数 学 年 刊 A 辑
2019, 40(1):93–104
DOI: 10.16205/j.cnki.cama.2019.0008

NSD 误差下线性模型 LAD 估计的线性表示*

王星惠[1] 韩瑞东[2] 胡舒合[3]

提要 考虑了 NSD 误差下的线性模型并建立回归参数 LAD 估计的线性表示. 这些结果将独立误差的情形推广和改进到 NSD 误差的情形. 作为一个应用, 获得了 LAD 估计量的收敛率.

关键词 LAD 估计量, 线性表示, NSD, 收敛率
MR (2000) 主题分类 62J05, 62F12
中图法分类 O212.1
文献标志码 A
文章编号 1000-8314(2019)01-0093-12

1 引 言

考虑线性回归模型

$$Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}_0 + e_i, \quad i = 1, 2, \cdots, n, \tag{1.1}$$

其中 $\boldsymbol{x}_i$ 是已知 $p \times 1$ 维设计向量, $\boldsymbol{\beta}_0 \in \mathbb{R}^p$ 是一个未知回归参数且 e_i 是随机误差.

$\boldsymbol{\beta}_0$ 的最小绝对偏差 (LAD) 估计量被定义为

$$\widehat{\boldsymbol{\beta}}_n = \arg\inf_{\boldsymbol{\beta} \in \mathbb{R}^n} \sum_{i=1}^n |Y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}|.$$

近些年来, 有关 LAD 估计的研究成为热点之一 (可见文 [1–8]).

全文记

$$\boldsymbol{S}_n = \sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i^\top, \quad d_n^2 = \max_{1 \leqslant i \leqslant n} \boldsymbol{x}_i^\top \boldsymbol{S}_n^{-1} \boldsymbol{x}_i,$$

假设某个整数 n_0, 对于 $n \geqslant n_0$ 有 $\boldsymbol{S}_{n_0} > 0$. 记 $\|\cdot\|$ 为 $\mathbb{R}^p$ 中的 L_2 范数. 设

$$\boldsymbol{x}_{ni} = \boldsymbol{S}_n^{-\frac{1}{2}} \boldsymbol{x}_i, \quad \boldsymbol{\beta}_{n0} = \boldsymbol{S}_n^{\frac{1}{2}} \boldsymbol{\beta}_0, \quad \widehat{\boldsymbol{\beta}}_n^* = \boldsymbol{S}_n^{\frac{1}{2}} \widehat{\boldsymbol{\beta}}_n,$$

则

$$\sum_{i=1}^n \boldsymbol{x}_{ni} \boldsymbol{x}_{ni}^\top = \boldsymbol{I}_p, \quad \sum_{i=1}^n ||\boldsymbol{x}_{ni}||^2 = p, \quad d_n^2 = \max_{1 \leqslant i \leqslant n} ||\boldsymbol{x}_{ni}||^2,$$

其中 $\boldsymbol{I}_p$ 是 $p \times p$ 维单位阵.

本文 2016 年 9 月 11 日收到, 2017 年 6 月 14 日收到修改稿.
[1]安徽大学统计学系, 合肥 230601. E-mail: wangxinghuial@163.com
[2]中国人民大学统计学院, 北京 100872. E-mail: 763162193@qq.com
[3]通信作者. 安徽大学数学科学学院, 合肥 230601. E-mail: hushuhe@263.net
*本文受到国家自然科学基金 (No. 11701005, No. 11871072, No. 11671012), 安徽省自然科学基金 (No. 1608085QA02, No. 1508085J06) 和安徽大学人才计划项目的资助.

模型 (1.1) 可以重新改写为

$$Y_i = \boldsymbol{x}_{ni}^{\top}\boldsymbol{\beta}_{n0} + e_i, \quad i = 1, 2, \cdots, n,$$

和

$$\sum_{i=1}^{n} |Y_i - \boldsymbol{x}_{ni}^{\top}\widehat{\boldsymbol{\beta}}_n^*| = \min_{\boldsymbol{\beta}\in\mathbb{R}^p} \sum_{i=1}^{n} |Y_i - \boldsymbol{x}_{ni}^{\top}\boldsymbol{\beta}|.$$

在假设 e_i 为独立的具有相同分布函数 F 的随机变量以及 F 的弱条件下, 对于 $p=1$, Bahadur[9] 证明了尽管 $\widehat{\boldsymbol{\beta}}_n$ 是 $\boldsymbol{\beta}_0$ 的一个非线性估计, 但是我们可以写

$$n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_0) = (2f(0))^{-1} n^{-\frac{1}{2}} \sum_{i=1}^{n} \mathrm{sign}(e_i) + \boldsymbol{r}_n,$$

其中 $f(0)$ 是 F 在 0 点处的一个导数, 且

$$\boldsymbol{r}_n = O(n^{-\frac{3}{4}}(\log n)^{\frac{1}{2}}(\log\log n)^{\frac{1}{4}}) \ \ \text{a.s.}.$$

模型 (1.1) 中的这个结果由 Babu[1] 推广到 $\boldsymbol{\beta}_0$ 的 LAD 估计上, 具体如下 (A1)–(A3).

首先, 我们罗列以下一些假设:

(A1) F 在 0 的某个领域有一个导数 f, $f(0) > 0$ 且 $F(0) = \frac{1}{2}$.

(A2) 当 $n \to \infty$ 时, $d_n(\log n)^{\frac{1}{2}} \to 0$.

