\widetilde{\boldsymbol{\chi}}_{1,\ i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}\psi_\gamma(\boldsymbol{Y}_i-\widetilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})\\ \displaystyle\sum_{i=1}^{n}\widetilde{\boldsymbol{\chi}}_{2,\ i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}\psi_\gamma(\boldsymbol{Y}_i-\widetilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})\end{pmatrix} \tag{11}$$

则 $\boldsymbol{U}(\boldsymbol{\xi})=(\boldsymbol{U}_1^{\mathrm{T}}(\boldsymbol{\xi}),\boldsymbol{U}_2^{\mathrm{T}}(\boldsymbol{\xi}),\boldsymbol{U}_{1,\ 3}^{\mathrm{T}}(\boldsymbol{\xi}),\boldsymbol{U}_{2,\ 3}^{\mathrm{T}}(\boldsymbol{\xi}))^{\mathrm{T}}$. 记 $\boldsymbol{\Gamma}_n=(\boldsymbol{\Gamma}_n^{kl})_{k,l=1,2,3,4}$ 为$\dfrac{\boldsymbol{U}(\boldsymbol{\xi}_0)}{\sqrt{n}}$ 的协方差阵. 当 $n\to\infty$ 时 $\boldsymbol{\Gamma}_n^{kl}\xrightarrow{p}\boldsymbol{\Gamma}^{kl}$，记 $\boldsymbol{\Gamma}=(\boldsymbol{\Gamma}^{kl})_{k,l=1,2,3,4}$. 根据 1.2 节的讨论，协方差阵 $\boldsymbol{V}_i$ 与参数 $\boldsymbol{\xi}$ 相关，为方便表示，本节将 $\boldsymbol{V}_i$

重记为 $\boldsymbol{V}_i(\boldsymbol{\xi})$. 定义如下矩阵

$$\boldsymbol{A}_n^{11}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{T}_i^{\mathrm{T}}\boldsymbol{D}_i^{-1}\frac{\partial\boldsymbol{\varepsilon}_i}{\partial\boldsymbol{\rho}}\Bigg|_{\xi=\xi_0}\qquad \boldsymbol{A}_n^{22}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{Z}_i^{\mathrm{T}}\boldsymbol{D}_i\boldsymbol{W}_i^{-1}\frac{\partial(\boldsymbol{\varepsilon}_i^2-\boldsymbol{d}_i^2)}{\partial\boldsymbol{\varrho}}\Bigg|_{\xi=\xi_0}$$

$$\boldsymbol{A}_n^{33}=\frac{1}{n}\sum_{i=1}^{n}\widetilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}(\boldsymbol{\xi})\frac{\partial\psi_\gamma(\boldsymbol{Y}_i-\widetilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})}{\partial\boldsymbol{\vartheta}}\Bigg|_{\xi=\xi_0}\qquad \boldsymbol{A}_n^{44}=\frac{1}{n}\sum_{i=1}^{n}\widetilde{\boldsymbol{\chi}}_{2,i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}(\boldsymbol{\xi})\frac{\partial\psi_\gamma(\boldsymbol{Y}_i-\widetilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})}{\partial\boldsymbol{\mu}}\Bigg|_{\xi=\xi_0}$$

给出如下条件：

（ⅰ）$\mathrm{E}(\psi_\gamma(e_{ij}))=0$，$\mathrm{E}(\psi'_\gamma(e_{ij}))>0$，且对于任意 $\gamma>0$，$\mathrm{E}(\psi_\gamma(e_{ij})^2)$ 有界.

（ⅱ）折点的真值 K_0 及协变量 $\boldsymbol{Z}_{ij}$ 的维数 p 是固定的，个体的观测次数 m_i 一致有界，且个体数 n 趋于无穷.

（ⅲ）X_{ij} 在定义域上有连续的密度函数. 当 $n\to\infty$，$\frac{1}{n}\sum_{ij}\|\boldsymbol{Z}_{ij}\|^2=O(1)$. 另外，$w_{ijk}$，$r_{ij}$ 及矩阵 $\boldsymbol{W}_i$ 均有界.

（ⅳ）记 $\boldsymbol{\Sigma}_{i,0}=\mathrm{Cov}(\psi_\gamma(\boldsymbol{e}_i))$，矩阵 $\boldsymbol{\Sigma}_{i,0}$ 及 $\boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0)\boldsymbol{\Sigma}_{i,0}\boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0)$ 的最大特征值均有界.

（ⅴ）对于 $k=1,\cdots,K$，折点参数的初始值 $\tau_k^{(0)}$ 满足 $\tau_{k,0}-\tau_k^{(0)}=O_p(n^{-\frac{1}{2}})$.

（ⅵ）参数空间 $\Theta\in\mathbb{R}^{2+p+2K}$ 是紧集，参数真值 $\boldsymbol{\eta}_0$ 是参数空间 Θ 的内点，$S_n(\boldsymbol{\eta})$ 在 $\boldsymbol{\eta}_0$ 处唯一取得全局最小值.

（ⅶ）当 $n\to\infty$ 时，$\dfrac{C_n\log n}{\sqrt{n}}\to 0$.

定理 1　假设真实折点个数 K_0 已知，且条件（ⅰ）—（ⅴ）成立，则当 $n\to\infty$ 时

$$\sqrt{n}\begin{pmatrix}\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}_0\\ \hat{\boldsymbol{\varrho}}-\boldsymbol{\varrho}_0\\ \hat{\boldsymbol{\vartheta}}-\boldsymbol{\vartheta}_0\\ \hat{\boldsymbol{\mu}}-\boldsymbol{\mu}_0\end{pmatrix}\xrightarrow{\mathrm{d}}N(\boldsymbol{0},\ \boldsymbol{A}^{-1}\boldsymbol{\Gamma}\boldsymbol{A}^{-1})$$

其中 $\boldsymbol{A}=\mathrm{diag}(\boldsymbol{A}^{11},\ \boldsymbol{A}^{22},\ \boldsymbol{A}^{33},\ \boldsymbol{A}^{44})$，$\boldsymbol{A}^{kk}=\lim\limits_{n\to\infty}\boldsymbol{A}_n^{kk}$，$k=1,2,3,4$.

