arXiv:2603.04752v1 [stat.ME] 5 Mar 2026

基于 γ-散度的扩散过程稳健估计

Tomoyuki Nakagawa[1,2] 和 Yusuke Shimizu[3]

[1]数据科学学院，MeiseiUniversity[*2]理化学研究所脑科学中心统计数学部[†3]城西大学数学系[‡]

最后更新：2026年3月6日

摘要

本文探讨扩散过程的高频观测数据中离群值的问题。由于离群值的存在可能导致扩散过程的统计推断出现错误，因此需要稳健估计方法。为构造稳健估计量，我们首先采用 Kessler方法将扩散过程的转移密度近似为高斯密度，然后采用两种类型的最小稳健散度估计方法。本文给出了基于 γ-散度的稳健估计量的渐近性质。此外，我们还推导了基于散度进行估计的条件影响函数，并讨论其有界性。

关键词：扩散过程；稳健统计； γ-散度；影响函数

数学学科分类：主分类62F35；次分类60J60

1 引言

本文探讨扩散过程高频观测数据中的离群值问题。扩散过程广泛应用于物理科学、生物科学、金融和工程学等领域。过去几十年来，许多论文致力于扩散过程的统计推断；参见Kessler (1997)；Masuda和Shimizu (2017)；Uchida和Yoshida (2012)；Yoshida (2011)及其中的参考文献。尤其是，离散观测扩散过程的估计受到广泛关注。具体而言，Kessler Kessler(1997)提出了扩散过程的准似然函数，并推导出最小对比估计量的渐近分布。然而，众所周知，基于似然的估计量会受到离群值或极端值的强烈影响。为解决这一问题，人们开发了多种稳健估计方法；参见La Vecchia和Trojani(2010)；Yoshida(1988)及其中的参考文献。在这些方法中，本文特别考虑将Basu等人提出的密度幂散度，以及 γ-散度用于扩散过程的稳健估计，后者由Jones等人(2001)提出。抗离群值稳健性已从许多方面得到研究，例如影响

[*]2-1-1 Hodokubo，Hino，东京 191-8506，日本
[†]2-1 Hirosawa 和光市，埼玉 351-0198 日本
[‡]Keyakidai 1-1，坂户市，埼玉 350-0246，日本

函数、崩溃点和再下降性（详情参见Hampel等人 (1986)；Huber和Ronchetti (2009)；Maronna等人 (2019等）。最近，Lee和Song (2013)证明了在Ornstein-Uhlenbeck型过程中，使用密度幂散度的稳健估计量具有渐近性质。类似地，Song等人 (2007)和Song(2017)证明了在离散观测扩散过程中，使用密度幂散度估计离散参数的稳健估计量具有渐近性质。然而，他们几乎没有提及密度幂散度估计量的稳健性。

本文旨在提出一种用于离散观测扩散过程的基于散度的稳健估计量。具体而言，我们展示了采用密度幂散度和 γ-散度的两种稳健估计量的稳健性（例如有界影响和再下降性等）。本文结构如下：第 2 节介绍离散观测扩散过程中基于密度幂散度和 γ-散度的估计。第3 节讨论所提出估计量的一致性和渐近正态性。第 4 节推导基于密度幂散度和 La Vecchia 和 Trojani (2010年) 定义的条件影响函数，并展示其稳健性。第 γ-散度的估计，并展示其稳健性。第 5 节考虑若干用于时间序列离群值的概率模型，包括 Maronna 等人(2019) 所述的加性离群值（AOs）和替换离群值（ROs），并使用蒙特卡罗方法比较基于稳健散度的估计与基于似然的估计。

2 稳健散度估计

考虑基于散度的估计方法，该方法用于评估任意两个概率分布之间的差异。基于散度的估计方法已成功用于构造稳健估计量。综述可参见Pardo (2006年), Cichocki 和 Amari (2010年），及其参考文献。特别地，密度幂散度与 γ-散度是众所周知的稳健散度，并在各类文献中具有良好的稳健性（例如有界影响和再下降性等）（参见 Fujisawa 和 Eguchi (2008年); Momozaki 和 Nakagawa (2022年）等）。

2.1 稳健散度

本节介绍密度幂散度与 γ-散度。设 f 和 g 为概率密度函数。

Basu 等人(1998年）提出了如下定义的密度幂散度

$$\begin{aligned}\mathcal{D}_\alpha(g,f) = &\frac{1}{\alpha(\alpha+1)}\int_\Omega g(x)^{\alpha+1}\mathrm{d}x - \frac{1}{\alpha}\int_\Omega g(x)f(x)^\alpha \mathrm{d}x \\ &+ \frac{1}{(1+\alpha)}\int_\Omega f(x)^{1+\alpha}\mathrm{d}x, \quad (\alpha>0).\end{aligned} \tag{1}$$

该散度将 Kullback-Leibler（KL）散度与 L_2-距离作为特例。Jones 等人(2001年）也提出了 γ-散度，其定义为

$$\begin{aligned}\mathcal{D}_\gamma(g,f) = &\frac{1}{\gamma(\gamma+1)}\int_\Omega g(x)^{\gamma+1}\mathrm{d}x - \frac{1}{\gamma}\log\int_\Omega g(x)f(x)^\gamma \mathrm{d}x \\ &+ \frac{1}{(1+\gamma)}\log\int_\Omega f(x)^{1+\gamma}\mathrm{d}x, \quad (\gamma>0).\end{aligned} \tag{2}$$

这里， γ-散度包含KL散度。

对于一族具有密度 $\{f_\theta\}$的参数分布 $\{F_\theta \mid \theta \in \Theta\}$，以及一个密度为 g的分布 G，我们将最小密度幂散度泛函与 γ-散度泛函 $T_\alpha(\cdot)$和 $T_\gamma(\cdot)$定义为

$$T_\alpha(g) = \underset{\theta\in\Theta}{\operatorname{argmin}}\, \mathcal{D}_\alpha(g, f_\theta), \quad T_\gamma(g) = \underset{\theta\in\Theta}{\operatorname{argmin}}\, \mathcal{D}_\gamma(g, f_\theta).$$

通过最小化密度幂散度和 γ-散度经验版本得到的估计量，对密度幂散度和 γ-散度进行最小化时，对于 $T_\alpha(g)$和$T_\gamma(g)$是弱相合的，并且具有渐近正态性；研究还证明，该估计量针对离群值和基础模型的误设具有很强的稳健性（详情请参见Basu等人（1998）；Jones等人（2001）；Fujisawa和Eguchi（2008）等）。

这些方法可以扩展到回归模型（详情请参见Basu等人（1998）；Fujisawa和Eguchi（2008）及其中的参考文献）。设 $\{f_\theta(y \mid x)\}$ 为由参数 $\theta \in \Theta$ 索引的回归模型参数族，并设 $g(y \mid x)$是在给定 $X = x$的条件下 Y的真实密度。分别用 $f_\theta(y \mid x)$和 $g(y \mid x)$替换式（1）中的 f 和 g，即可得到如下 x-条件版本的密度幂散度族：

$$\begin{aligned}\mathcal{D}_\alpha\left(g(\cdot \mid x), f(\cdot \mid x)\right) = &\frac{1}{\alpha(\alpha+1)}\int_\Omega g(y \mid x)^{\alpha+1}\mathrm{d}y - \frac{1}{\alpha}\int_\Omega g(y \mid x) f(y \mid x)^\alpha \mathrm{d}y \\ &+ \frac{1}{(1+\alpha)}\int_\Omega f(y \mid x)^{1+\alpha}\mathrm{d}y, \quad (\alpha > 0).\end{aligned} \tag{3}$$

同样地，将 f 与 g 在（2）中分别替换为 f_θ（$y \mid x$）与 g（$y \mid x$），即可得到一族 x条件版本的 γ散度，如下所示：

$$\begin{aligned}\mathcal{D}_\gamma\left(g(\cdot \mid x), f(\cdot \mid x)\right) = &\frac{1}{\gamma(\gamma+1)}\log\int_\Omega g(y \mid x)^{\gamma+1}\mathrm{d}y - \frac{1}{\gamma}\log\int_\Omega g(y \mid x) f_\theta(y \mid x)^\gamma \mathrm{d}y \\ &+ \frac{1}{(1+\gamma)}\log\int_\Omega f_\theta(y \mid x)^{1+\gamma}\mathrm{d}y, \quad (\gamma > 0).\end{aligned} \tag{4}$$

利用（3）和（4），我们考虑在下一节将这些方法推广到扩散过程。

2.2 使用稳健散度估计随机微分方程

考虑一个一维遍历扩散过程，它由维纳过程驱动的随机微分方程描述：

$$dX_t = b(X_t, \mu)dt + a(X_t, \sigma)dw_t, \quad X_0 = x_0, \tag{5}$$

其中，$\{w_t\}_{t\geq 0}$ 是一维维纳过程，x_0 是与 $\{w_t\}_{t\geq 0}$独立的随机变量，$\sigma \in \Theta_\sigma$ 和 $\mu \in \Theta_\mu$ 是未知参数，$\Theta_\sigma \subset \mathbb{R}^p$ 和 $\Theta_\mu \subset \mathbb{R}^q$ 是紧集，且 $a: \mathbb{R}\times\mathbb{R}^p \to \mathbb{R}$ 和 $b: \mathbb{R}\times\mathbb{R}^q \to \mathbb{R}$ 是已知函数。$\theta_0 := (\mu_0, \sigma_0)$ 是 $\theta := (\mu, \sigma)$ 的真值。数据为离散观测$\{X_{t_i^n}\}_{i=0,1,\ldots,n}$，其中$t_i^n = ih_n$（$h_n > 0$为离散化步长）。我们将考虑 $h_n \to 0$、$nh_n \to \infty$ 和 $nh_n^2 \to 0$ $(n \to \infty)$ 的情形。

Lee和Song (2013) 和 Song (2017年) 给出了扩散过程的密度幂交叉熵，如下所示：

$$Q_{n,\alpha}(\theta) = \sum_{i=1}^n \left\{-\frac{1}{\alpha}f_\theta(x_{t_i^n} \mid x_{t_{i-1}^n})^\alpha + \frac{1}{\alpha+1}\int f_\theta(x \mid x_{t_{i-1}^n})^{1+\alpha}\mathrm{d}x\right\} = \sum_{i=1}^n q_{\alpha,i}(\theta),$$