(A3) 存在某个 $c > 0$, 使得对于 0 的某个领域内的所有 y, 有

$$|f(y) - f(0)| \leqslant c|y|^{\frac{1}{2}}.$$

定理 A 在模型 (1.1) 中, 假设 $\{e_i, i \geqslant 1\}$ 为一个具有相同分布函数 F 的独立随机变量序列. 如果假设 (A1)–(A3) 满足, 那么

$$2f(0)(\widehat{\boldsymbol{\beta}}_n^* - \boldsymbol{\beta}_{n0}) = \sum_{i=1}^{n} \boldsymbol{x}_{ni}\mathrm{sign}(e_i) + O(d_n^{\frac{1}{2}}(\log n)^{\frac{3}{4}}) \ \ \text{a.s.}. \tag{1.2}$$

Rao 和 Zhao[10] 和 Chen[11] 建立了线性模型 M-估计的线性表示 (详细可见文 [12]).

下面我们回顾由 Hu[13] 介绍的 NSD 随机变量的概念.

定义 1.1 (见 [14]) 一个函数 $\phi: \mathbb{R}^n \to \mathbb{R}$ 被称为超可加的, 如果对于所有的 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\phi(\boldsymbol{x}\vee\boldsymbol{y}) + \phi(\boldsymbol{x}\wedge\boldsymbol{y}) \geqslant \phi(\boldsymbol{x}) + \phi(\boldsymbol{y})$, 其中 $\vee$ 和 $\wedge$ 分别表示分量的最大值和最小值.

Kemperman[14] 指出, 如果 ϕ 有连续的二阶偏导数, 那么 ϕ 的超可加性等价于 $\frac{\partial^2\phi}{\partial x_i\partial x_j} \geqslant 0, 1 \leqslant i \neq j \leqslant n$. 超可加函数的例子有很多. 例如, 具有数值 $\sum\limits_{i=1}^{n} x_i$, $\max\limits_{1\leqslant k\leqslant n} \sum\limits_{i=1}^{k} x_i$ 和 $\min\limits_{1\leqslant i\leqslant n} x_i$ (见 [15]) 以及对所有 $t > 0$, $\exp\Big\{t \sum\limits_{1\leqslant i<j\leqslant n} x_i x_j\Big\}$ (见 [16]) 的函数 $\phi(x_1, x_2, \cdots, x_n)$ 均为超可加函数.

定义 1.2 一个随机向量 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ 被称为 NSD, 若

$$\mathrm{E}\phi(X_1, X_2, \cdots, X_n) \leqslant \mathrm{E}\phi(Y_1, Y_2, \cdots, Y_n), \tag{1.3}$$

其中 $Y_1, Y_2, \cdots, Y_n$ 为独立随机变量且对每个 i, $X_i \stackrel{d}{=} Y_i$, ϕ 是一个超可加函数, 且使得 (1.3) 中的期望存在. 随机变量序列 $\{X_n, n \geqslant 1\}$ 被称为 NSD 的, 若对每个 $n \geqslant 2$, $(X_1, X_2, \cdots, X_n)$ 是 NSD 的.

Hu[13] 给出了一个例子阐释 NSD 得不出由 Alam 和 Saxena[17] 介绍的 NA, 且设置了一个 NA 能否得出 NSD 的开放性问题. 此外, Hu 提供了有关 NSD 的一些基本性质和 3 个结构定理, 并且展示了一些著名的多元分布具有 NSD 性质. Christofides 和 Vaggelatou[15] 解决了上述关于 NA 可以得到 NSD 的问题. 作为一个结果, 我们可以说 NSD 是 NA 的一种推广. 有关 NSD 随机变量的研究可见文 [9, 16, 18–21] 等.

本文的主要目的是考虑同分布 NSD 误差下的模型 (1.1), 并建立回归参数 $\boldsymbol{\beta}_0$ 的 LAD 估计量的强线性表示. 这些结果将推广和改进 Babu 的相应结果. 作为一个应用, 获得了 LAD 估计的收敛率.

本文将进行如下安排. 第 2 节给出主要结果. 第 3 节给出 NSD 误差下 LAD 估计的一些模拟结果. 主要结果证明需要的一些预备引理和主要结果的证明分别在第 4 节和第 5 节.

通篇全文, 符号 c 表示在不同地方可以取不同值的正常数. 假设 $I(A)$ 是集合 A 的示性函数.

2 主要结果

本小节将给出以下主要结果.

定理 2.1 在模型 (1.1) 中, 假设 $\{e_i, i \geqslant 1\}$ 为同分布 NSD 随机变量序列. 如果 (A1)–(A2) 满足, 那么 (1.2) 成立.

注 2.1 与定理 A 相比较, 定理 2.1 中假设的 0 的某领域内的所有 y, 有

$$|f(y) - f(0)| \leqslant c|y|^{\frac{1}{2}}$$

这一条件去掉了. 所以定理 2.1 在较弱的条件下将 Babu[1] 的相应结果推广和改进到 NSD 误差的情形.

由定理 2.1 立即能得出有关模型 (1.1) 的下列推论.

推论 2.1 在定理 2.1 的条件下,

$$2f(0)\boldsymbol{S}_n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_0) = \boldsymbol{S}_n^{-\frac{1}{2}} \sum_{i=1}^{n} \boldsymbol{x}_i \mathrm{sign}(e_i) + O(d_n^{\frac{1}{2}}(\log n)^{\frac{3}{4}}) \quad \text{a.s..}$$

作为线性表示的一个应用, 我们获得 LAD 估计的收敛率.