注 2　由定理 1，回归系数估计 $\hat{\boldsymbol{\eta}}=(\hat{\boldsymbol{\vartheta}}^{\mathrm{T}},\ \hat{\boldsymbol{\mu}}^{\mathrm{T}})^{\mathrm{T}}$ 的渐近协方差阵可通过下式对其进行估计，即

$$\widehat{\mathrm{Cov}(\hat{\boldsymbol{\eta}})}=\hat{\boldsymbol{H}}_{\gamma0}^{-1}\hat{\boldsymbol{H}}_{\gamma1}\hat{\boldsymbol{H}}_{\gamma0}^{-1}\tag{12}$$

其中

$$\hat{\boldsymbol{H}}_{\gamma0}=\sum_{i=1}^{n}\widetilde{\boldsymbol{\chi}}_i^{\mathrm{T}}\boldsymbol{V}_i^{-1}(\boldsymbol{\xi})\boldsymbol{\Lambda}_i(\boldsymbol{\eta})\widetilde{\boldsymbol{\chi}}_i\Big|_{\xi=\hat{\xi}}\tag{13}$$

$$\hat{\boldsymbol{H}}_{\gamma1}=\sum_{i=1}^{n}\widetilde{\boldsymbol{\chi}}_i^{\mathrm{T}}\boldsymbol{V}_i^{-1}(\boldsymbol{\xi})\psi_\gamma(\boldsymbol{Y}_i-\widetilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})\times\{\psi_\gamma(\boldsymbol{Y}_i-\widetilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})\}^{\mathrm{T}}\boldsymbol{V}_i^{-1}(\boldsymbol{\xi})\widetilde{\boldsymbol{\chi}}_i\Big|_{\xi=\hat{\xi}}\tag{14}$$

选择使参数估计值的渐近协方差阵行列式，即 $\det(\widehat{\mathrm{Cov}(\hat{\boldsymbol{\eta}})})$ 最小的 γ 为最优调谐参数 γ_{opt}. 在本文后续的模拟与实证中，采用网格搜索法寻找 γ_{opt}，在 2 ～ 50 之间以步长 3 进行搜索.

定理 2　假设真实折点个数 K_0 已知，且条件（ⅰ）—（ⅴ）成立，则当 $n\to\infty$ 时

$$\sqrt{n}(\hat{\boldsymbol{\tau}}-\boldsymbol{\tau}_0)\xrightarrow{\mathrm{d}}N(\boldsymbol{0},\ \boldsymbol{B}^{-1}\boldsymbol{\Gamma}^{44}\boldsymbol{B}^{-1})$$

其中 $\boldsymbol{B}=\mathrm{diag}\left(\dfrac{1}{b_{1,0}},\ \cdots,\ \dfrac{1}{b_{K,0}}\right)$.

定理 3　假设条件（ⅰ）—（ⅶ）成立，则对于 $\hat{K}=\arg\min\limits_{k=0,\cdots,K^*}\mathrm{BIC}(k)$，当 $n\to\infty$ 时

$$P(\hat{K}=K_0)\to 1$$

3 折点的存在性检验

对于多折点回归模型的折点存在性检验，文献[4]基于无折点的原假设 H_0: $b_k=0$, $k=1,\cdots,K$ 提出 CUSUM(cumulative summation)型统计量. 本文沿用文献[4]所提方法，提出基于指数平方损失的纵向折点回归模型的折点存在性检验方法，具体步骤如下：

步骤1 计算 H_0 成立时系数参数 $\zeta=(a_0, a_1, \boldsymbol{\beta}^{\mathrm{T}})^{\mathrm{T}}$ 的估计 $\hat{\zeta}$，即关于 ζ 最小化如下目标函数

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}\varphi_\gamma(Y_{ij}-\zeta^{\mathrm{T}}\boldsymbol{M}_{ij})$$

其中 $\boldsymbol{M}_{ij}=(1, X_{ij}, \boldsymbol{Z}_{ij}^{\mathrm{T}})^{\mathrm{T}}$, $\varphi_\gamma(x)=1-\exp\left(-\dfrac{x^2}{\gamma}\right)$ 为指数平方损失函数.

步骤2 基于原始样本数据计算 $T_n(\gamma)=\sup\limits_{\tau\in T}|R_n(\tau, \gamma, \hat{\zeta})|$，其中

$$R_n(\tau, \gamma, \hat{\zeta})=n^{-\frac{1}{2}}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\psi_\gamma(Y_{ij}-\hat{\zeta}^{\mathrm{T}}\boldsymbol{M}_{ij})(X_{ij}-\tau)_-$$

$\psi_\gamma(x)=\dfrac{2x}{\gamma}\exp\left(-\dfrac{x^2}{\gamma}\right)$, $(a)_-=a\cdot I(a<0)$，T 是折点参数的取值范围.