其中

$$q_{\alpha,i}(\theta)=\begin{cases}\dfrac{1}{(2\pi a_{i-1}(\sigma)^2h_n)^{\alpha/2}}\left[-\dfrac{1}{\alpha}\exp\left\{-\dfrac{\alpha(X_{t_i^n}-X_{t_{i-1}^n}-h_nb_{i-1}(\mu))^2}{2a_{i-1}(\sigma)^2h_n}\right\}+\dfrac{1}{(1+\alpha)^{3/2}}\right], & (\alpha>0),\\ \dfrac{1}{2}\log(2\pi a_{i-1}(\sigma)^2)+\dfrac{\left(X_{t_i^n}-X_{t_{i-1}^n}-h_nb_{i-1}(\mu)\right)^2}{2a_{i-1}(\sigma)^2h_n}, & (\alpha=0),\end{cases}$$

$a_{i-1}(\sigma)=a(X_{t_{i-1}^n},\sigma)$以及 $b_{i-1}(\mu)=b(X_{t_{i-1}^n},\mu)$。随后，他们利用密度幂散度证明了估计量的一致性和渐近正态性。此外，Ghosh等人(2022年)讨论了其在贝叶斯背景下的指数收敛性。

类似地，我们将扩散过程模型的 γ-交叉熵定义如下：

$$Q_{n,\gamma}(\theta)=\sum_{i=1}^{n}\frac{-f_\theta(X_{t_i^n}\mid X_{t_{i-1}^n})^\gamma}{\left\{\int f_\theta(x\mid X_{t_{i-1}^n})^{1+\gamma}\mathrm{d}x\right\}^{\gamma/(1+\gamma)}}=\sum_{i=1}^{n}q_{\gamma,i}(\theta),$$

其中

$$q_{\gamma,i}(\theta)=\begin{cases}-\left(\dfrac{1+\gamma}{2\pi a_{i-1}(\sigma)^2h_n}\right)^{\gamma/2(1+\gamma)}\exp\left\{-\dfrac{\gamma(X_{t_i^n}-X_{t_{i-1}^n}-h_nb_{i-1}(\mu))^2}{2a_{i-1}(\sigma)^2h_n}\right\}, & (\gamma>0),\\ \dfrac{1}{2}\log(2\pi a_{i-1}(\sigma)^2)+\dfrac{\left(X_{t_i^n}-X_{t_{i-1}^n}-h_nb_{i-1}(\mu)\right)^2}{2a_{i-1}(\sigma)^2h_n}, & (\gamma=0).\end{cases}\tag{6}$$

需要注意的是，$q_{0,i}(\theta)=\lim_{\alpha\to0}q_{\alpha,i}(\theta)=\lim_{\gamma\to0}q_{\gamma,i}(\theta)$。

3 渐近性质

本节将推导基于 γ-散度的估计量的渐近性质。我们考虑随机微分方程

$$dX_t=b(X_t,\mu)dt+\sigma dw_t,\quad X_0=x_0,\tag{7}$$

其中 $a(x,\sigma)=\sigma$ 在（5）中。接下来，定义估计量 $\hat{\theta}_n^{(\gamma)}=(\hat{\mu}_n^{(\gamma)},\hat{\sigma}_n^{(\gamma)})$，其对应于 θ 如下

$$Q_{n,\gamma}(\hat{\theta}_n^{(\gamma)})=\inf_{\theta\in\Theta}\sum_{i=1}^{n}q_{\gamma,i}(\theta),$$

其中

$$q_{\gamma,i}(\theta)=\begin{cases}-\sigma^{-\gamma/(1+\gamma)}\exp\left\{-\dfrac{\gamma(X_{t_i^n}-X_{t_{i-1}^n}-h_nb_{i-1}(\mu))^2}{2\sigma^2h_n}\right\}, & (\gamma>0),\\ \log\sigma^2+\dfrac{\left(X_{t_i^n}-X_{t_{i-1}^n}-h_nb_{i-1}(\mu)\right)^2}{\sigma^2h_n}, & (\gamma=0).\end{cases}$$

注意，我们省略了（6）中的常数项。令

$$\mathscr{P}=\left\{f(x,\mu)\;\middle|\;|f|\le C(1+|x|)^C\text{ for some }C\right\},$$

其中 C 不依赖于 θ。然后，我们作出以下假设：

(A1) 存在一个常数 $C > 0$，使得对所有 $x, y \in \mathbb{R}$, $|b(x, \mu_0) - b(y, \mu_0)| \leq C|x - y|$。

(A2) 过程 X 关于 (μ_0, σ_0) 是遍历的，其不变测度为 ν_0，且满足 $\int x^k d\nu_0(x) < \infty$，对所有 $k \geq 0$. 均成立

(A3) 对所有 $k > 0$，$\sup_t \mathbb{E}[|X_t|^k] < \infty$。

(A4) 函数 b 关于 x 连续可微，对所有 μ 均成立，且其导数属于 $\mathscr{P}$。

(A5) 函数 b 及其所有关于 x 的导数，关于 μ 均三次可微，对所有 x 均成立。此外，这些关于 μ 的三阶及以下导数属于 $\mathscr{P}$。

(A6) 若 $b(x, \mu) = b(x, \mu_0)$ 对 ν_0 几乎处处且对所有 x 均成立，则 $\mu = \mu_0$。

(A7) $S := \int \partial_\mu b(x, \mu_0) \partial_{\mu^\top} b(x, \mu_0) d\nu_0(x)$ 是正定的，其中 $\partial_\mu b = \partial b / \partial \mu$。

下面的定理给出了 $\hat{\theta}_n^{(\gamma)}$ 的渐近性质。

定理 3.1。 假设 *(A1)–(A7)*，$h_n \to 0$，$nh_n \to \infty$ 并且 $nh_n^2 \to 0$。那么，对于任意 $\gamma \geq 0$,

$$\left(\sqrt{nh_n}(\hat{\mu}_n^{(\gamma)} - \mu_0), \sqrt{n}(\hat{\sigma}_n^{(\gamma)} - \sigma_0) \right) \xrightarrow{\mathcal{L}} N_{p+1}(0, \Sigma_0^{(\gamma)}),$$

其中

$$\Sigma_0^{(\gamma)} = \sigma_0^2 \begin{pmatrix} \left(\frac{1+\gamma}{\sqrt{1+2\gamma}} \right)^3 S^{-1} & 0 \\ 0 & \frac{(1+\gamma)^3(3\gamma^2+4\gamma+2)}{4(1+2\gamma)^{5/2}} \end{pmatrix}.$$

证明见附录 A.1 中的定理3.1的证明。

注 1。 随机微分方程（7）可扩展为以下形式：

$$dX_t = b(X_t, \mu)dt + \sigma a(X_t)dw_t. \tag{8}$$

由于由变换 $Y_t = H(X_t)$，其中 H 满足 $\partial_x H(x) = 1/a(x)$，我们得到

$$dY_t = M(Y_t, \mu, \sigma)dt + \sigma dw_t,$$

其中

$$M(y, \mu, \sigma) = \frac{b\left(H^{-1}(y), \mu\right)}{a\left(H^{-1}(y)\right)} - \frac{\sigma^2}{2} \partial_x a\left(H^{-1}(y), \sigma\right),$$

and随后，针对（8）的估计过程可简化为针对随机微分方程（7）的估计过程。

4 稳健性

在本节中，我们讨论针对扩散过程的无穷小稳健统计程序。统计文献普遍认识到，稳健统计方法十分必要。重要的研究探讨了针对独立同分布数据的无穷小稳健（即具有有界影响函数的）估计量；例如，Hampel 等人(1986); Huber 和 Ronchetti (2009).

La Vecchia 和 Trojani (2010) 定义了扩散过程的影响函数和条件影响函数。设

$$G_{\varepsilon,\nu} = (1-\varepsilon)G_{\theta_0} + \varepsilon\nu$$

其中 ν 是污染分布，构成参数扩散概率 G_{θ_0}的一般 ε-污染，其中 $0 \le \varepsilon \le \eta$ 针对固定的 $0 < \eta < 1$。

定义 4.1 （条件影响函数）。 对于条件无偏估计量 $\hat{\ }\ \theta\ (\cdot)$ （即（5）中扩散过程的估计量）的影响函数 (IF) 是任意函数 IF（$\cdot$， $\cdot$; θ_0），使得

$$\partial_\varepsilon \hat{\theta}(G_{\varepsilon,\nu})\Big|_{\varepsilon=0} = \int \mathrm{IF}(x_1, x_0; \theta_0)\nu(\mathrm{d}x_0, \mathrm{d}x_1)$$

对所有 $\nu \in \mathcal{M}$成立，其中 $\mathcal{M}$ 是平稳过程的二维边缘分布族。此外，条件影响函数（IF^c）是任意影响函数 IF^c（$\cdot$， $\cdot$; θ_0），使得 $\mathbb{E}[\mathrm{IF}^c(X_{t_i^n}, X_{t_{i-1}^n}; \theta_0) \mid X_{t_{i-1}^n}] = 0$。

此外，La Vecchia 和 Trojani （2010年）证明了如下定理。

定理 4.2 （推论 5，见La Vecchia 和 Trojani，2010年）。 设 θ 为条件无偏 M-估计量，用于扩散过程，由可积的估计函数 ψ所定义。则 θ 的条件影响函数由下式唯一给出

$$\mathrm{IF}^c_\psi(x_i, x_{i-1}; \theta_0) = -D(\psi, \theta_0)^{-1}\psi(x_i, x_{i-1}; \theta_0)$$

其中 $D(\psi, \theta_0) = E_{\theta_0}\left[\partial_\theta \psi(X_i, X_{i-1}; \theta_0)\right]$。

由定理4.2，我们得到稳健散度的条件影响函数为

$$\begin{aligned}
\mathrm{IF}^c_{\psi_\alpha}(x_i, x_{i-1}; \theta_0) &= -D(\psi_\alpha, \theta_0)^{-1}\psi_\alpha(x_i, x_{i-1}; \theta_0), \\
\mathrm{IF}^c_{\psi_\gamma}(x_i, x_{i-1}; \theta_0) &= -D(\psi_\gamma, \theta_0)^{-1}\psi_\gamma(x_i, x_{i-1}; \theta_0).
\end{aligned}$$