定理 2.2 在定理 2.1 的条件下,

$$\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_0 = O(d_n(\log n)^{\frac{1}{2}}) \quad \text{a.s..} \tag{2.1}$$

3 模　拟

在本小节, 我们对线性模型的 LAD 估计的有限样本的优势通过实验进行了验证. 样本数据由下面模型产生:

$$Y_i = \beta_0 + \beta_1 \frac{i}{n} + e_i, \quad i = 1, 2, \cdots, n, \tag{3.1}$$

误差 $(e_1,e_2,\cdots,e_n)\sim N(\mathbf{0},\Sigma)$, 其中 $\mathbf{0}$ 是一个零向量且

$$\Sigma=\begin{pmatrix}1+\rho & -\rho & -\rho^2 & 0 & \cdots & 0 & 0 & 0 & 0\\ -\rho & 1+\rho & -\rho & -\rho^2 & \cdots & 0 & 0 & 0 & 0\\ -\rho^2 & -\rho & 1+\rho & -\rho & \cdots & 0 & 0 & 0 & 0\\ 0 & -\rho^2 & -\rho & 1+\rho & \cdots & 0 & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & 1+\rho & -\rho & -\rho^2 & 0\\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1+\rho & -\rho & -\rho^2\\ 0 & 0 & 0 & 0 & \cdots & -\rho^2 & -\rho & 1+\rho & -\rho\\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho^2 & -\rho & 1+\rho\end{pmatrix}_{n\times n},$$

$0<\rho<1$. 这个固定的设计向量满足我们的假设. 由文 [22] 可知, $(e_1,e_2,\cdots,e_n)$ 是一个 NA 向量, 进而也是一个 NSD 向量.

取 $\rho=0.2$ 和 $\rho=0.4$, 表 1 和表 2 罗列了模型 (3.1) 中 $\boldsymbol{\beta}_0=(\beta_0,\beta_1)^\top$ 的 LAD 估计 $\widehat{\boldsymbol{\beta}}_n$ 的均值、样本标准误 (SE) 和均方误差 (MSE). 取 $n=200$, 500, $(\beta_0,\beta_1)=(2,2)$ 和 $(2,4)$. 重复 1000 次. 从此两表中我们可以看出所有的偏差都较小. 随着样本量从 200 增加到 500, 估计的精度提高, 且 SE 和 MSE 的所有值都在变小. 模拟结果显示了线性模型在 NSD 误差下的 LAD 估计的有限样本的优势.

表 1 对于 $\rho=0.2$ 的模型 (3.1) 的 LAD 估计量的均值, SE 和 MSE

β_0	β_1		$n=200$		$n=500$	
			$\widehat{\beta_0}$	$\widehat{\beta_1}$	$\widehat{\beta_0}$	$\widehat{\beta_1}$
2	2	均值	2.00068	2.00162	1.99519	2.00084
		SE	0.08642	0.10794	0.06933	0.10194
		MSE	0.00746	0.01164	0.00482	0.01038
2	4	均值	1.99978	4.00435	1.99778	4.00431
		SE	0.08768	0.10875	0.07147	0.10197
		MSE	0.00768	0.01183	0.00511	0.01041

4 预备引理

本小节, 为了主要结果的证明, 我们给出一些引理.

引理 4.1 (见 [13]) 如果 $(X_1,X_2,\cdots,X_n)$ 是 NSD 的且 $f_1,f_2,\cdots,f_n$ 均为单调不减 (或单调不增) 函数, 那么 $(f_1(X_1),f_2(X_2),\cdots,f_n(X_n))$ 仍为 NSD 的.

表 2　对于 $\rho = 0.4$ 的模型 (3.1) 的 LAD 估计量的均值, SE 和 MSE

β_0	β_1		$n=200$ $\widehat{\beta_0}$	$n=200$ $\widehat{\beta_1}$	$n=500$ $\widehat{\beta_0}$	$n=500$ $\widehat{\beta_1}$
		均值	1.99512	2.00110	1.99629	2.00249
2	2	SE	0.08274	0.10749	0.06609	0.09867
		MSE	0.00686	0.01154	0.00438	0.00973
		均值	2.00293	3.99880	2.00261	3.99691
2	4	SE	0.08345	0.10751	0.06670	0.10009
		MSE	0.00697	0.01155	0.00445	0.01002

引理 4.2 (见 [23])　设 $\{X_n, n \geqslant 1\}$ 为 NSD 随机变量序列且满足 $EX_n = 0$, 对 $n \geqslant 1$ 和某个 $b > 0$, $|X_n| \leqslant b$ a.s., 记 $B_n^2 = \sum\limits_{i=1}^{n} \mathrm{E}X_i^2$, 则对任意的 $\varepsilon > 0$, 有

$$\mathrm{P}\Big(\Big|\sum_{i=1}^{n} X_i\Big| > \varepsilon\Big) \leqslant 2\exp\Big\{-\frac{\varepsilon^2}{2(2B_n^2 + b\varepsilon)}\Big\}.$$

引理 4.3 (见 [23])　设 $\{X_n, n \geqslant 1\}$ 为同分布 NSD 随机变量序列且满足 $\mathrm{E}X_n = 0$. $\{a_{nk}, k \geqslant 1, n \geqslant 1\}$ 为一个实数阵列且满足 $\max\limits_{1\leqslant k\leqslant n} |a_{nk}| = O(n^{-1})$, 则

$$\sum_{k=1}^{n} a_{nk}X_k \to 0 \quad \text{a.s.}, \quad n \to \infty.$$

取任意序列 $\mu_n \uparrow \infty$ 满足

$$\mu_n d_n^{\frac{1}{4}} (\log n)^{\frac{1}{8}} \to 0.$$

设

$$\nu_{n1} = \mu_n d_n^{\frac{1}{2}} (\log n)^{\frac{3}{4}}.$$

考虑 p-维矩形

$$D = \{\boldsymbol{\beta} \in \mathbb{R}^p : \|\boldsymbol{\beta}\| \leqslant \mu_n (\log n)^{\frac{1}{2}}\}.$$

将 D 剖分为一些子矩形 $D_1, D_2, \cdots, D_N$, 使得这些 D_k 的直径都小于 $\nu_{n0} = \mu_n d_n (\log n)^{\frac{1}{2}}$. 这个剖分可以选择 $N \leqslant cn^p$.