步骤3 生成服从标准正态分布的随机样本 $\{v_1, \cdots, v_n\}$，计算 $T_n^*(\gamma)=\sup\limits_{\tau\in T}|R_n^*(\tau, \gamma, \hat{\zeta})|$，其中

$$R_n^*(\tau, \gamma, \hat{\zeta})=n^{-\frac{1}{2}}\sum_{i=1}^{n}v_i\sum_{j=1}^{m_i}\psi_\gamma(Y_{ij}-\hat{\zeta}^{\mathrm{T}}\boldsymbol{M}_{ij})\{(X_{ij}-\tau)_--\hat{\boldsymbol{G}}_{n1}^{\mathrm{T}}(\tau)\hat{\boldsymbol{G}}_n^{-1}\boldsymbol{M}_{ij}\}$$

$$\hat{\boldsymbol{G}}_n=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\psi'_\gamma(Y_{ij}-\hat{\zeta}^{\mathrm{T}}\boldsymbol{M}_{ij})\boldsymbol{M}_{ij}\boldsymbol{M}_{ij}^{\mathrm{T}},\ \hat{\boldsymbol{G}}_{n1}(\tau)=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\psi'_\gamma(Y_{ij}-\hat{\zeta}^{\mathrm{T}}\boldsymbol{M}_{ij})\boldsymbol{M}_{ij}(X_{ij}-\tau)_-$$

步骤4 重复步骤3 L 次，计算

$$\hat{p}_n=\frac{1}{L}\sum_{l=1}^{L}I\{T_n^{*(b)}(\gamma)\geqslant T_n(\gamma)\}$$

作为统计量 $T_n(\gamma)$ 的 p 值. 一般情况，若 $\hat{p}_n$ 小于显著性水平0.05，则拒绝原假设.

注3 在步骤2和步骤3中 $R_n^*(\tau, \gamma, \hat{\zeta})$ 与 $R_n(\tau, \gamma, \hat{\zeta})$ 渐近等价，其证明过程与文献[6]类似.

4 模拟研究

根据如下模型生成数据

$$Y_{ij}=a_0+a_1X_{ij}+\sum_{k=1}^{K}b_k(X_{ij}-\tau_k)_++\beta Z_{ij}+e_{ij},\ i=1,\cdots,n,\ j=1,\cdots,m_i \tag{15}$$

其中 $Z_{ij}\overset{\text{i.i.d}}{\sim}N(5, 1)$. 对于门限变量 X_{ij}，时间变量 t_{ij} 以及随机误差 e_{ij}，考虑以下3种情形：

情形1 考虑时间 t_{ij} 为门限变量，残差服从正态分布的情况. 设置 $t_{ij}=X_{ij}\sim U(1, 10)$, $m_i=10$, $i=1,\cdots,n$. $\boldsymbol{e}_i=(e_{i1}, \cdots, e_{im_i})^{\mathrm{T}}$ 服从多元正态分布 $N(\boldsymbol{0}, \boldsymbol{\Sigma}_i)$，$\boldsymbol{\Sigma}_i=\nabla_i^{-1}\boldsymbol{B}_i(\nabla_i^{\mathrm{T}})^{-1}$，其中 $\boldsymbol{B}_i$ 对角元素为 $\exp(-0.5+0.5u_{ij})$, $u_{ij}\sim N(0, 2)$. ∇_i 是单位下三角矩阵，j 行 k 列元素 $\delta_{jk}^{(i)}=-(0.2+0.8(t_{ij}-t_{ik}))$, $k<j$.

情形2 考虑门限变量不为时间，且数据存在异常值的情况. 设置 $X_{ij}\sim U(-5, 5)$, $t_{ij}=\{1, 2, \cdots, 10\}$，随后每个观测以20%的概率缺失以生成不平衡数据. $\boldsymbol{e}_i$ 服从自由度为3，协方差阵为情形1中 $\boldsymbol{\Sigma}_i$ 的多元 t 分布，随后随机选取10%的 e_{ij} 服从标准柯西分布.

情形3 为公平起见，以等相关结构生成残差来说明修正的Cholesky分解法估计参数的有效性. 设置 $\boldsymbol{e}_i$ 服从自由度为3，协方差阵为 $\boldsymbol{\Sigma}_i=\boldsymbol{A}_i^{\frac{1}{2}}\boldsymbol{C}_i\boldsymbol{A}_i^{\frac{1}{2}}$ 的多元正态分布，其中 $\boldsymbol{A}_i=\mathrm{diag}(\sigma_{i1}^2, \cdots, \sigma_{im_i}^2)$, $\sigma_{ij}^2=2t_{ij}^2$, $\boldsymbol{C}_i$ 是相关系数为0.85的等相关结构. 随后随机选取10%的 e_{ij} 服从标准柯西分布. 其他设置与情形2相同.

每种情形均考虑3种折线效应，对于情形1，设定：①$K=1$, $\boldsymbol{\tau}=7$, $\boldsymbol{\vartheta}=(-2, 1, -3, 1)^{\mathrm{T}}$；②$K=2$,

$\boldsymbol{\tau}=(5,7)^{\mathrm{T}}$，$\boldsymbol{\vartheta}=(2,1,-4,5,1)^{\mathrm{T}}$；③$K=3$，$\boldsymbol{\tau}=(2,5,7)^{\mathrm{T}}$，$\boldsymbol{\vartheta}=(0,1,-3,2,-6,1)^{\mathrm{T}}$. 对于情形 2、3，系数 $\boldsymbol{\vartheta}$ 设置与情形 1 相同，折点个数和位置设置为：①$K=1$，$\boldsymbol{\tau}=-1$；②$K=2$，$\boldsymbol{\tau}=(-1,1)^{\mathrm{T}}$；③$K=3$，$\boldsymbol{\tau}=(-4,-1,4)^{\mathrm{T}}$. 设定样本量 $n=400$，重复模拟次数为 100 次. 使用 1.2 节所提算法求解最小化目标函数(4) 及方程组(8)，分别记为"ESL.IND" 和"ESL.CHO"，即独立工作矩阵结构及基于 Cholesky 分解的模型参数的估计方法.

4.1 最优调谐参数 γ_{opt} 的选择

表 1 给出 3 种情形下所有模拟中选择的最优参数 γ_{opt} 的均值. 可以看到，无异常值的情形下，γ_{opt} 均值较大；情形 2,3 的 γ_{opt} 较小，以降低异常值产生的影响. 这说明在指数平方损失函数中，可以通过选择合适的调谐参数实现回归系数的自适应稳健估计.