此处 ψ_α 与 ψ_γ 分别为密度幂散度和 γ散度的估计函数，并给出如下：

$$\begin{aligned}
\psi_\alpha(x_i, x_{i-1}; \theta) &= (\psi_{\alpha,1}(x_i, x_{i-1}; \theta), \psi_{\alpha,2}(x_i, x_{i-1}; \theta))^\top, \\
\psi_\gamma(x_i, x_{i-1}; \theta) &= (\psi_{\gamma,1}(x_i, x_{i-1}; \theta), \psi_{\gamma,2}(x_i, x_{i-1}; \theta))^\top,
\end{aligned}$$

其中

$$\begin{aligned}
\psi_{\alpha,1}(x_i, x_{i-1}; \theta) &= \partial_\mu q_{\alpha,i}(\theta) = -w_i^{(\alpha)} \frac{(x_i - x_{i-1} - h_n b_{i-1}(\mu))}{a_{i-1}(\sigma)^2}\partial_\mu b_{i-1}(\mu), \\
\psi_{\alpha,2}(x_i, x_{i-1}; \theta) &= \partial_\sigma q_{\alpha,i}(\theta) \\
&= \left\{ w_i^{(\alpha)} - \frac{\alpha}{(1+\alpha)^{3/2}} \frac{1}{(2\pi a_{i-1}(\sigma)^2 h_n)^{\alpha/2}} \right\} \frac{1}{a_{i-1}(\sigma)}\partial_\sigma a_{i-1}(\sigma) \\
&\quad - w_i^{(\alpha)} \frac{(x_i - x_{i-1} - h_n b_{i-1}(\mu))^2}{a_{i-1}(\sigma)^3 h_n}\partial_\sigma a_{i-1}(\sigma), \\
\psi_{\gamma,1}(x_i, x_{i-1}; \theta) &= \partial_\mu q_{\gamma,i}(\theta) = -w_i^{(\gamma)} \frac{(x_i - x_{i-1} - h_n b_{i-1}(\mu))}{a_{i-1}(\sigma)^2}\partial_\mu b_{i-1}(\mu), \\
\psi_{\gamma,2}(x_i, x_{i-1}; \theta) &= \partial_\sigma q_{\gamma,i}(\theta) \\
&= \frac{1}{\gamma+1} w_i^{(\gamma)} \frac{1}{a_{i-1}(\sigma)}\partial_\sigma a_{i-1}(\sigma) \\
&\quad - w_i^{(\gamma)} \frac{(x_i - x_{i-1} - h_n b_{i-1}(\mu))^2}{a_{i-1}(\sigma)^3 h_n}\partial_\sigma a_{i-1}(\sigma),
\end{aligned}$$

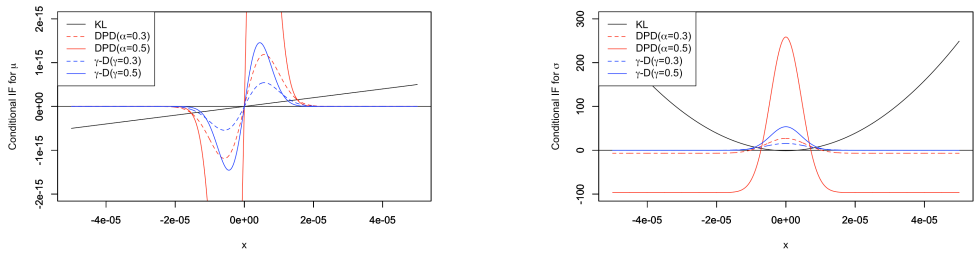


图1：左图为模型A中 μ的条件影响函数，右图为 σ 的条件影响函数，其中 $n = 100$、$h_n = n^{-0.55}$、 $x_{i-1} = h_n$、 $\mu_0 = 1.0$ 和$\sigma_0 = 1.0$。横轴表示 x_i。

and

$$w_i^{(\alpha)} = (2\pi h_n a_{i-1}(\sigma)^2)^{-\alpha/2} \exp\left\{-\frac{\alpha(x_i - x_{i-1} - h_n b_{i-1}(\mu))^2}{2a_{i-1}(\sigma)^2 h_n}\right\},$$
$$w_i^{(\gamma)} = \left(\frac{1+\gamma}{2\pi h_n a_{i-1}(\sigma)^2}\right)^{\gamma/2(1+\gamma)} \exp\left\{-\frac{\gamma(x_i - x_{i-1} - h_n b_{i-1}(\mu))^2}{2a_{i-1}(\sigma)^2 h_n}\right\}.$$

在图1中，黑线表示最大似然估计的条件影响函数，红线表示基于密度幂散度的估计量的条件影响函数，蓝线表示基于 γ-散度的估计量的条件影响函数。从图1中可以很容易看出，最大似然估计的条件影响函数无界。另一方面，基于密度幂散度和 γ-散度的估计量的条件影响函数似乎有界。在图1中，我们注意到，基于 γ-散度的估计量的条件影响函数似乎具有再下降性质，如 $x_i \to \infty$。

5 模拟研究

本节展示在离散观测扩散过程中使用散度进行稳健估计的表现。这里，我们设离散化步长为 $h_n = n^{-0.55}$。在每种情形下，我们考虑以下两种扩散模型。

模型A： $a(x, \sigma) = \sigma$， $b(x, \mu) = -\mu x$。

模型B： $a(x, \sigma) = 1 + \sigma/(1 + x^2)$ $b(x, \mu) = -\mu x$。

注意，模型A 是奥恩斯坦-乌伦贝克过程。我们在每种情形下采用两种离群值生成结构。样本量 n 和散度参数 α， γ的配置为 $n = 50$、100、200、500和 α、 $\gamma = 0.0$、0。3、0。5。对于每种配置和每个模型，各表中的偏差和均方误差（MSE）均通过重复2000次的蒙特卡罗模拟获得。

5.1 情景 1

将 AO 结构视为时间序列的离群值生成器。AO 模型的一个特例最初由 Fox（1972）提出。设观测值 $\{Y_{t_i^n}\}_{i=0,1,\ldots,n}$ 是我们感兴趣的宽平稳“核心”过程，$\{Z_{t_i^n}\}_{i=0,1,\ldots,n}$ 是平稳离群过程，$\{R_{t_i^n}\}_{i=0,1,\ldots,n}$ 是一个零一过程，且 $\Pr(R_{t_i^n} = 0) = 1 - \varepsilon$

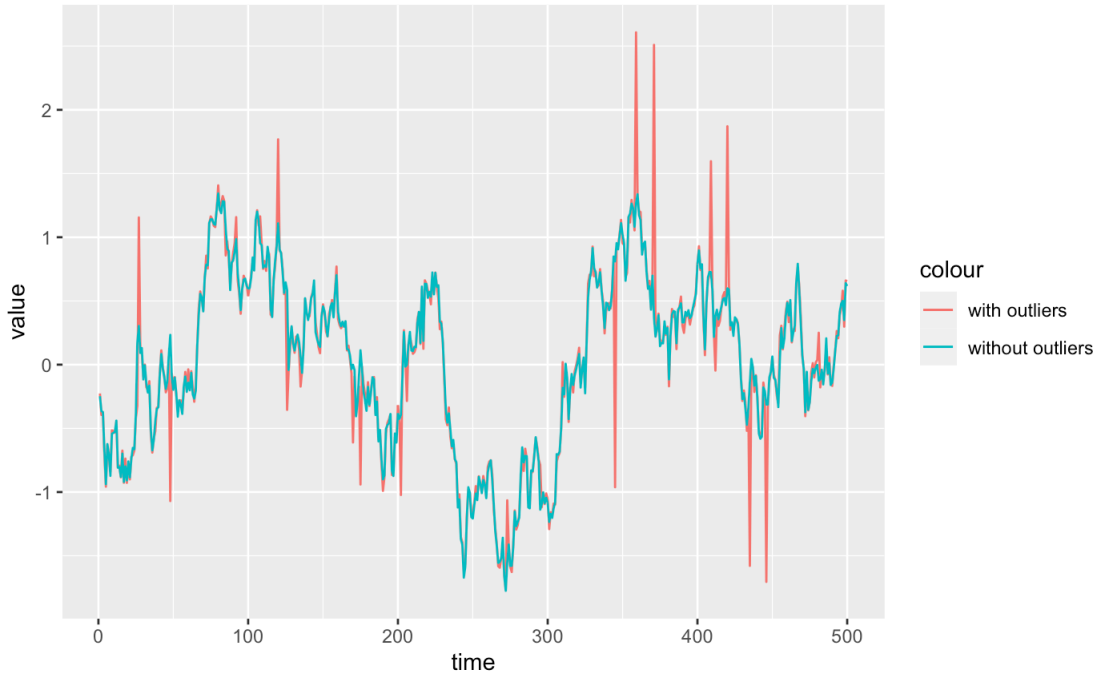


图2：蓝线和红线是模型的样本路径 A ，其中 $\sigma_z^2 = 1.0$，$n = 500$，$\mu_0 = 1.0$ 和 $\sigma_0 = 1.0$。蓝线是无离群值的样本路径，红线也是无离群值的样本路径。

且 $\Pr(R_{t_i^n} = 1) = \varepsilon$。在实践中，比例 ε 通常为正但较小。在加性离群模型下，实际观测到的不是 $Y_{t_i^n}$，而是

$$Y_{t_i^n} = X_{t_i^n} + R_{t_i^n} Z_{t_i^n},$$

其中，过程 $\{X_{t_i^n}\}_{i=0,\ 1,\ldots,n}$，$\{Z_{t_i^n}\}_{i=0,\ 1,\ldots,n}$，以及 $\{R_{t_i^n}\}_{i=0,\ 1,\ldots,n}$ 均假设彼此独立。因此，当 $Y_{t_i^n}$ 时，$R_{t_i^n} = 1$是离群值。在此情形下，我们假设 $Z_{t_i^n}$独立同分布于正态分布 $N(0,\ \sigma_z^2)$，其中 $\sigma_z^2 = 1.0$，1。5，且 $\varepsilon = 0.05$。样本路径如图 2 所示。实际上，观测到的是红色样本路径，而蓝色样本路径未被观测到。我们关注的是，能否利用红色样本路径估计蓝色样本路径的生成过程。

表1 和 3 给出了各模型和配置下的偏差。此外，表 2 和 4 给出了各模型和配置下的均方误差。从表1–4可以看出，最大似然估计受到离群值的显著影响。另一方面，基于密度幂和 γ-散度的估计量受离群值的影响较小。我们还可以看到，在不存在离群值时，基于密度幂和 γ-散度的估计量与最大似然估计具有几乎相同的准确性。此外，从表 2 和4可以看出，在不存在离群值时，估计量的均方误差会随着 n 增加而减小，也就是说，在不存在离群值时，这些估计量是相合的。当包含离群值时，最大似然估计的均方误差不但没有减小，反而会随着样本量 n 增大而增加。因此，最大似然估计并不相合。另一方面，根据定理3.1，即使包含离群值，基于密度幂和 γ-散度的估计量仍然相合。因此，从表 2 和 4可以看出，这些估计量的均方误差会随着 n 增加而减小。

表 1：模型的经验偏差 A 在情景 1 中，条件为 $\mu_0 = 1.0$ 且 $\sigma_0 = 1.$ 0.