对于 $\boldsymbol{\beta} \in D$, 定义

$$R_i(\boldsymbol{\beta}) = 2\int_0^{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\beta}} \{(I(e_i \leqslant s) - I(e_i \leqslant 0)) - \mathrm{E}(I(e_i \leqslant s) - I(e_i \leqslant 0))\}\mathrm{d}s.$$

引理 4.4 设 (A1)–(A2) 满足且 D_l 为一个上述定义的子矩形, 则对任意的 $\varepsilon > 0$ 以及充分大 n, 有

$$\mathrm{P}\Big(\sup_{\boldsymbol{\alpha},\boldsymbol{\beta}\in D_l}\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta})-R_i(\boldsymbol{\alpha}))\Big|\geqslant \varepsilon v_{n1}^2\Big)\leqslant 8\exp(-c\mu_n\log n).$$

证 记 $\overline{D}_l$ 为 D_l 的闭包. 对每个 i, 在 $\overline{D}_l$ 中有 $\boldsymbol{\gamma}_{i1}$ 和 $\boldsymbol{\gamma}_{i2}$, 使得对所有的 $\boldsymbol{\beta}\in\overline{D}_l$, $\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i1}\leqslant\boldsymbol{x}_{ni}^{\top}\boldsymbol{\beta}\leqslant\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i2}$. 若 $(\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i1})(\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i2})<0$, 存在 $\boldsymbol{\gamma}_{i0}\in\overline{D}_l$, 使得 $\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i0}=0$, 则

$$\Big|\int_{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\alpha}}^{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\beta}}(I(e_i\leqslant s)-I(e_i\leqslant 0))\mathrm{d}s\Big|\leqslant\sum_{j=1}^{2}\int_{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i0}}^{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{ij}}(I(e_i\leqslant s)-I(e_i\leqslant 0))\mathrm{d}s,\tag{4.1}$$

其中 $\boldsymbol{\alpha}\in\overline{D}_l$ 和 $\boldsymbol{\beta}\in\overline{D}_l$. 若 $0\leqslant\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i1}\leqslant\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i2}$, $\boldsymbol{\gamma}_{i0}=\boldsymbol{\gamma}_{i1}$, 则有 (4.1). 若 $\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i1}\leqslant\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i2}\leqslant 0$, $\boldsymbol{\gamma}_{i0}=\boldsymbol{\gamma}_{i2}$, 则有 (4.1). 无论何种情况, 对于每个 i, 存在 $\boldsymbol{\gamma}_{i0}\in\overline{D}_l$, 使得

$$(\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i0})(\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{ij})\geqslant 0,\quad j=1,2,$$

且 (4.1) 成立. 注意到 $\|\boldsymbol{x}_{ni}\|\leqslant d_n$, 由 (A2), 有

$$|\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{ij}|\leqslant\|\boldsymbol{x}_{ni}\|\|\boldsymbol{\gamma}_{ij}\|\leqslant d_n\nu_{n0}=\mu_n d_n(\log n)^{\frac{1}{2}}\to 0,\quad j=0,1,2.$$

从而

$$\begin{aligned}&\sum_{i=1}^{n}\Big|\int_{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i0}}^{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{ij}}|\mathrm{E}(I(e_i\leqslant s)-I(e_i\leqslant 0))|\mathrm{d}s\Big|\\ &\leqslant c\sum_{i=1}^{n}\Big|\int_{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i0}}^{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{ij}}|s|\mathrm{d}s\Big|\\ &\leqslant c\sum_{i=1}^{n}\|\boldsymbol{x}_{ni}\|^2\|\boldsymbol{\gamma}_{ij}+\boldsymbol{\gamma}_{i0}\|\|\boldsymbol{\gamma}_{ij}-\boldsymbol{\gamma}_{i0}\|\\ &\leqslant cp\mu_n(\log n)^{\frac{1}{2}}\nu_{n0}=o(\nu_{n1}^2).\end{aligned}\tag{4.2}$$

由 (4.1) 和 (4.2), 对充分大的 n, 可得

$$\begin{aligned}&\mathrm{P}\Big(\sup_{\boldsymbol{\alpha},\boldsymbol{\beta}\in D_l}\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta})-R_i(\boldsymbol{\alpha}))\Big|\geqslant\varepsilon v_{n1}^2\Big)\\ &\leqslant\sum_{j=1}^{2}\mathrm{P}\Big(\sum_{i=1}^{n}(R_i(\boldsymbol{\gamma}_{ij})-R_i(\boldsymbol{\gamma}_{i0}))\geqslant\frac{\varepsilon v_{n1}^2}{3}\Big).\end{aligned}$$