表 1 最优调谐参数 γ_{opt} 的均值

	情形 1			情形 2			情形 3		
	$K=1$	$K=2$	$K=3$	$K=1$	$K=2$	$K=3$	$K=1$	$K=2$	$K=3$
ESL.IND	45.11	47.75	47.30	7.97	9.26	10.10	16.64	16.46	15.65
ESL. CHO	44.21	47.78	46.85	7.55	8.24	9.32	14.03	14.12	13.43

4.2 选择相合性及参数估计

表 2 给出了真实折点个数为 2 时，不同 C_n 在 3 种情形下的折点个数的正确选择率. 可以看到，所有情形下，本文所提出的折点个数选择方法均具有较高的正确选择率，并且"ESL. CHO" 能够实现更高的选择正确率. 而且样本量不变时，较大的 C_n 能略微提高正确选择折点数的概率.

表 2 $K=2$，$C_n=1$，$\log(\log n)$，$\log n$，$2\log n$ 时的折点个数正确选择率

C_n	情形 1		情形 2		情形 3	
	ESL.IND	ESL.CHO	ESL.IND	ESL.CHO	ESL.IND	ESL.CHO
1	0.97	1.00	0.92	0.98	0.96	1.00
$\log(\log n)$	0.97	1.00	0.92	0.98	0.96	1.00
$\log n$	0.97	1.00	0.92	0.98	0.97	1.00
$2\log n$	1.00	1.00	0.93	0.98	0.99	1.00

为说明所提方法优势，与以下估计量进行比较：

1）方程(5) 的工作协方差矩阵 $\mathbf{V}_i$ 为等相关及 AR(1) 的特定结构，利用文献[7] 提出的 QIF 方法对(5) 式进行求解，同时分别考虑指数平方损失及经典的平方损失两种损失函数，记所得估计量分别为"ESL. EXCH""OLS.EXCH""ESL.AR1""OLS.AR1"；

2）文献[4] 提出的基于分位数回归的纵向多折点模型的估计量，指定分位水平为 0.5，记为"MKQR"，使用 R 程序包 MultiKink 实现；

3）文献[3] 提出的多折点模型的最小二乘估计量，记为"SEG"，使用 R 程序包 segmented 实现.

表 3 展示了情形 2 在 K=2 时的参数估计结果，汇总了 100 次模拟中估计量的平均偏差、标准差及均方误差. 其余情形的参数估计的结果已上传至 Github(https://github.com/Tangming-hub). 可以发现：

1）当残差向量服从正态分布时(情形 1)，几种方法所得到的估计量的估计效果相近. 然而当数据中存在异常值时(情形 2 和情形 3)，基于指数平方损失和分位数回归的估计量"ESL.IND""ESL.CHO""ESL. EXCH""ESL.AR1" 和"MKQR" 均能提供回归参数的有效估计，其中相对于分位数回归方法，基于指数平方损失函数的估计量表现更佳，而基于经典的平方损失函数的估计量的估计效果均较差；

2）仅考虑指数平方损失函数的估计方法时，相对于独立工作矩阵的估计量"ESL.IND"，融合组内相关性的估计量"ESL.CHO""ESL.EXCH""ESL.AR1" 均具有更优良的估计效果，且本文所提出的"ESL.CHO" 在各指标上表现最佳. 这说明有效考虑纵向数据个体重复观测有利于提高回归参数的估计效率. 综上，本文所提出的估计方法，可以为纵向多折点回归模型的参数提供更为稳健的、有效的估计.

此外，不同情形下 $K=2$ 时，3 种估计方法"ESL.CHO""ESL.IND""MKQR" 的标准差、标准误、95%

的 Wald 型置信区间的平均长度及经验覆盖率亦上传至 Github. 由结果可见，3 种估计量的标准差与经验标准误差接近，并且所构造的置信区间的经验覆盖率均在置信水平 95% 左右. 并且，相对于“MKQR”方法，本文所提方法的置信区间的长度更短.

表 3　情形 2，$K=2$ 的模拟结果

	方法	τ_1	τ_2	a_0	a_1	b_1	b_2	β
Bias	ESL. IND	0.002 1	−0.004 0	0.003 3	0.000 8	−0.006 8	0.001 0	0.001 1
	ESL. CHO	0.000 6	−0.001 7	0.006 8	0.001 7	−0.007 2	0.006 5	0.001 1
	ESL. AR1	−0.000 9	−0.001 4	0.013 3	0.004 0	−0.003 5	−0.001 2	0.000 8
	ESL. EXCH	0.000 9	−0.000 5	0.010 8	0.002 1	−0.006 4	0.006 8	0.000 4
	OLS. AR1	0.005 6	−0.006 9	−0.005 6	0.005 7	−0.027 1	0.011 1	0.004 4
	OLS. EXCH	0.012 8	−0.000 2	−0.056 0	−0.011 0	−0.020 3	0.032 3	0.006 0
	MKQR	0.003 4	−0.005 2	−0.000 4	−0.003 6	−0.003 7	0.001 0	−0.001 1
	SEG	−0.001 8	0.020 3	−0.080 3	0.011 9	−0.035 1	0.051 3	0.026 4
SD	ESL. IND	0.032 5	0.024 7	0.156 6	0.030 7	0.093 2	0.101 8	0.022 7
	ESL. CHO	0.018 1	0.013 2	0.090 3	0.017 0	0.056 1	0.058 8	0.011 5
	ESL. AR1	0.029 2	0.019 9	0.148 7	0.026 5	0.084 5	0.089 7	0.019 9
	ESL. EXCH	0.022 1	0.016 4	0.123 2	0.019 5	0.068 8	0.070 4	0.014 7
	OLS. AR1	0.057 2	0.043 1	0.293 5	0.060 6	0.162 5	0.170 7	0.042 3
	OLS. EXCH	0.086 7	0.083 4	0.411 6	0.088 4	0.271 4	0.252 4	0.066 1
	MKQR	0.034 3	0.029 7	0.165 7	0.034 7	0.103 0	0.109 4	0.022 9
	SEG	0.149 1	0.199 9	1.222 8	0.204 2	0.587 7	0.530 2	0.189 9
MSE	ESL. IND	0.001 1	0.000 6	0.024 3	0.000 9	0.008 7	0.010 3	0.000 5
	ESL. CHO	0.000 3	0.000 2	0.008 1	0.000 3	0.003 2	0.003 5	0.000 1
	ESL. AR1	0.000 8	0.000 4	0.022 1	0.000 7	0.007 1	0.008 0	0.000 4
	ESL. EXCH	0.000 5	0.000 3	0.015 2	0.000 4	0.004 7	0.005 0	0.000 2
	OLS. AR1	0.003 3	0.001 9	0.085 3	0.003 7	0.026 9	0.029 0	0.001 8
	OLS. EXCH	0.007 6	0.006 9	0.170 8	0.007 9	0.073 3	0.064 1	0.004 4
	MKQR	0.001 2	0.000 9	0.027 2	0.001 2	0.010 5	0.011 9	0.000 5
	SEG	0.022 0	0.040 0	1.486 7	0.041 4	0.343 1	0.280 9	0.036 4