		无离群值		$\sigma_z^2 = 1.0$		$\sigma_z^2 = 1.5$	
n	α, γ	μ	σ	μ	σ	μ	σ
				MLE			
50	0.0	−0.0720	0.2383	0.2222	1.0408	0.3227	1.3100
100	0.0	−0.0452	0.1806	0.3889	1.4168	0.5380	1.8764
200	0.0	−0.0300	0.1322	0.5991	1.9954	0.8104	2.7239
500	0.0	−0.0174	0.1063	0.9268	3.1066	1.2292	4.3395
				基于密度幂散度的估计量			
50	0.3	−0.0704	0.2407	0.0475	0.6322	0.0545	0.6750
100	0.3	−0.0446	0.1807	0.0792	0.5497	0.0788	0.5546
200	0.3	−0.0299	0.1246	0.0922	0.4537	0.0858	0.4318
500	0.3	−0.0170	0.1057	0.0957	0.3662	0.0865	0.3366
50	0.5	−0.0727	0.2417	0.0162	0.5161	0.0180	0.5208
100	0.5	−0.0460	0.1792	0.0466	0.4287	0.0453	0.4195
200	0.5	−0.0306	0.1200	0.0603	0.3327	0.0569	0.3149
500	0.5	−0.0171	0.1055	0.0705	0.2725	0.0661	0.2537
				基于γ-散度的估计量			
50	0.3	−0.0704	0.2406	0.0432	0.6289	0.0481	0.6682
100	0.3	−0.0446	0.1809	0.0729	0.5448	0.0702	0.5485
200	0.3	−0.0299	0.1245	0.0834	0.4486	0.0748	0.4252
500	0.3	−0.0170	0.1055	0.0845	0.3613	0.0734	0.3315
50	0.5	−0.0734	0.2418	0.0047	0.5100	0.0028	0.5115
100	0.5	−0.0464	0.1791	0.0323	0.4234	0.0274	0.4122
200	0.5	−0.0308	0.1195	0.0423	0.3272	0.0355	0.3079
500	0.5	−0.0172	0.1054	0.0490	0.2675	0.0418	0.2489

表 2：模型的经验均方误差A 在情景 1 中，当 $\mu_0 = 1.0$ 且 $\sigma_0 = 1.$ 0.

		无离群值		$\sigma_z^2 = 1.0$		$\sigma_z^2 = 1.5$	
n	α, γ	μ	σ	μ	σ	μ	σ
				MLE			
50	0.0	0.0146	0.5570	0.0639	1.6402	0.1187	2.2732
100	0.0	0.0066	0.3880	0.1579	2.3954	0.2961	3.9091
200	0.0	0.0034	0.2529	0.3624	4.2347	0.6601	7.6726
500	0.0	0.0012	0.1624	0.8602	9.8134	1.5121	18.9940
			基于密度幂散度的估计量				
50	0.3	0.0161	0.6051	0.0267	1.4553	0.0308	1.6694
100	0.3	0.0074	0.4135	0.0192	1.0768	0.0199	1.1419
200	0.3	0.0038	0.2659	0.0156	0.6937	0.0143	0.6610
500	0.3	0.0014	0.1682	0.0118	0.3928	0.0100	0.3620
50	0.5	0.0186	0.6568	0.0221	1.2373	0.0231	1.3035
100	0.5	0.0085	0.4437	0.0131	0.8766	0.0130	0.8772
200	0.5	0.0043	0.2860	0.0094	0.5224	0.0088	0.4996
500	0.5	0.0016	0.1800	0.0072	0.3097	0.0065	0.2939
			基于 γ-散度的估计量				
50	0.3	0.0162	0.6051	0.0261	1.4481	0.0297	1.6495
100	0.3	0.0074	0.4137	0.0180	1.0674	0.0183	1.1287
200	0.3	0.0038	0.2659	0.0138	0.6841	0.0123	0.6490
500	0.3	0.0014	0.1681	0.0097	0.3874	0.0078	0.3569
50	0.5	0.0192	0.6578	0.0223	1.2307	0.0231	1.2845
100	0.5	0.0087	0.4439	0.0120	0.8688	0.0116	0.8657
200	0.5	0.0044	0.2853	0.0074	0.5166	0.0067	0.4924
500	0.5	0.0016	0.1800	0.0046	0.3068	0.0038	0.2915

表 3：模型的经验偏差 B 在情景 1 中，满足 $\mu_0 = 1.0$ 且 $\sigma_0 = 1.$ 0.

		无离群值		$\sigma_z^2 = 1.0$		$\sigma_z^2 = 1.5$	
n	α, γ	μ	σ	μ	σ	μ	σ
				MLE			
50	0.0	−0.1672	0.1515	0.2730	0.6262	0.4734	0.8472
100	0.0	−0.1090	0.1171	0.5286	0.8322	0.8361	1.2111
200	0.0	−0.0730	0.0843	0.8792	1.1909	1.3658	1.8420
500	0.0	−0.0423	0.0713	1.4996	1.9372	2.3352	3.1722
				基于密度幂散度的估计量			
50	0.3	−0.1487	0.0475	0.0690	0.2639	0.0965	0.2907
100	0.3	−0.1010	0.0046	0.1587	0.2304	0.1765	0.2398
200	0.3	−0.0705	−0.0356	0.2158	0.1833	0.2238	0.1802
500	0.3	−0.0414	−0.0426	0.2623	0.1557	0.2566	0.1425
50	0.5	−0.1444	0.0023	0.0389	0.1699	0.0543	0.1825
100	0.5	−0.0994	−0.0440	0.1159	0.1286	0.1240	0.1303
200	0.5	−0.0693	−0.0854	0.1604	0.0753	0.1618	0.0688
500	0.5	−0.0403	−0.0891	0.1998	0.0523	0.1938	0.0410
				基于γ-散度的估计量			
50	0.3	−0.1483	0.0459	0.0640	0.2593	0.0889	0.2850
100	0.3	−0.1007	0.0029	0.1516	0.2261	0.1662	0.2349
200	0.3	−0.0704	−0.0372	0.2051	0.1790	0.2093	0.1753
500	0.3	−0.0413	−0.0438	0.2468	0.1515	0.2368	0.1378
50	0.5	−0.1433	−0.0025	0.0267	0.1601	0.0369	0.1707
100	0.5	−0.0987	−0.0486	0.0990	0.1190	0.1009	0.1197
200	0.5	−0.0689	−0.0899	0.1359	0.0655	0.1301	0.0581
500	0.5	−0.0399	−0.0930	0.1659	0.0425	0.1528	0.0305

表4：情景 1 中模型 B 在 $\mu_0 = 1.0$ 和 $\sigma_0 = 1.$ 0.

		无离群值		$\sigma_z^2 = 1.0$		$\sigma_z^2 = 1.5$	
n	α, γ	μ	σ	μ	σ	μ	σ
				MLE			
50	0.0	0.0976	0.3761	0.1721	0.7683	0.3217	1.0938
100	0.0	0.0466	0.2586	0.3260	0.9512	0.7456	1.7253
200	0.0	0.0240	0.1643	0.7970	1.5827	1.8894	3.5575
500	0.0	0.0090	0.1052	2.2577	3.8579	5.4623	10.1679
				基于密度幂散度的估计量			
50	0.3	0.1014	0.3438	0.1350	0.5846	0.1573	0.6600
100	0.3	0.0493	0.2265	0.0971	0.4319	0.1094	0.4559
200	0.3	0.0259	0.1420	0.0891	0.2657	0.0955	0.2647
500	0.3	0.0099	0.0888	0.0877	0.1613	0.0854	0.1552
50	0.5	0.1138	0.3554	0.1366	0.5400	0.1474	0.5814
100	0.5	0.0560	0.2343	0.0857	0.3736	0.0907	0.3889
200	0.5	0.0292	0.1515	0.0661	0.2160	0.0677	0.2135
500	0.5	0.0112	0.0963	0.0569	0.1283	0.0546	0.1256
				基于γ-散度的估计量			
50	0.3	0.1016	0.3435	0.1329	0.5789	0.1530	0.6507
100	0.3	0.0495	0.2263	0.0940	0.4273	0.1044	0.4498
200	0.3	0.0260	0.1420	0.0838	0.2624	0.0881	0.2612
500	0.3	0.0099	0.0888	0.0793	0.1592	0.0749	0.1530
50	0.5	0.1167	0.3550	0.1372	0.5332	0.1460	0.5693
100	0.5	0.0574	0.2341	0.0823	0.3675	0.0853	0.3815
200	0.5	0.0298	0.1518	0.0584	0.2123	0.0577	0.2098
500	0.5	0.0115	0.0967	0.0441	0.1262	0.0398	0.1237