注意对 $j=1,2$, 有

$$\begin{aligned}&R_i(\boldsymbol{\gamma}_{ij})-R_i(\boldsymbol{\gamma}_{i0})\\ &=2\int_{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i0}}^{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{ij}}\{(I(e_i\leqslant s)-I(e_i\leqslant 0))-\mathrm{E}(I(e_i\leqslant s)-I(e_i\leqslant 0))\}\mathrm{d}s\\ &=2\int_{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i0}}^{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{ij}}\{I(e_i\leqslant s)-\mathrm{E}I(e_i\leqslant s)\}\mathrm{d}s\\ &\quad-2\int_{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{i0}}^{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\gamma}_{ij}}\{I(e_i\leqslant 0)-\mathrm{E}I(e_i\leqslant 0)\}\mathrm{d}s\end{aligned}$$

$$=: \xi_i - \eta_i.$$

由引理 4.1 可知, $\{\xi_i, i \geqslant 1\}$ 和 $\{\eta_i, i \geqslant 1\}$ 为两个 NSD 随机变量序列. 很容易验证 $\mathrm{E}\xi_i = 0$,

$$|\xi_i| \leqslant c|\boldsymbol{x}_{ni}^\top(\boldsymbol{\gamma_{ij}} - \boldsymbol{\gamma_{i0}})| \leqslant cd_n\nu_{n0},$$

$$\sum_{i=1}^n \mathrm{E}\xi_i^2 \leqslant c\sum_{i=1}^n (\boldsymbol{x}_{ni}^\top(\boldsymbol{\gamma_{ij}} - \boldsymbol{\gamma_{i0}}))^2 \leqslant cp\mu_n(\log n)^{\frac{1}{2}}\nu_{n0}.$$

由引理 4.2, 对于充分大的 n, 可得

$$\begin{aligned}\mathrm{P}\Big(\sum_{i=1}^n \xi_i > \frac{\varepsilon\nu_{n1}^2}{6}\Big) &\leqslant \exp\Big(-\frac{\frac{\varepsilon^2\nu_{n1}^4}{36}}{2(c\mu_n(\log n)^{\frac{1}{2}}\nu_{n0} + \frac{\varepsilon\nu_{n1}^2}{6}\cdot cd_n\nu_{n0})}\Big)\\ &\leqslant \exp(-\mu_n\log n).\end{aligned} \tag{4.3}$$

类似于 (4.3) 的证明, 有

$$\mathrm{P}\Big(\sum_{i=1}^n \eta_i > \frac{\varepsilon\nu_{n1}^2}{6}\Big) \leqslant \exp(-\mu_n\log n).$$

引理 4.4 得证.

为简单起见, 在下面引理中将 $\sup\limits_{\substack{\boldsymbol{\alpha},\boldsymbol{\beta}\in D\\ \|\boldsymbol{\alpha}-\boldsymbol{\beta}\|\leqslant\nu_{n1}}}$ 简记为 $\sup\limits_{\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|\leqslant\nu_{n1}}$. 对于 $\boldsymbol{\beta}\in\mathbb{R}^p$, 定义

$$\widetilde{R}_i(\boldsymbol{\beta}) = |e_i - \boldsymbol{x}_{ni}^\top\boldsymbol{\beta}| - |e_i| + \boldsymbol{x}_{ni}^\top\boldsymbol{\beta}\mathrm{sign}(e_i) - f(0)(\boldsymbol{x}_{ni}^\top\boldsymbol{\beta})^2. \tag{4.4}$$

引理 4.5 若 (A1)–(A2) 满足, 则有

$$\mathrm{P}\Big(\sup_{\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|\leqslant\nu_{n1}}\Big|\sum_{i=1}^n (R_i(\boldsymbol{\beta}) - R_i(\boldsymbol{\alpha}))\Big| \geqslant \varepsilon\nu_{n1}^2\Big) = o(\nu_{n1}^2) \quad \text{a.s.}, \tag{4.5}$$

且

$$\sup_{\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|\leqslant\nu_{n1}}\Big|\sum_{i=1}^n (\widetilde{R}_i(\boldsymbol{\beta}) - \widetilde{R}_i(\boldsymbol{\alpha}))\Big| = o(\nu_{n1}^2) \quad \text{a.s.}. \tag{4.6}$$

证 注意到

$$\begin{aligned}R_i(\boldsymbol{\beta}) - R_i(\boldsymbol{\alpha}) &= 2\int_{\boldsymbol{x}_{ni}^\top\boldsymbol{\alpha}}^{\boldsymbol{x}_{ni}^\top\boldsymbol{\beta}} \{I(e_i \leqslant s) - \mathrm{E}I(e_i \leqslant s)\}\mathrm{d}s\\ &\quad - 2\int_{\boldsymbol{x}_{ni}^\top\boldsymbol{\alpha}}^{\boldsymbol{x}_{ni}^\top\boldsymbol{\beta}} \{I(e_i \leqslant 0) - \mathrm{E}I(e_i \leqslant 0)\}\mathrm{d}s\\ &=: \xi_i' - \eta_i'.\end{aligned} \tag{4.7}$$

对固定的 $\boldsymbol{\alpha}\in D$ 和 $\boldsymbol{\beta}\in D$ 且满足 $\|\boldsymbol{\alpha}-\boldsymbol{\beta}\| \leqslant 2\nu_{n1}$, 有

$$|\xi_i'| \leqslant cd_n\|\boldsymbol{\alpha}-\boldsymbol{\beta}\| \leqslant cd_n\nu_{n1},$$

和

$$\sum_{i=1}^n \mathrm{Var}(\xi_i')^2 \leqslant c\sum_{i=1}^n (\boldsymbol{x}_{ni}^\top(\boldsymbol{\beta}-\boldsymbol{\alpha}))^2 \leqslant cp\nu_{n1}^2.$$