4.3　折线效应存在性检验的功效性分析

为研究本文第 3 节所提出的折线效应存在性检验统计量 T_n 的有限样本性质，考虑折点个数为 2 时检验统计量的功效. 对于情形 1，设置 $-b_1=b_2=0, 0.1, 0.2, 0.3$；对于情形 2 和情形 3，设置 $-b_1=b_2=0, 0.15, 0.3, 0.45$，其中当系数 b_1 和 b_2 都等于 0 时，不存在折线效应. 设置 $L=300$，显著性水平 $\alpha=0.05$. 表 4 展示所得检验统计量均值(Mean of T_n，简记为 Mean-T_n)及经验 p 值(Power). 根据统计量的定义，当原假设成立时，统计量 T_n 应接近于 0，模拟结果与理论一致. 注意到，当折线效应不存在时，经验 P 值接近名义上的显著性水平 0.05，而随着折线效应增强，检验功效增加，并趋近于 1，这说明本文提出的检验统计量能够有效识别折线效应.

表 4　$K=2$ 时各情形在不同折线效应下统计量 T_n 均值及检验功效的势

b_1	情形 1				情形 2				情形 3			
	0.00	−0.10	−0.20	−0.30	0.00	−0.15	−0.30	−0.45	0.00	−0.15	−0.30	−0.45
Mean-T_n	0.128 3	0.224 8	0.351 4	0.548 7	0.425 9	0.598 7	0.982 6	1.280 9	0.218 6	0.348 8	0.556 9	0.695 2
p	0.07	0.33	0.89	1.00	0.05	0.10	0.52	0.96	0.05	0.22	0.73	0.99

本文亦使用所提方法“ESL.IND”“ESL.CHO”分析文献[4]的纵向黄体酮数据，参数的估计值、平均绝对误差、置信区间，以及拟合曲线(结果见 Github). 实证结果显示，相较“SEG”“MKQR”方法，所提估计具有明显竞争优势.

5　证明

为了证明定理 1，我们给出如下引理.

引理 1　若真实折点个数 K_0 已知，且条件(ⅰ)—(ⅴ)成立，则当 $n \to \infty$ 时 $\boldsymbol{\eta} \xrightarrow{\mathrm{p}} \boldsymbol{\eta}_0$.

引理 1 的证明　根据文献[16]结论，证 $\hat{\boldsymbol{\eta}}$ 为 $\boldsymbol{\eta}_0$ 的一致估计量，只需证明存在常数 $C > 0$，对任意 ε，

$$p\Big(\sup_{\|\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}_0\| \leqslant Cn^{-\frac{1}{2}}} (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0)^{\mathrm{T}} \boldsymbol{U}(\hat{\boldsymbol{\xi}}) < 0\Big) > 1 - \varepsilon \tag{16}$$

(16)式等价于

$$p\left(\sup_{\|\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}_0\| \leqslant Cn^{-\frac{1}{2}}} ((\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}_0)^{\mathrm{T}}, (\hat{\boldsymbol{\varrho}} - \boldsymbol{\varrho}_0)^{\mathrm{T}}, (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}}, (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)^{\mathrm{T}}) \begin{pmatrix} \boldsymbol{U}_1(\hat{\boldsymbol{\xi}}) \\ \boldsymbol{U}_2(\hat{\boldsymbol{\xi}}) \\ \boldsymbol{U}_{1,3}(\hat{\boldsymbol{\xi}}) \\ \boldsymbol{U}_{2,3}(\hat{\boldsymbol{\xi}}) \end{pmatrix} < 0\right) > 1 - \varepsilon$$

下面证明 $\hat{\boldsymbol{\vartheta}}$ 的渐近一致性，$\hat{\boldsymbol{\rho}}$，$\hat{\boldsymbol{\varrho}}$，$\hat{\boldsymbol{\mu}}$ 的性质证明可类似进行. 使用中值定理，经计算，有

$$\begin{aligned}
&(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}} \boldsymbol{U}_{1,3}(\hat{\boldsymbol{\xi}}) = \\
&(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}} \sum_{i=1}^{n} \tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}} \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0) \psi_\gamma(\boldsymbol{e}_{i,0}) + (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}} \sum_{i=1}^{n} \tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}} [\boldsymbol{V}_i^{-1}(\hat{\boldsymbol{\xi}}) - \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0)] \psi_\gamma(\boldsymbol{e}_{i,0}) + \\
&(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}} \sum_{i=1}^{n} \tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}} \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0) \boldsymbol{E}_i (\tilde{e}_i - \boldsymbol{e}_{i,0}) + (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}} \sum_{i=1}^{n} \tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}} [\boldsymbol{V}_i^{-1}(\hat{\boldsymbol{\xi}}) - \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0)] \boldsymbol{E}_i (\tilde{e}_i - \boldsymbol{e}_{i,0}) = \\
&I_1 + I_2 + I_3 + I_4
\end{aligned}$$