表5：情景2中模型 A 的经验偏差，其中 $\mu_0 = 1.0$ 和 $\sigma_0 = 1.0$。

		无离群值		$\sigma_z^2 = 1.0$		$\sigma_z^2 = 1.5$	
n	α, γ	μ	σ	μ	σ	μ	σ
				MLE			
50	0.0	−0.0720	0.2383	0.2823	1.3795	0.3726	1.6356
100	0.0	−0.0452	0.1806	0.5010	1.9756	0.6337	2.4232
200	0.0	−0.0300	0.1322	0.7577	2.8292	0.9417	3.5368
500	0.0	−0.0174	0.1063	1.1807	4.5580	1.4397	5.7696
				基于密度幂散度的估计量			
50	0.3	−0.0704	0.2407	0.0561	0.6537	0.0594	0.6806
100	0.3	−0.0446	0.1807	0.0798	0.5390	0.0767	0.5332
200	0.3	−0.0299	0.1246	0.0870	0.4211	0.0808	0.4012
500	0.3	−0.0170	0.1057	0.0864	0.3331	0.0798	0.3157
50	0.5	−0.0727	0.2417	0.0211	0.5221	0.0211	0.5262
100	0.5	−0.0460	0.1792	0.0465	0.4105	0.0437	0.3991
200	0.5	−0.0306	0.1200	0.0575	0.3088	0.0547	0.2963
500	0.5	−0.0171	0.1055	0.0659	0.2523	0.0630	0.2437
				基于γ散度的估计量			
50	0.3	−0.0704	0.2406	0.0503	0.6489	0.0522	0.6749
100	0.3	−0.0446	0.1809	0.0716	0.5339	0.0667	0.5265
200	0.3	−0.0299	0.1245	0.0763	0.4144	0.0686	0.3955
500	0.3	−0.0170	0.1055	0.0735	0.3279	0.0657	0.3110
50	0.5	−0.0734	0.2418	0.0074	0.5151	0.0042	0.5171
100	0.5	−0.0464	0.1791	0.0292	0.4043	0.0237	0.3918
200	0.5	−0.0308	0.1195	0.0367	0.3027	0.0316	0.2893
500	0.5	−0.0172	0.1054	0.0419	0.2477	0.0372	0.2393

5.2 情景2

将RO结构视为时间序列的离群值生成器。RO模型的形式为

$$Y_{t_i^n} = (1 - R_{t_i^n})X_{t_i^n} + R_{t_i^n} Z_{t_i^n},$$

其中，$R_{t_i^n}$是一个零一过程，满足 $\Pr(R_{t_i^n} = 0) = 1 - \varepsilon$ 和 $\Pr(R_{t_i^n} = 1) = \varepsilon$；$Z_{t_i^n}$ 是一个“替换”过程，不一定独立于 $X_{t_i^n}$。与情景 1 类似，我们假设 $Z_{t_i^n}$独立同分布于正态分布 $N(0, \sigma_z^2)$，其中 $\sigma_z^2 = 1.0, 1.5$，且 $\varepsilon = 0.05$。

从表5–8可以看出，与情景 1 类似，最大似然估计受到离群值的显著影响。另一方面，基于密度幂和 γ-散度的估计量受离群值的影响很小。我们还发现，在没有离群值时，基于密度幂和γ-散度的估计量与最大似然估计具有几乎相同的准确性。

6 结论性说明

本文介绍了使用散度对扩散过程进行稳健估计。我们提出了 γ-散度估计方法，用于扩散过程模型，这与Lee和Song(2013)；Song等人(2007)；Song(2017)类似，并证明了一致性和

表6：模型的经验均方误差A 在情景2中，当 $\mu_0 = 1.0$ 且 $\sigma_0 = 1.$ 0.

		无离群值		$\sigma_z^2 = 1.0$		$\sigma_z^2 = 1.5$	
n	α, γ	μ	σ	μ	σ	μ	σ
				MLE			
50	0.0	0.0146	0.5570	0.0943	2.4601	0.1534	3.2323
100	0.0	0.0066	0.3880	0.2576	4.2911	0.4082	6.2597
200	0.0	0.0034	0.2529	0.5775	8.2573	0.8903	12.7616
500	0.0	0.0012	0.1624	1.3953	20.9381	2.0740	33.4507
				基于密度幂散度的估计量			
50	0.3	0.0161	0.6051	0.0312	1.5489	0.0338	1.7288
100	0.3	0.0074	0.4135	0.0202	1.0342	0.0200	1.0718
200	0.3	0.0038	0.2659	0.0142	0.6244	0.0130	0.5973
500	0.3	0.0014	0.1682	0.0098	0.3570	0.0086	0.3386
50	0.5	0.0186	0.6568	0.0233	1.2671	0.0235	1.3408
100	0.5	0.0085	0.4437	0.0134	0.8215	0.0130	0.8216
200	0.5	0.0043	0.2860	0.0087	0.4861	0.0084	0.4737
500	0.5	0.0016	0.1800	0.0064	0.2915	0.0060	0.2835
				基于γ-散度的估计量			
50	0.3	0.0162	0.6051	0.0300	1.5393	0.0325	1.7164
100	0.3	0.0074	0.4137	0.0186	1.0250	0.0181	1.0584
200	0.3	0.0038	0.2659	0.0123	0.6137	0.0109	0.5886
500	0.3	0.0014	0.1681	0.0076	0.3519	0.0065	0.3345
50	0.5	0.0192	0.6578	0.0229	1.2566	0.0231	1.3234
100	0.5	0.0087	0.4439	0.0120	0.8149	0.0114	0.8122
200	0.5	0.0044	0.2853	0.0066	0.4810	0.0062	0.4671
500	0.5	0.0016	0.1800	0.0037	0.2894	0.0033	0.2818

表7：模型的经验偏差 B 在情景2中，当 $\mu_0 = 1.0$ 且 $\sigma_0 = 1.$ 0.

		无离群值		$\sigma_z^2 = 1.0$		$\sigma_z^2 = 1.5$	
n	α, γ	μ	σ	μ	σ	μ	σ
				MLE			
50	0.0	−0.1672	0.1515	0.6031	0.9204	0.8144	1.1395
100	0.0	−0.1090	0.1171	1.1254	1.2922	1.4698	1.6862
200	0.0	−0.0730	0.0843	1.7963	1.8345	2.3346	2.5017
500	0.0	−0.0423	0.0713	3.0908	2.9636	4.0263	4.2140
				基于密度幂散度的估计量			
50	0.3	−0.1487	0.0475	0.1346	0.3011	0.1407	0.3075
100	0.3	−0.1010	0.0046	0.2136	0.2403	0.2113	0.2383
200	0.3	−0.0705	−0.0356	0.2615	0.1805	0.2488	0.1680
500	0.3	−0.0414	−0.0426	0.2826	0.1395	0.2624	0.1243
50	0.5	−0.1444	0.0023	0.0837	0.1868	0.0850	0.1857
100	0.5	−0.0994	−0.0440	0.1501	0.1253	0.1471	0.1219
200	0.5	−0.0693	−0.0854	0.1865	0.0642	0.1765	0.0549
500	0.5	−0.0403	−0.0891	0.2095	0.0372	0.1971	0.0268
				基于γ-散度的估计量			
50	0.3	−0.1483	0.0459	0.1259	0.2977	0.1298	0.3029
100	0.3	−0.1007	0.0029	0.2011	0.2371	0.1964	0.2344
200	0.3	−0.0704	−0.0372	0.2437	0.1768	0.2286	0.1637
500	0.3	−0.0413	−0.0438	0.2575	0.1352	0.2354	0.1200
50	0.5	−0.1433	−0.0025	0.0638	0.1783	0.0612	0.1757
100	0.5	−0.0987	−0.0486	0.1219	0.1170	0.1148	0.1124
200	0.5	−0.0689	−0.0899	0.1478	0.0548	0.1339	0.0447
500	0.5	−0.0399	−0.0930	0.1589	0.0275	0.1437	0.0165

表8：情景2下模型 B 无离群值时的经验均方误差 $\mu_0 = 1.0$ 和 $\sigma_0 = 1.$ 0.

		无离群值		$\sigma_z^2 = 1.0$		$\sigma_z^2 = 1.5$	
n	α, γ	μ	σ	μ	σ	μ	σ
				MLE			
50	0.0	0.0976	0.3761	0.4614	1.2232	0.7609	1.6746
100	0.0	0.0466	0.2586	1.3131	1.9285	2.2070	3.1019
200	0.0	0.0240	0.1643	3.2508	3.5298	5.4742	6.4226
500	0.0	0.0090	0.1052	9.5617	8.8883	16.2201	17.8629
				基于密度幂散度的估计量			
50	0.3	0.1014	0.3438	0.1841	0.6493	0.1923	0.6924
100	0.3	0.0493	0.2265	0.1343	0.4228	0.1357	0.4358
200	0.3	0.0259	0.1420	0.1193	0.2599	0.1121	0.2528
500	0.3	0.0099	0.0888	0.1004	0.1548	0.0885	0.1454
50	0.5	0.1138	0.3554	0.1613	0.5560	0.1634	0.5682
100	0.5	0.0560	0.2343	0.1040	0.3577	0.1037	0.3641
200	0.5	0.0292	0.1515	0.0784	0.2115	0.0738	0.2068
500	0.5	0.0112	0.0963	0.0611	0.1241	0.0555	0.1198
				基于γ-散度的估计量			
50	0.3	0.1016	0.3435	0.1785	0.6453	0.1851	0.6846
100	0.3	0.0495	0.2263	0.1272	0.4204	0.1273	0.4322
200	0.3	0.0260	0.1420	0.1088	0.2579	0.1007	0.2503
500	0.3	0.0099	0.0888	0.0859	0.1528	0.0740	0.1439
50	0.5	0.1167	0.3550	0.1575	0.5506	0.1583	0.5597
100	0.5	0.0574	0.2341	0.0956	0.3546	0.0939	0.3598
200	0.5	0.0298	0.1518	0.0641	0.2094	0.0589	0.2046
500	0.5	0.0115	0.0967	0.0417	0.1229	0.0365	0.1187

所提出估计量的渐近正态性。我们推导了基于稳健散度的估计量的条件影响函数，并在数值研究中考察了有界性。

致谢

本研究受日本学术振兴会青年科学家科研资助金资助，项目编号为JP23K13019和JP23K16852。

附录

A.1 定理证明 3.1

我们假设 $\gamma > 0$ ，因为 $\gamma = 0$ 情形的证明与 $\gamma > 0$的证明类似。首先，我们将证明

$$\frac{Q_{n,\gamma}(\theta)}{n} \xrightarrow{p} -\sigma^{-\gamma/(1+\gamma)} \left(1 + \frac{\gamma\sigma_0^2}{\sigma^2}\right)^{-1/2} \tag{A.1}$$