利用引理 4.2, 对任意的 $\varepsilon>0$, 有

$$\mathrm{P}\Big(\Big|\sum_{i=1}^{n}\xi_i'\Big|\geqslant\varepsilon\nu_{n1}^2\Big)\leqslant 2\exp(-c\mu_n\log n). \tag{4.8}$$

类似于 (4.8) 的证明, 有

$$\mathrm{P}\Big(\Big|\sum_{i=1}^{n}\eta_i'\Big|\geqslant\varepsilon\nu_{n1}^2\Big)\leqslant 2\exp(-c\mu_n\log n). \tag{4.9}$$

由 (4.7)–(4.9), 可得

$$\mathrm{P}\Big(\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta})-R_i(\boldsymbol{\alpha}))\Big|\geqslant\varepsilon\nu_{n1}^2\Big)\leqslant 4\exp(-c\mu_n\log n).$$

取 $\boldsymbol{\beta}_j\in D_j$, $j=1,2,\cdots,N$. 由引理 4.4, 对任意的 $\varepsilon>0$ 和充分大的 n, 有

$$\begin{aligned}
&\mathrm{P}\Big(\sup_{\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|\leqslant\nu_{n1}}\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta})-R_i(\boldsymbol{\alpha}))\Big|\geqslant\varepsilon\nu_{n1}^2\Big)\\
&\leqslant\sum_{j=1}^{N}\mathrm{P}\Big(\sup_{\boldsymbol{\beta}\in D_j}\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta}_j)-R_i(\boldsymbol{\beta}))\Big|\geqslant\frac{\varepsilon\nu_{n1}^2}{3}\Big)\\
&\quad+\sum_{k=1}^{N}\mathrm{P}\Big(\sup_{\boldsymbol{\beta}\in D_k}\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta}_k)-R_i(\boldsymbol{\beta}))\Big|\geqslant\frac{\varepsilon\nu_{n1}^2}{3}\Big)\\
&\quad+\sum_{(j,k):\|\boldsymbol{\beta}_j-\boldsymbol{\beta}_k\|\leqslant 2\nu_{n1}}\mathrm{P}\Big(\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta}_j)-R_i(\boldsymbol{\beta}_k))\Big|\geqslant\frac{\varepsilon\nu_{n1}^2}{3}\Big)\\
&=2\sum_{j=1}^{N}\mathrm{P}\Big(\sup_{\boldsymbol{\beta}\in D_j}\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta}_j)-R_i(\boldsymbol{\beta}))\Big|\geqslant\frac{\varepsilon\nu_{n1}^2}{3}\Big)\\
&\quad+\sum_{(j,k):\|\boldsymbol{\beta}_j-\boldsymbol{\beta}_k\|\leqslant 2\nu_{n1}}\mathrm{P}\Big(\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta}_j)-R_i(\boldsymbol{\beta}_k))\Big|\geqslant\frac{\varepsilon\nu_{n1}^2}{3}\Big)\\
&\leqslant cn^p\exp(-c\mu_n\log n),
\end{aligned}$$

这可以推出

$$\sum_{n=1}^{\infty}\mathrm{P}\Big(\sup_{\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|\leqslant\nu_{n1}}\Big|\sum_{i=1}^{n}(R_i(\boldsymbol{\beta})-R_i(\boldsymbol{\alpha}))\Big|\geqslant\varepsilon\nu_{n1}^2\Big)<\infty.$$

根据 Borel-Cantelli 引理, (4.5) 证毕.

注意

$$\begin{aligned}
\widetilde{R}_i(\boldsymbol{\beta})&=2\int_0^{\boldsymbol{x}_{ni}^\top\boldsymbol{\beta}}(I(e_i\leqslant s)-I(e_i\leqslant 0))\mathrm{d}s-f(0)(\boldsymbol{x}_{ni}^\top\boldsymbol{\beta})^2\\
&=R_i(\boldsymbol{\beta})+2\int_0^{\boldsymbol{x}_{ni}^\top\boldsymbol{\beta}}(F(u)-F(0)-f(0)u)\mathrm{d}u.
\end{aligned} \tag{4.10}$$

对任意的 $\boldsymbol{\alpha}\in D$ 和 $\boldsymbol{\beta}\in D$ 满足 $\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|\leqslant\nu_{n1}$, 有

$$\sum_{i=1}^{n}\Big|\int_{\boldsymbol{x}_{ni}^\top\boldsymbol{\alpha}}^{\boldsymbol{x}_{ni}^\top\boldsymbol{\beta}}(F(u)-F(0)-f(0)u)\mathrm{d}u\Big|$$

$$
\begin{aligned}
&\leqslant c\sum_{i=1}^{n}\Big|\int_{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\alpha}}^{\boldsymbol{x}_{ni}^{\top}\boldsymbol{\beta}} u^2\mathrm{d}u\Big| \\
&\leqslant c\sum_{i=1}^{n}|(\boldsymbol{x}_{ni}^{\top}\boldsymbol{\alpha})^3-(\boldsymbol{x}_{ni}^{\top}\boldsymbol{\beta})^3| \\
&\leqslant cd_n(\log n)\nu_{n1}=o(\nu_{n1}^2).
\end{aligned}\tag{4.11}
$$

由 (4.5), (4.10)–(4.11), 立即可得 (4.6). 引理 4.5 得证.