其中 $\tilde{e}_i^*$ 位于 $\tilde{e}_i$ 及 $\boldsymbol{e}_{i,0}$ 之间，$\boldsymbol{E}_i = \mathrm{diag}(\psi'_\gamma(\tilde{e}_i^*))$. 对于 $\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0 = CO_p(n^{-\frac{1}{2}})$，条件(ⅰ)成立时，显然有 $\mathrm{E}(I_1) = 0$，$\mathrm{E}(I_2) = 0$. 且在条件(ⅰ)—(ⅳ)成立时，有

$$\sum_{i=1}^{n} \tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}} \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0) \boldsymbol{\Sigma}_{i,0} \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0) \tilde{\boldsymbol{\chi}}_{1,i} = O_p(n)$$

以及

$$\sum_{i=1}^{n} \tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}} [\boldsymbol{V}_i^{-1}(\hat{\boldsymbol{\xi}}) - \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0)] \boldsymbol{\Sigma}_{i,0} [\boldsymbol{V}_i^{-1}(\hat{\boldsymbol{\xi}}) - \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0)] \tilde{\boldsymbol{\chi}}_{1,i} = C^2 O_p(1)$$

另一方面，由于 $\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0 = CO_p(n^{-\frac{1}{2}})$，因此，$I_1 = CO_p(1)$，$I_2 = C^2 O_p(n^{-\frac{1}{2}})$. 注意到，$\tilde{e}_{ij}^{(0)} - e_{ij,0} = \tilde{\boldsymbol{\chi}}_{1,ij}(\boldsymbol{\vartheta}_0 - \hat{\boldsymbol{\vartheta}}) + \sum_{k=1}^{K} b_{k,0}\{(X_{ij} - \tau_{k,0})_+ - (X_{ij} - \tau_k^{(0)})_+ + (\tau_{k,0} - \tau_k^{(0)}) I(X_{ij} > \tau_k^{(0)})\}$，记 $b_{k,0}\{(X_{ij} - \tau_{k,0})_+ - (X_{ij} - \tau_k^{(0)})_+ + (\tau_{k,0} - \tau_k^{(0)}) I(X_{ij} > \tau_k^{(0)})\}$ 为 $\Delta_{ij,k}$，因此，

$$\begin{aligned}
I_3 = {} & (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}} \sum_{i=1}^{n} \tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}} \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0) \boldsymbol{E}_i \tilde{\boldsymbol{\chi}}_{1,i} (\boldsymbol{\vartheta}_0 - \hat{\boldsymbol{\vartheta}}) + (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}} \sum_{i=1}^{n} \tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}} \boldsymbol{V}_i^{-1}(\boldsymbol{\xi}_0) \boldsymbol{E}_i \tilde{\boldsymbol{\chi}}_{1,i} \boldsymbol{\Delta}_i = \\
& I_{3,1} + I_{3,2}
\end{aligned}$$

其中 $\boldsymbol{\Delta}_{i,k} = (\sum_{k=1}^{K} \Delta_{i1,k}, \cdots, \sum_{k=1}^{K} \Delta_{im_i,k})^{\mathrm{T}}$. 当条件(ⅰ)—(ⅳ)成立时，有 $I_{3,1} = -C^2 O_p(1)$. 经计算，$|\Delta_{ij}| < |b_{k,0}||\tau_{k,0} - \tau_k^{(0)}|$，因此，当条件(ⅴ)成立时，$\Delta_{ij,k} = O_p(n^{-\frac{1}{2}})$. 由条件(ⅰ)—(ⅴ)可得 $I_{3,2} = CO_p(1)$. 因此，$I_3 = -C^2 O_p(1)$. 类似可得，$I_4 = C^3 O_p(n^{-\frac{1}{2}})$. 综合上述结论，可得

$$(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}} \boldsymbol{U}_{1,3}(\hat{\boldsymbol{\xi}}) = CO_p(1) + C^2 O_p(n^{-\frac{1}{2}}) - C^2 O_p(1) + C^3 O_p(n^{-\frac{1}{2}})$$

因此，存在常数 $C > 0$ 使得 $(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^{\mathrm{T}} \boldsymbol{U}_{1,3}(\hat{\boldsymbol{\xi}}) < 0$. 故而引理 1 可证.

定理 1 的证明　下面证明 $\hat{\boldsymbol{\vartheta}}$ 渐近服从正态分布，$\hat{\boldsymbol{\mu}}$，$\hat{\boldsymbol{\rho}}$，$\hat{\boldsymbol{\varrho}}$ 的证明可类似进行. 由中值定理，根据估计方程(8)所得的估计量 $\hat{\boldsymbol{\vartheta}}$ 满足

$$\sqrt{n}(\hat{\boldsymbol{\vartheta}}-\boldsymbol{\vartheta}_0)=-\left[\frac{\partial \boldsymbol{U}_{1,3}(\boldsymbol{\xi})}{\partial \boldsymbol{\vartheta}}\bigg|_{\xi=\xi^*}\right]^{-1}\times \boldsymbol{U}_{1,3}(\boldsymbol{\xi}_0)=$$

$$\left[\frac{1}{n}\sum_{i=1}^{n}\tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}\boldsymbol{\Lambda}_i(\boldsymbol{\eta})\tilde{\boldsymbol{\chi}}_{1,i}\bigg|_{\xi=\xi^*}\right]^{-1}\times\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}\psi_\gamma(\boldsymbol{Y}_i-\tilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})\bigg|_{\xi=\xi_0}\right]$$