关于 θ 一致成立。由欧拉近似可得

$$X_{t_i^n} = X_{t_{i-1}^n} + b_{i-1}(\mu_0)h_n + \sigma_0 Z_{n,i}\sqrt{h_n} + \Delta_{n,i},$$

其中

$$Z_{n,i} = \frac{1}{\sqrt{h_n}}(W_{t_i^n} - W_{t_{i-1}^n}), \quad \Delta_{n,i} = \int_{t_{i-1}^n}^{t_i^n} \left(b_s(\mu_0) - b_{i-1}(\mu_0)\right) ds.$$

定义

$$\begin{aligned} K_{n,i}(\mu,\sigma) :=& \frac{\gamma(X_{t_i^n} - X_{t_{i-1}^n} - h_n b_{i-1}(\mu))^2}{2h_n\sigma^2} - \frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2} \\ =& \frac{\gamma B_{i-1}(\mu)^2 h_n}{2\sigma^2} + \frac{\gamma\Delta_{n,i}^2}{2h_n\sigma^2} + \frac{\gamma\sigma_0 B_{i-1}(\mu) Z_{n,i}\sqrt{h_n}}{\sigma^2} + \frac{\gamma B_{i-1}(\mu)\Delta_{n,i}}{\sigma^2} \\ &+ \frac{\gamma\sigma_0 Z_{n,i}\Delta_{n,i}}{\sigma^2\sqrt{h_n}}, \end{aligned}$$

其中 $B_{i-1}(\mu) := b_{i-1}(\mu_0) - b_{i-1}(\mu)$。类似于 Lee和Song(2013) 的定理 1 的证明，标准估计给出

$$\sup_\theta \max_{1\le i\le n} |K_{n,i}(\mu,\sigma)| = o_p(h_n^r), \quad 0 \le r < \frac{1}{2},$$

and

$$\begin{aligned} &\sup_\theta \left| -\sigma^{-\gamma/(1+\gamma)} \frac{1}{n}\sum_{i=1}^n \left[\exp\left\{ -\frac{\gamma\left(X_{t_i^n} - X_{t_{i-1}^n} - h_n b_{i-1}(\mu)\right)^2}{2h_n\sigma^2} \right\} - \exp\left(-\frac{\gamma\sigma_0^2}{2\sigma^2} Z_{n,i}^2 \right) \right] \right| \\ &= \sup_\theta \left| -\sigma^{-\gamma/(1+\gamma)} \frac{1}{n}\sum_{i=1}^n \exp\left(-\frac{\gamma\sigma_0^2}{2\sigma^2} Z_{n,i}^2 \right) \left(e^{-K_{n,i}(\mu,\sigma)} - 1 \right) \right| \\ &\le \sup_\theta \sigma^{-\gamma/(1+\gamma)} \max_{1\le i\le n} \left| e^{-K_{n,i}(\mu,\sigma)} - 1 \right| = o_p(1). \end{aligned}$$

因此，由 Lee和Song (2013) 中的引理4可知，对 θ,

$$\begin{aligned}\frac{Q_{n,\gamma}(\theta)}{n} &= -\sigma^{-\gamma/(1+\gamma)}\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\gamma\left(X_{t_i^n}-X_{t_{i-1}^n}-h_n b_{i-1}(\mu)\right)^2}{2h_n\sigma^2}\right\}\\ &\xrightarrow{p} -\sigma^{-\gamma/(1+\gamma)}\left(1+\frac{\gamma\sigma_0^2}{\sigma^2}\right)^{-1/2}\end{aligned}$$

从而得到（A.1）。注意

$$Q_\gamma(\sigma) := -\sigma^{-\gamma/(1+\gamma)}\left(1+\frac{\gamma\sigma_0^2}{\sigma^2}\right)^{-1/2}$$

在 $\sigma=\sigma_0$ 处取得最小值。

接下来，我们将验证

$$\begin{aligned}&\frac{Q_{n,\gamma}(\theta)}{nh_n}-\frac{Q_{n,\gamma}(\mu_0,\sigma)}{nh_n}\\ &\xrightarrow{p}\frac{\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}}}{2}\left(1+\frac{\gamma\sigma_0^2}{\sigma^2}\right)^{-3/2}\int\left(b(x,\mu)-b(x,\mu_0)\right)^2 d\nu_0(x).\end{aligned}$$

根据泰勒展开，我们有

$$\begin{aligned}&e^{-K_{n,i}(\mu,\sigma)}-e^{-K_{n,i}(\mu_0,\sigma)}\\ =&-\frac{\gamma B_{i-1}(\mu)^2 h_n}{2\sigma^2}-\frac{\gamma\sigma_0 B_{i-1}(\mu)Z_{n,i}\sqrt{h_n}}{\sigma^2}-\frac{\gamma B_{i-1}(\mu)\Delta_{n,i}}{\sigma^2}+\frac{1}{2}K_{n,i}(\mu,\sigma)^2\\ &+\frac{K_{n,i}(\mu_0,\sigma)^2}{2!}e^{\xi_{1,i}}+\frac{K_{n,i}(\mu,\sigma)^3}{3!}e^{\xi_{2,i}}\\ =&-\frac{\gamma\sigma_0 B_{i-1}(\mu)Z_{n,i}\sqrt{h_n}}{\sigma^2}-\frac{\gamma}{2\sigma^2}\left(1-\frac{\gamma\sigma_0^2}{\sigma^2}Z_{n,i}^2\right)B_{i-1}(\mu)^2 h_n+H_{n,i}(\mu,\sigma)\\ &+\frac{K_{n,i}(\mu_0,\sigma)^2}{2!}e^{\xi_{1,i}}+\frac{K_{n,i}(\mu,\sigma)^3}{3!}e^{\xi_{2,i}},\end{aligned}$$

其中 $|\xi_{1,i}|\le|K_{n,i}(\mu_0,\sigma)|$, $|\xi_{2,i}|\le|K_{n,i}(\mu,\sigma)|$, 且

$$H_{n,i}(\mu,\sigma) := -\frac{\gamma B_{i-1}(\mu)\Delta_{n,i}}{\sigma^2}+\frac{1}{2}K_{n,i}(\mu,\sigma)^2-\frac{1}{2}\frac{\gamma^2\sigma_0^2}{\sigma^4}Z_{n,i}^2 B_{i-1}(\mu)^2 h_n.$$

定义

$$R_{n,i}(\mu,\sigma) := H_{n,i}(\mu,\sigma)+\frac{K_{n,i}(\mu_0,\sigma)^2}{2!}e^{\xi_{1,i}}+\frac{K_{n,i}(\mu,\sigma)^3}{3!}e^{\xi_{2,i}}.$$

于是，由引理2、4和5在Lee和Song （2013）中，我们有

$$\begin{aligned}\sup_\theta\max_{1\le i\le n}|H_{n,i}(\mu,\sigma)| &= o_p(h_n^{r_1}),\quad 0\le r_1<1.5,\\ \sup_\theta\max_{1\le i\le n}|K_{n,i}(\mu,\sigma)^3| &= o_p(h_n^{r_2}),\quad 0\le r_2<1.5,\\ \sup_\theta\max_{1\le i\le n}|K_{n,i}(\mu_0,\sigma)^2| &= o_p(h_n^{r_3}),\quad 0\le r_3<2,\\ \sup_\theta\max_{1\le i\le n}|e^{\xi_{1,i}}| &\le \exp\left\{\sup_\theta\max_{1\le i\le n}|K_{n,i}(\mu_0,\sigma)|\right\}=O_p(1),\\ \sup_\theta\max_{1\le i\le n}|e^{\xi_{2,i}}| &\le \exp\left\{\sup_\theta\max_{1\le i\le n}|K_{n,i}(\mu,\sigma)|\right\}=O_p(1),\\ \sup_\theta\max_{1\le i\le n}R_{n,i}(\mu,\sigma) &= o_p(h_n^{r}),\quad 0\le r<1.5,\end{aligned}$$

因此

$$
\begin{aligned}
&\frac{Q_{n,\gamma}(\theta)}{nh_n} - \frac{Q_{n,\gamma}(\mu_0,\sigma)}{nh_n} \\
=&-\frac{\sigma^{-\gamma/(1+\gamma)}}{nh_n}\sum_{i=1}^{n}\exp\left(-\frac{\gamma\sigma_0^2}{2\sigma^2}Z_{n,i}^2\right)\left(e^{-K_{n,i}(\mu,\sigma)}-e^{-K_{n,i}(\mu_0,\sigma)}\right) \\
=&\frac{\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sigma_0}{n\sqrt{h_n}}\sum_{i=1}^{n}B_{i-1}(\mu)Z_{n,i}\exp\left(-\frac{\gamma\sigma_0^2}{2\sigma^2}Z_{n,i}^2\right) \\
&+\frac{\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}}}{2n}\sum_{i=1}^{n}\left(1-\frac{\gamma\sigma_0^2}{\sigma^2}Z_{n,i}^2\right)B_{i-1}(\mu)^2\exp\left(-\frac{\gamma\sigma_0^2}{2\sigma^2}Z_{n,i}^2\right) \\
&-\frac{\sigma^{-\gamma/(1+\gamma)}}{nh_n}\sum_{i=1}^{n}R_{n,i}(\mu,\sigma)\exp\left(-\frac{\gamma\sigma_0^2}{2\sigma^2}Z_{n,i}^2\right) \\
\xrightarrow{p}&\frac{\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}}}{2}\left(1+\frac{\gamma\sigma_0^2}{\sigma^2}\right)^{-3/2}\int\left\{b(x,\mu)-b(x,\mu_0)\right\}^2 d\nu_0(x)
\end{aligned}
$$