5　主要结果的证明

本小节将给出主要结果的证明. 不失一般性, 我们可以假设 (1.1) 中的真实参数 $\boldsymbol{\beta}_0$ 为零向量, 进而 $\boldsymbol{\beta}_{n0}$ 也同样成立.

定理 2.1 的证明 记

$$
\boldsymbol{r}_n=2f(0)(\widehat{\boldsymbol{\beta}}_n^*-\boldsymbol{\beta}_{n0})-\sum_{i=1}^{n}\boldsymbol{x}_{ni}\mathrm{sign}(e_i).
$$

我们将分两步证明 (1.2). 首先证明

$$
\widetilde{\boldsymbol{\beta}}_n=\sum_{i=1}^{n}\boldsymbol{x}_{ni}\mathrm{sign}(e_i)=O(\log n)^{\frac12}\quad\text{a.s.}. \tag{5.1}
$$

记

$$
\boldsymbol{x}_{ni}=\begin{pmatrix}x_{ni1}\\ x_{ni2}\\ \vdots\\ x_{nip}\end{pmatrix},\quad a_{ni}=\begin{cases}\dfrac{x_{nik}}{(\log n)^{\frac12}c_n}, & 1\leqslant i\leqslant n,\\ 0, & i>n,\end{cases}\qquad k=1,2,\cdots,p,
$$

其中 $\{c_n\}$ 为任意的趋向于 ∞ 的数列, 则

$$
|a_{ni}|\leqslant cn^{-1}.
$$

由引理 4.3, 有

$$
\sum_{i=1}^{n}\frac{x_{nik}}{(\log n)^{\frac12}c_n}\mathrm{sign}(e_i)=\sum_{i=1}^{n}a_{ni}\mathrm{sign}(e_i)\to 0\quad\text{a.s.}.
$$

所以对 $k=1,2,\cdots,p$,

$$
\sum_{i=1}^{n}x_{nik}\mathrm{sign}(e_i)=O((\log n)^{\frac12})\quad\text{a.s.}.
$$

由此立即推出

$$
\sum_{i=1}^{n}\boldsymbol{x}_{ni}\mathrm{sign}(e_i)=O((\log n)^{\frac12})\ \text{a.s.},
$$

即 (5.1) 证毕.

第 2 步, 我们证明

$$\min_{\|\boldsymbol{\beta}-\widetilde{\boldsymbol{\beta}}_n\|=\nu_{n1}}\sum_{i=1}^{n}|e_i-\boldsymbol{x}_{ni}^{\top}\boldsymbol{\beta}|\geqslant\sum_{i=1}^{n}|e_i-\boldsymbol{x}_{ni}^{\top}\widetilde{\boldsymbol{\beta}}|+\frac{1}{2}f(0)\nu_{n1}^2$$
$$>\sum_{i=1}^{n}|e_i-\boldsymbol{x}_{ni}^{\top}\widetilde{\boldsymbol{\beta}}|,\tag{5.2}$$

由此结合 $|\cdot|$ 的凸性, 可得

$$\|\widehat{\boldsymbol{\beta}}_n^*-\widetilde{\boldsymbol{\beta}}_n\|<\nu_{n1},$$

则对任意的 $\mu_n\uparrow\infty$,

$$\boldsymbol{r}_n=O(\mu_n d_n^{\frac{1}{2}}(\log n)^{\frac{3}{4}})\ \ \text{a.s.},$$

立即可得 (1.2).

令 $A_n=\{\|\widetilde{\boldsymbol{\beta}}_n\|\leqslant b(\log n)^{\frac{1}{2}}\}$. 由 (A2), 对充分大的 n, 有 $\nu_{n1}\leqslant(\log n)^{\frac{1}{2}}$. 由引理 4.5 以及 (4.4) 和 (4.10), 有

$$\sup_{\|\boldsymbol{\beta}-\widetilde{\boldsymbol{\beta}}\|=\nu_{n1}}I(A_n)\Big|\sum_{i=1}^{n}(|e_i-\boldsymbol{x}_{ni}^{\top}\boldsymbol{\beta}|-|e_i-\boldsymbol{x}_{ni}^{\top}\widetilde{\boldsymbol{\beta}}_n|)-f(0)\|\boldsymbol{\beta}-\widetilde{\boldsymbol{\beta}}_n\|^2\Big|$$
$$=\sup_{\|\boldsymbol{\beta}-\widetilde{\boldsymbol{\beta}}\|=\nu_{n1}}I(A_n)\Big|\sum_{i=1}^{n}(\widetilde{R}_i(\boldsymbol{\beta})-\widetilde{R}_i(\widetilde{\boldsymbol{\beta}}_n))\Big|=o(\nu_{n1}^2)\ \ \text{a.s.}.$$

因此, 以概率 1 对充分大的 n, 事件 A_n 能推出 (5.2). (1.2) 证毕.

定理 2.2 的证明 注意到

$$2f(0)(\widehat{\boldsymbol{\beta}}_n-\boldsymbol{\beta}_0)=\boldsymbol{S}_n^{-\frac{1}{2}}\widetilde{\boldsymbol{\beta}}_n+\boldsymbol{S}_n^{-\frac{1}{2}}\boldsymbol{r}_n.\tag{5.3}$$

因为 $|\boldsymbol{S}_n^{-\frac{1}{2}}|=O(d_n)$, 其中 $|\boldsymbol{A}|$ 记为矩阵 $\boldsymbol{A}$ 中所有元素的绝对值的最大值, 所以

$$\boldsymbol{S}_n^{-\frac{1}{2}}\boldsymbol{r}_n=o(d_n)=o(d_n(\log n)^{\frac{1}{2}})\ \ \text{a.s.}.\tag{5.4}$$

由 (5.1), 可得 $\widetilde{\boldsymbol{\beta}}_n=O((\log n)^{\frac{1}{2}})$ a.s.. 从而,

$$\boldsymbol{S}_n^{-\frac{1}{2}}\widetilde{\boldsymbol{\beta}}_n=O(d_n(\log n)^{\frac{1}{2}})\ \ \text{a.s.}.\tag{5.5}$$

由 (5.3)–(5.5), 可得 (2.1).