其中 $\boldsymbol{\xi}^*$ 位于 $\hat{\boldsymbol{\xi}}$ 与 $\boldsymbol{\xi}_0$ 之间，因此，据引理 1，$n\to\infty$ 时 $\boldsymbol{\xi}^*\xrightarrow{\mathrm{p}}\boldsymbol{\xi}_0$. 由连续映射定理可知，

$$\frac{1}{n}\sum_{i=1}^{n}\tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}\boldsymbol{\Lambda}_i(\boldsymbol{\eta})\tilde{\boldsymbol{\chi}}_{1,i}\bigg|_{\xi=\xi^*}\xrightarrow{\mathrm{p}}\boldsymbol{A}^{33} \tag{17}$$

记 $\boldsymbol{U}_{1,3}^{(i)}(\boldsymbol{\xi}_0)=\tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}\psi_\gamma(\boldsymbol{Y}_i-\tilde{\boldsymbol{\chi}}_{1,i}\boldsymbol{\eta})\,|_{\xi=\xi_0}$，条件(ⅰ)成立时，其期望 $\mathrm{E}\boldsymbol{U}_{1,3}^{(i)}(\boldsymbol{\xi}_0)=\boldsymbol{0}$，协方差阵

$$\mathrm{Cov}(\boldsymbol{U}_{1,3}^{(i)}(\boldsymbol{\xi}_0))=\tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}\tilde{\boldsymbol{\chi}}_{1,i}$$

条件(ⅱ)—(ⅲ)的成立保证了存在常数 κ_0 使得 $\mathrm{Cov}(\boldsymbol{U}_{1,3}^{(i)}(\boldsymbol{\xi}_0))\leqslant\kappa_0\boldsymbol{I}_{m_i\times m_i}$，因此 $\sum_{i=1}^{\infty}\frac{\mathrm{Cov}(\boldsymbol{U}_{1,3}^{(i)}(\boldsymbol{\xi}_0))}{i^2}<\infty$. 应用李雅普洛夫中心极限定理，有

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\boldsymbol{\chi}}_{1,i}^{\mathrm{T}}\boldsymbol{V}_i^{-1}\psi_\gamma(\boldsymbol{Y}_i-\tilde{\boldsymbol{\chi}}_i\boldsymbol{\eta})\,|_{\xi=\xi_0}\xrightarrow{\mathrm{d}}N(\boldsymbol{0},\boldsymbol{\Gamma}^{33}) \tag{18}$$

结合(17)式及(18)式，

$$\sqrt{n}(\hat{\boldsymbol{\vartheta}}-\boldsymbol{\vartheta}_0)\xrightarrow{\mathrm{d}}N(\boldsymbol{0},(\boldsymbol{A}^{33})^{-1}\boldsymbol{\Gamma}^{33}(\boldsymbol{A}^{33})^{-1})$$

故而定理 1 可证.

定理 2 的证明 记折点位置真值为 $\boldsymbol{\tau}_0=(\tau_{1,0},\cdots,\tau_{K,0})^{\mathrm{T}}$. 由(9)式可得等式

$$\tau_k-\tau_{k,0}=\frac{\hat{\mu}_k-\mu_{k,0}}{b_{k,0}}+(\hat{\mu}_k-\mu_{k,0})\left(\frac{1}{\hat{b}_k}-\frac{1}{b_{k,0}}\right)+\mu_{k,0}\left(\frac{1}{\hat{b}_k}-\frac{1}{b_{k,0}}\right)$$

根据定理 1，$\hat{\mu}_k$ 与 $\hat{b}_k$ 分别为 $\mu_{k,0}$，$b_{k,0}$ 的 $\sqrt{n}$ 相合估计，因此

$$\sqrt{n}(\tau_k-\tau_{k,0})=\frac{\sqrt{n}(\hat{\mu}_k-\mu_{k,0})}{b_{k,0}}+o_p(1)$$

根据 Slutsky 定理可得 $\sqrt{n}(\hat{\tau}_k-\tau_{k,0})$ 与 $\frac{1}{b_{k,0}}\sqrt{n}(\hat{\mu}_k-\mu_{k,0})$ 具有相同的渐近分布. 定理 2 可证.

定理 3 的证明 为证 $\hat{K}$ 为折点个数真值 K_0 的相合估计，我们只需证明当 $n\to\infty$ 时

$$P(\min_{K\neq K_0}\mathrm{BIC}(K)>\mathrm{BIC}(K_0))\to 1$$

经计算，

$$\mathrm{BIC}(K)-\mathrm{BIC}(K_0)=\log\left(\frac{S_n(\hat{\boldsymbol{\eta}}_K)}{S_n(\hat{\boldsymbol{\eta}}_{K_0})}\right)+\frac{\log n}{n}(K-K_0)C_n=$$

$$\log\left(1+\frac{S_n(\hat{\boldsymbol{\eta}}_K)-S_n(\hat{\boldsymbol{\eta}}_{K_0})}{S_n(\hat{\boldsymbol{\eta}}_{K_0})}\right)+\frac{\log n}{n}(K-K_0)C_n=$$

$$\log\left(1+\frac{\{S_n(\hat{\boldsymbol{\eta}}_K)-S_n(\boldsymbol{\eta}_0)\}+\{S_n(\boldsymbol{\eta}_0)-S_n(\hat{\boldsymbol{\eta}}_{K_0})\}}{S_n(\hat{\boldsymbol{\eta}}_{K_0})}\right)+\frac{\log n}{n}(K-K_0)C_n$$