关于 θ一致成立。至此完成了$\hat{}\theta_n^{(\gamma)} \xrightarrow{p} \theta_0$的证明。

其次，我们将推导 $\left(\sqrt{nh_n}(\hat{}\mu_n^{(\gamma)}-\mu_0)$的渐近分布，$\sqrt{n}(\hat{}\sigma_n^{(\gamma)}-\sigma_0)\right)$。由泰勒展开，我们得到

$$
0=\begin{pmatrix}\partial_\mu Q_{n,\gamma}(\theta_0)\\ \partial_\sigma Q_{n,\gamma}(\theta_0)\end{pmatrix}+\int_0^1\nabla^2 Q_{n,\gamma}\left(\theta_0+u(\hat{\theta}_n^{(\gamma)}-\theta_0)\right)du\begin{pmatrix}\hat{\mu}_n^{(\gamma)}-\mu_0\\ \hat{\sigma}_n^{(\gamma)}-\sigma_0\end{pmatrix}
$$

由此我们可以写出

$$
L_{n,\gamma}=\int_0^1 C_{n,\gamma}\left(\theta_0+u(\hat{\theta}_n^{(\gamma)}-\theta_0)\right)du\begin{pmatrix}\sqrt{nh_n}(\hat{\mu}_n^{(\gamma)}-\mu_0)\\ \sqrt{n}(\hat{\sigma}_n^{(\gamma)}-\sigma_0)\end{pmatrix},
$$

其中

$$
\begin{aligned}
L_{n,\gamma}&=\begin{pmatrix}-\frac{1}{\sqrt{nh_n}}\partial_\mu Q_{n,\gamma}(\theta_0)\\ -\frac{1}{\sqrt{n}}\partial_\sigma Q_{n,\gamma}(\theta_0)\end{pmatrix},\\
C_{n,\gamma}(\theta)&=\begin{pmatrix}\frac{1}{nh_n}\partial_\mu^2 Q_{n,\gamma}(\theta) & \frac{1}{n\sqrt{h_n}}\partial_{\mu\sigma}^2 Q_{n,\gamma}(\theta)\\ \frac{1}{n\sqrt{h_n}}\partial_{\sigma\mu}^2 Q_{n,\gamma}(\theta) & \frac{1}{n}\partial_\sigma^2 Q_{n,\gamma}(\theta)\end{pmatrix}.
\end{aligned}
$$

我们定义

$$
m_{n,i}^{(\gamma)}(\theta)=-\sigma^{-\gamma/(1+\gamma)}\exp\left\{-\frac{\gamma\left(B_{i-1}(\mu)h_n+\sigma_0 Z_{n,i}\sqrt{h_n}\right)^2}{2h_n\sigma^2}\right\},
$$

and

$$
\zeta_{n,i}=\begin{pmatrix}\frac{1}{\sqrt{nh_n}}\partial_\mu m_{n,i}^{(\gamma)}(\theta_0)\\ \frac{1}{\sqrt{n}}\partial_\sigma m_{n,i}^{(\gamma)}(\theta_0)\end{pmatrix}.
$$

我们注意到

$$\partial_\mu m_{m,i}^{(\gamma)}(\theta_0) = -\sqrt{h_n}\gamma\sigma_0^{-\frac{1+2\gamma}{1+\gamma}} Z_{n,i}\partial_\mu b_{i-1}(\mu_0)e^{-\frac{\gamma}{2}Z_{n,i}^2},$$
$$\partial_\sigma m_{m,i}^{(\gamma)}(\theta_0) = \gamma\sigma_0^{-\frac{1+2\gamma}{1+\gamma}}\left(\frac{1}{1+\gamma} - Z_{n,i}^2\right)e^{-\frac{\gamma}{2}Z_{n,i}^2}.$$

利用以下论证:

$$E_{\theta_0}\left[Z_{n,i}e^{-\frac{\gamma}{2}Z_{n,i}^2}|\mathscr{G}_{i-1}^n\right] = 0,$$
$$E_{\theta_0}\left[Z_{n,i}^2e^{-\gamma Z_{n,i}^2}|\mathscr{G}_{i-1}^n\right] = (1+2\gamma)^{-3/2},$$
$$E_{\theta_0}\left[\left(\frac{1}{1+\gamma} - Z_{n,i}^2\right)e^{-\frac{\gamma}{2}Z_{n,i}^2}|\mathscr{G}_{i-1}^n\right] = 0,$$
$$E_{\theta_0}\left[Z_{n,i}\left(\frac{1}{1+\gamma} - Z_{n,i}^2\right)e^{-\gamma Z_{n,i}^2}|\mathscr{G}_{i-1}^n\right] = 0,$$
$$E_{\theta_0}\left[\left(\frac{1}{1+\gamma} - Z_{n,i}^2\right)^2e^{-\gamma Z_{n,i}^2}|\mathscr{G}_{i-1}^n\right] = \frac{3\gamma^2+4\gamma+2}{(1+2\gamma)^{5/2}(1+\gamma)^2},$$

其中 $\mathscr{G}_i^n$表示由 $\{w_s:\ s \le t_i^n\}$生成的 σ-代数，并且根据 Lee和Song(2013)中的引理 6，我们可以证明

$$\sum_{i=1}^n E_{\theta_0}[\zeta_{n,i}|\mathscr{G}_{i-1}^n] \xrightarrow{p} 0,$$
$$\sum_{i=1}^n E_{\theta_0}[\zeta_{n,i}\zeta_{n,i}^T|\mathscr{G}_{i-1}^n] \xrightarrow{p} \Sigma_1^{(\gamma)},$$
$$\sum_{i=1}^n E_{\theta_0}[|\zeta_{n,i}|^4|\mathscr{G}_{i-1}^n] \xrightarrow{p} 0,$$

其中

$$\Sigma_1^{(\gamma)} = \gamma^2\sigma_0^{-\frac{2(1+2\gamma)}{1+\gamma}}\begin{pmatrix} \frac{1}{(1+2\gamma)^{3/2}}S & 0 \\ 0 & \frac{3\gamma^2+4\gamma+2}{(1+2\gamma)^{5/2}(1+\gamma)^2} \end{pmatrix},$$

并且 S 是假设(A7)中定义的那个。这意味着 $L_{n,\ \gamma} \xrightarrow{\mathcal{L}} N_{p+1}(0,\ \Sigma_1^{(\gamma)})$。

现在，若我们有以下结果，则定理的证明即告完成

$$C_{n,\gamma}(\theta_0) \xrightarrow{p} \sigma_0^{-\frac{2+3\gamma}{1+\gamma}}\begin{pmatrix} \gamma(1+\gamma)^{-3/2}S & 0 \\ 0 & 2\gamma(1+\gamma)^{-5/2} \end{pmatrix}. \tag{A.2}$$

我们注意到

$$
\begin{aligned}
\partial_\mu^2 Q_{n,\gamma}(\theta) =& -\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}\Bigg\{\left(X_{t_i^n}-X_{t_{i-1}^n}-h_n b_{i-1}(\mu)\right)^2\partial_\mu b_{i-1}(\mu)\partial_{\mu^\top} b_{i-1}\overset{(\mu)}{\frac{\gamma}{\sigma^2}}-\\
&-h_n\partial_\mu b_{i-1}(\mu)\partial_{\mu^\top} b_{i-1}(\mu)+\left(X_{t_i^n}-X_{t_{i-1}^n}-h_n b_{i-1}(\mu)\right)\partial_\mu^2 b_{i-1}\Bigg\}^{(\mu)}\\
&\times\exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)\exp\left(-K_{n,i}(\mu,\sigma)\right)\\
=& -\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}\sigma_0\partial_\mu^2 b_{i-1}(\mu)Z_{n,i}\sqrt{h_n}\exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)\\
&-\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}\left\{J_{1,n,i}(\theta)-\frac{\gamma\sigma_0^2}{\sigma^2}\partial_\mu^2 b_{i-1}(\mu)Z_{n,i}^2 B_{i-1}(\mu)\right\}h_n\\
&\times\exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)-\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}R_{1,n,i}(\theta),\\
\partial_\sigma^2 Q_{n,\gamma}(\theta) =& -\frac{\gamma(1+2\gamma)}{(1+\gamma)^2}\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}\exp\left(-K_{n,i}(\mu,\sigma)-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)\\
&+\frac{2(3+5\gamma)}{1+\gamma}\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}\exp\left(-K_{n,i}(\mu,\sigma)-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)\\
&\times\left(K_{n,i}(\mu,\sigma)+\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)\\
&-4\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}\exp\left(-K_{n,i}(\mu,\sigma)-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)\\
&\times\left(K_{n,i}(\mu,\sigma)+\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)^2\\
=&\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}\left(-\frac{\gamma(1+2\gamma)}{(1+\gamma)^2}+\frac{\gamma(3+5\gamma)}{1+\gamma}\frac{\sigma_0^2 Z_{n,i}^2}{\sigma^2}-\frac{\gamma^2\sigma_0^4 Z_{n,i}^4}{\sigma^4}\right)\\
&\times\exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)\\
&+\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}\left(\frac{2(3+5\gamma)}{1+\gamma}K_{n,i}(\mu,\sigma)+J_{3,n,i}(\theta)\right)\exp\left(-\frac{\gamma\sigma_0^2 Z_{n\,i}^2}{2\sigma^2}\right)\\
&+\sigma^{-\frac{2+3\gamma}{1+\gamma}}\sum_{i=1}^{n}R_{2,n,i}(\theta)\exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)K_{n,i}(\theta)e^{\xi},
\end{aligned}
$$

$$
\begin{aligned}
\partial^2_{\mu\sigma} Q_{n,\gamma}(\theta) =& \frac{\gamma^2}{1+\gamma}\sigma^{-\frac{2+3\gamma}{1+\gamma}} \sum_{i=1}^{n} \exp\left(-K_{n,i}(\mu,\sigma) - \frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right) \\
& \times (h_n B_{i-1}(\mu) + \sigma_0 Z_{n,i}\sqrt{h_n} + \Delta_{n,i})\partial_\mu b_{i-1}(\mu) \\
& + 2\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}} \sum_{i=1}^{n} \exp\left(-K_{n,i}(\mu,\sigma) - \frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right) \\
& \times (h_n B_{i-1}(\mu) + \sigma_0 Z_{n,i}\sqrt{h_n} + \Delta_{n,i})\partial_\mu b_{i-1}(\mu) \\
& \times \left(K_{n,i}(\mu,\sigma) + \frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right) \\
& + 2\gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}} \sum_{i=1}^{n} \exp\left(-K_{n,i}(\mu,\sigma) - \frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right) \\
& \times (h_n B_{i-1}(\mu) + \sigma_0 Z_{n,i}\sqrt{h_n} + \Delta_{n,i})\partial_\mu b_{i-1}(\mu) \\
=& \gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}} \sum_{i=1}^{n} \partial_\mu b_{i-1}(\mu) \left(\frac{\gamma}{1+\gamma} + 2 + \frac{\gamma\sigma_0^2 Z_{n,i}^2}{\sigma^2}\right) \sigma_0 Z_{n,i}\sqrt{h_n} \\
& \times \exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right) \\
& + \gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}} \sum_{i=1}^{n} \partial_\mu b_{i-1}(\mu) J_{4,n,i}(\theta) \exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right) \\
& - \gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}} \sum_{i=1}^{n} \partial_\mu b_{i-1}(\mu) R_{3,n,i}(\theta) \exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right) K_{n,i}(\theta) e^{\xi},
\end{aligned}
$$