根据条件(ⅵ)，有 $S_n(\hat{\boldsymbol{\eta}}_K)-S_n(\boldsymbol{\eta}_0)\geqslant 0$. 对 $S_n(\boldsymbol{\eta}_0)-S_n(\hat{\boldsymbol{\eta}}_{K_0})$ 使用泰勒展开，

$$S_n(\boldsymbol{\eta}_0)-S_n(\hat{\boldsymbol{\eta}}_{K_0})=\sum_{i=1}^{n}\sum_{j=1}^{m_i}\{\varphi_\gamma(Y_{ij}-\tilde{\boldsymbol{\chi}}_{ij}\boldsymbol{\eta}_0)(\boldsymbol{\eta}_0-\hat{\boldsymbol{\eta}}_{K_0})+\mathrm{o}(\boldsymbol{\eta}_0-\hat{\boldsymbol{\eta}}_{K_0})\}$$

由定理 1 知 $\hat{\boldsymbol{\eta}}_{K_0}$ 为 $\boldsymbol{\eta}_0$ 的 $\sqrt{n}$ 相合估计，因此，$S_n(\boldsymbol{\eta}_0)-S_n(\hat{\boldsymbol{\eta}}_{K_0})=\mathrm{o}(n^{\frac{1}{2}})$. 计算可得

$$\mathrm{BIC}(K)-\mathrm{BIC}(K_0)\geqslant\mathrm{o}(n^{-\frac{1}{2}})+\frac{\log n}{n}(K-K_0)C_n$$

考虑 $K>K_0$ 和 $K<K_0$ 两种情况. $K>K_0$ 时，由于 $K-K_0>0$，$n\to\infty$ 时显然有 $\mathrm{BIC}(K)-\mathrm{BIC}(K_0)>0$ 依概率趋于 1；对于 $K<K_0$，条件(ⅶ) 成立时，当 $n\to\infty$ 时

$$\mathrm{BIC}(K)-\mathrm{BIC}(K_0)\geqslant \mathrm{o}(n^{-\frac{1}{2}})-\mathrm{o}\left(C_n\frac{\log n}{n}\right)>0$$

定理 3 得证.

6　总结

本文基于指数平方损失提出纵向多折点回归模型参数估计和统计推断方法. 为处理折点回归模型，本文首先基于局部线性平滑方法将折点回归模型转化为普通的纵向线性模型. 然后为融合纵向数据中重复观测间的相关性，本文利用修正的 Cholesky 分解方法对重复观测间的协方差阵进行建模，以提高回归模型参数的估计效率. 本文讨论了参数估计的大样本性质，并同时讨论了指数平方损失函数中调谐参数的选择、折点个数的确定方法和折线效应的检验问题等. 数值模拟和实证分析结果显示本文所提方法可以为纵向多折点回归模型的参数提供更为稳健的、有效的估计.

参考文献：

[1] LERMAN P M. Fitting Segmented Regression Models by Grid Search [J]. Journal of the Royal Statistical Society Series C：Applied Statistics，1980，29(1)：77-84.

[2] HINKLEY D，CHAPMAN P，RUNGER G. Change-Point Problems [R]. Minnesota：University of Minnesota，1980.

[3] MUGGEO V M R. Estimating Regression Models with Unknown Break-Points [J]. Statistics in Medicine，2003，22(19)：3055-3071.

[4] ZHONG W，WAN C，ZHANG W Y. Estimation and Inference for Multi-Kink Quantile Regression [J]. Journal of Business & Economic Statistics，2022，40(3)：1123-1139.

[5] SHA N. On Testing the Change-Point in the Longitudinal Bent Line Quantile Regression Model [D]. Columbia：Columbia University，2011.

[6] WAN C，ZHONG W，ZHANG W Y，et al. Multikink Quantile Regression for Longitudinal Data with Application to Progesterone Data Analysis [EB/OL]. (2021-12-20)[2022-08-02]. https：//arxiv. org/pdf/2112. 05045. pdf.

[7] QU A N，LINDSAY B G，LI B. Improving Generalised Estimating Equations Using Quadratic Inference Functions [J]. Biometrika，2000，87(4)：823-836.

[8] YE H J，PAN J X. Modelling of Covariance Structures in Generalised Estimating Equations for Longitudinal Data [J]. Biometrika，2006，93(4)：927-941.

[9] HUBER P J. Robust Estimation of a Location Parameter [J]. The Annals of Mathematical Statistics，1964，35(1)：73-101.

[10] JURECKOVA J. Nonparametric Estimate of Regression Coefficients [J]. The Annals of Mathematical Statistics，1971，42(4)：1328-1338.

[11] FOX M，RUBIN H. Admissibility of Quantile Estimates of a Single Location Parameter [J]. The Annals of Mathematical Statistics，1964，35(3)：1019-1030.

[12] WANG X Q，JIANG Y L，HUANG M，et al. Robust Variable Selection with Exponential Squared Loss [J]. Journal of the American Statistical Association，2013，108(502)：632-643.

[13] WANG K N，LIN L. Robust Structure Identification and Variable Selection in Partial Linear Varying Coefficient Models [J]. Journal of Statistical Planning and Inference，2016，174：153-168.

[14] SONG Y Q，JIAN L，LIN L. Robust Exponential Squared Loss-Based Variable Selection for High-Dimensional Single-Index Varying-Coefficient Model [J]. Journal of Computational and Applied Mathematics，2016，308：330-345.

[15] LV J，GUO C H，WU J B. Subject-Wise Empirical Likelihood Inference for Robust Joint Mean-Covariance Model with Longitudinal Data [J]. Statistics and Its Interface，2019，12(4)：617-630.

[16] WANG L. GEE Analysis of Clustered Binary Data with Diverging Number of Covariates [J]. The Annals of Statistics，2011，39(1)：389-417.

责任编辑　张枸