其中 $|\xi| \leq |K_{n,\ i}(\theta)|$ 和

$$
\begin{aligned}
J_{1,n,i}(\theta) =& B_i(\theta)\partial_\mu^2 b_{i-1}(\mu) + \left(\frac{\gamma\sigma_0^2}{\sigma^2}Z_{n,i}^2 - 1\right)\partial_\mu b_{i-1}(\mu)\partial_{\mu^\top} b_{i-1}(\mu),\\
J_{2,n,i}(\theta) =& \Delta_{n,i}\partial_\mu^2 b_{i-1}(\mu) + \frac{\gamma}{\sigma^2}\left\{2\sigma_0 Z_{n,i}B_{i-1}(\mu)h_n^{3/2} + 2\sigma_0 Z_{n,i}\Delta_{n,i}\sqrt{h_n}\right.\\
&\left. + B_{i-1}(\mu)^2 h_n^2 + 2B_{i-1}(\mu)\Delta_{n,i}h_n + \Delta_{n,i}^2\right\}\partial_\mu b_{i-1}(\mu)\partial_{\mu^\top} b_{i-1}(\mu),\\
J_{3,n,i}(\theta) =& -\frac{4\gamma\sigma_0^2 Z_{n,i}^2}{\sigma^2}K_{n,i}(\theta) - 4K_{n,i}^2(\theta)\\
M_{n,i}(\theta) =& \sigma_0\partial_\mu^2 b_{i-1}(\mu)Z_{n,i}\sqrt{h_n} + J_{1,n,i}(\theta)h_n + J_{2,n,i}(\theta)\\
J_{4,n,i}(\theta) =& 2\sigma_0 Z_{n,i}K_{n,i}(\mu,\sigma)\sqrt{h_n} + (h_n B_{i-1}(\mu) + \Delta_{n,i})\\
&\times\left(\frac{\gamma}{1+\gamma} + 2 + \frac{\gamma\sigma_0^2 Z_{n,i}^2}{\sigma^2} + 2K_{n,i}(\mu,\sigma)\right)\\
R_{1,n,i}(\theta) =& \{J_{2,n,i}(\theta) - (J_{1,n,i}(\theta)h_n + J_{2,n,i}(\theta))\frac{\gamma\sigma_0}{\sigma^2}Z_{n,i}B_{i-1}(\mu)\sqrt{h_n}\}\exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right)\\
&+ M_i(\theta)\left\{\frac{\gamma\sigma_0}{\sigma^2}Z_{n,i}B_{i-1}(\mu)\sqrt{h_n} - K_{n,i}(\theta) + \frac{1}{2}K_{n,i}^2(\theta)e^\xi\right\}\\
&\times\exp\left(-\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2}\right),\\
R_{2,n,i}(\theta) =& -\frac{\gamma(1+2\gamma)}{(1+\gamma)^2} + \frac{2(3+5\gamma)}{1+\gamma}\left(\frac{\gamma\sigma_0^2 Z_{n,i}^2}{2\sigma^2} + K_{n,i}(\theta)\right) - \frac{\gamma^2\sigma_0^4 Z_{n,i}^4}{\sigma^4} + J_{3,n,i}(\theta),\\
R_{3,n,i}(\theta) =& \sigma_0\left(\frac{\gamma}{1+\gamma} + 2 + \frac{\gamma\sigma_0^2 Z_{n,i}^2}{\sigma^2}\right)Z_{n,i}\sqrt{h_n} + J_{4,n,i}(\theta).
\end{aligned}
$$

以与定理2的证明相同的方式，在Lee和Song（2013）中，我们可以得到

$$
\begin{aligned}
\sup_\theta \max_{1\leq i\leq n}|J_{2,n,i}(\theta)| &= o_p(h_n^r), \quad 0 \leq r < 1.5,\\
\sup_\theta \max_{1\leq i\leq n}|J_{3,n,i}(\theta)| &= o_p(h_n^r), \quad 0 \leq r < 0.5,\\
\sup_\theta \max_{1\leq i\leq n}|J_{4,n,i}(\theta)| &= o_p(h_n^r), \quad 0 \leq r < 1,\\
\sup_\theta \max_{1\leq i\leq n}|R_{1,n,i}(\theta)| &= o_p(h_n^r), \quad 0 \leq r < 1.5,\\
\sup_\theta \max_{1\leq i\leq n}|R_{2,n,i}(\theta)K_{n,i}(\mu,\sigma)e^\xi| &= o_p(h_n^r), \quad 0 \leq r < 0.5,\\
\sup_\theta \max_{1\leq i\leq n}|R_{3,n,i}(\theta)K_{n,i}(\mu,\sigma)e^\xi| &= o_p(h_n^r), \quad 0 \leq r < 1.
\end{aligned}
$$

因此，我们得到

$$\frac{1}{nh_n}\partial_\mu^2 Q_{n,\gamma}(\theta) \xrightarrow{p} \gamma\sigma^{-\frac{2+3\gamma}{1+\gamma}}\left(1+\frac{\gamma\sigma_0^2}{\sigma^2}\right)^{-3/2}$$
$$\times\left\{S-\int(b(x,\mu_0)-b(x,\mu))\partial_\mu^2 b(x,\mu)\nu_0(x)\right\},$$
$$\frac{1}{n}\partial_\sigma^2 Q_{n,\gamma}(\theta) \xrightarrow{p} \sigma^{-\frac{2+3\gamma}{1+\gamma}}\left\{-\frac{\gamma(1+2\gamma)}{(1+\gamma)^2}\left(1+\frac{\gamma\sigma_0^2}{\sigma^2}\right)^{-1/2}\right.$$
$$\left.+\frac{\gamma(3+5\gamma)}{1+\gamma}\frac{\sigma_0^2}{\sigma^2}\left(1+\frac{\gamma\sigma_0^2}{\sigma^2}\right)^{-3/2}-\frac{3\gamma^2\sigma_0^4}{\sigma^4}\left(1+\frac{\gamma\sigma_0^2}{\sigma^2}\right)^{-5/2}\right\},$$
$$\frac{1}{n\sqrt{h_n}}\partial_{\mu\sigma} Q_{n,\gamma}(\theta) \xrightarrow{p} 0,$$

在 θ 上一致成立。由此完成了（A.2）的证明，因此定理 3.1得以证明。

参考文献

Basu, A.、Harris, I. R.、Hjort, N. L.和Jones, M.（1998年）。通过最小化密度幂散度进行稳健且高效的估计。*Biometrika*，85(3):549–559。Cichocki, A.和Amari, S.（2010年）。α-β-γ散度族：灵活且稳健的相似性度量。*Entropy*，12(6):1532–1568。Fox, A. J.（1972年）。时间序列中的离群值。皇家统计学会会刊：*B*辑（方法论），34(3):350–363。Fujisawa, H.和Eguchi, S.（2008年）。针对严重污染具有小偏差的稳健参数估计。多元分析杂志，99(9):2053–2081。Ghosh, A.、Majumder, T.和Basu, A.（2022年）。一般稳健贝叶斯伪后验：指数收敛结果及其应用。《中华统计学志》，32:787–823。Hampel, F. R.、Ronchetti, E. M.、Rousseeuw, P. J.和Stahel, W. A.（1986年）。稳健统计。Wiley Online Library。Huber, J.和Ronchetti, E. M.（2009年）。《稳健统计学》第二版。Wiley。Jones, M.、Hjort, N. L.、Harris, I. R.和Basu, A.（2001年）。相关基于密度的最小散度估计量的比较。*Biometrika*，88(3):865–873。Kessler, M.（1997年）。基于离散观测对遍历扩散过程的估计。《斯堪的纳维亚统计学杂志》，24(2):211–229。La Vecchia, D.和Trojani, F.（2010年）。扩散过程的无穷小稳健性。美国统计学会杂志，105(490):703–712。Lee, S.和Song, J.（2013）。扩散过程的最小密度幂散度估计量。《统计数学研究所年刊》，65(2):213–236。Maronna, R. A.、Martin, R. D.、Yohai, V. J.和Salibi´an-Barrera, M.（2019年）。《稳健统计学：理论与方法（含 *R*）》。约翰·威立父子出版公司。

Masuda, H.和Shimizu, Y.（2017年）。多重与混合速率渐近理论下正则化估计中的矩收敛。《统计学数学方法》，26:81–110。Momozaki, T.和Nakagawa, T.（2022年）。通过散度方法实现有序响应模型的抗离群值稳健性。*arXiv*预印本 *arXiv:2209.11965*。Pardo, L.（2006年）。基于散度度量的统计推断。CRC出版社。Song, J.（2017年）。离散观测扩散过程中离散参数的稳健估计。《中华统计学志》，第373–388页。Song, J.-M.、Lee, S.-Y.、Na, O.-Y.和Kim, H.-J.（2007年）。离散观测扩散过程中扩散参数的最小密度幂散度估计量。《统计应用与方法通讯》，14(2):267–280。Uchida, M.和Yoshida, N.（2012年）。基于抽样数据的遍历扩散过程自适应估计。《随机过程及其应用》，122(8):2885–2924。Yoshida, N.（1988年）。扩散过程中的稳健M估计量。《统计数学研究所年刊》，40:799–820。Yoshida, N.（2011年）。随机微分方程的多项式型大偏差不等式与准似然分析。《统计数学研究所年刊》，63(3):431–479。