致谢 作者感谢编委和匿名审稿专家对本文提出的重要的改进意见.

参 考 文 献

[1] Babu G J. Strong representations for LAD estimators in linear models [J]. *Probab Theory Related Fields*, 1989, 83:547–558.

[2] Bloomfield P, Steiger W L. Least absolute deviations, theory, applications, and algorithms [M]. Basel: Birkhäuser, 1983.

[3] Chen X R, Bai Z D, Zhao L C, et al. Asymptotic normality of minimum L_1 norm estimates in linear models[J]. *Sci in China Ser A*, 1990, 11:1311–1328.

[4] Herce M A. Asymptotic theory of LAD estimation in a unit root process with finite variance errors [J]. *Econometric Theory*, 1996, 12:129–153.

[5] Ling S. Self-weighted least absolute deviation estimation for infinite variance autoregressive models [J]. *J Roy Statist Soc Ser B*, 2005, 67:381–393.

[6] Pan J Z, Wang H, Yao Q W. Weighted least absolute deviations estimation for ARMA models with infinite variance [J]. *Econometric Theory*, 2007, 23:852–879.

[7] Rogers A J. Least absolute deviations regression under nonstandard conditions [J]. *Econometric Theory*, 2001, 17:820–852.

[8] Zhou Z Y, Lin Z Y. Asymptotic theory for LAD estimation of moderate deviations from a unit root [J]. *Stat Probab Lett*, 2014, 90:25–32.

[9] Bahadur R R. A note on quantiles in large samples [J]. *Ann Math Statist*, 1966, 37:577–580.

[10] Rao C R, Zhao L C. Linear representation of M-estimates in linear models [J]. *Canad J Statist*, 1992, 20:359–368.

[11] Chen X R. Linear representations for M-estimations of the parameter in linear models [J]. *Sci China Ser A*, 1983, 23:1264–1275.

[12] Chen X R, Zhao L C. M-methods in linear model [M]. Shanghai: Shanghai Scientific and Technical Publishers, 1996.

[13] Hu T Z. Negatively supperadditive dependence of random variables with applications [J]. *Chinese J Appl Probab Statist*, 2000, 16:133–144.

[14] Kemperman J H B. On the FKG-inequalities for measures on a partially ordered space [J]. *Proc Akad Wetenschappen Ser A*, 1977, 80:313–331.

[15] Christofides T C, Vaggelatou E. A connection between supermodular ordering and positive/negative association [J]. *J Multivariate Anal*, 2004, 88:138–151.

[16] Eghbal N, Amini M, Bozorgnia A. On the Kolmogorov inequalities for quadratic forms of dependent uniformly bounded random variables [J]. *Stat Probab Lett*, 2011, 81:1112–1120.

[17] Alam K, Saxena K M L. Positive dependence in multivariate distributions [J]. *Commun Stat Theory Methods*, 1981, 10:1183–1196.

[18] Eghbal N, Amini M, Bozorgnia A. Some maximal inequalities for quadratic forms of negative superadditive dependence random variables [J]. *Stat Probab Lett*, 2010, 80:587–591.

[19] Shen A T, Zhang Y, Volodin A. Applications of the Rosenthal-type inequality for negatively superadditive dependent random variables [J]. *Metrika*, 2015, 78:295–311.

[20] Shen Y, Wang X J, Yang W Z, et al. Almost sure convergence theorem and strong stability for weighted sums of NSD random variables [J]. *Acta Math Sin Engl Ser*, 2013, 29:743–756.

[21] Wang X J, Deng X, Zheng L L, et al. Complete convergence for arrays of rowwise negatively supperadditive-dependent random variables and its applications [J]. *Statistics*, 2014, 48:834–850.

[22] Joag-Dev K, Proschan F. Negative association of random variables with applications [J]. *Ann Stat*, 1983, 11:286–295.

[23] Wang X H, Hu S H. On the strong consistency of M-estimates in linear models for negatively superadditive dependent errors [J]. *Aust N Z J Stat*, 2015, 57:259–274.

Linear Representations of LAD Estimators in Linear Models with NSD Errors

WANG Xinghui[1] HAN Ruidong[2] HU Shuhe[3]

[1]Department of Statistics, Anhui University, Hefei 230601, China.
E-mail: wangxinghuial@163.com

[2]School of Statistics, Renmin University of China, Beijing 100872, China.
E-mail: 763162193@qq.com

[3]Corresponding author. School of Mathematical Science, Anhui University, Hefei 230601, China. E-mail: hushuhe@263.net

Abstract This paper deals with the linear model with NSD errors and establish the linear representation of least absolute deviation (LAD) estimators of the regressive parameter, which generalizes and improves the result of independent errors to the case of negatively superadditive dependent errors. As an application, the convergence rate of LAD estimators is obtained.

Keywords LAD estimator, Linear representation, NSD, Rate of convergence

2000 MR Subject Classification 62J05, 62F12

The English translation of this paper will be published in
Chinese Journal of Contemporary Mathematics, Vol. 40 No. 1, 2019
by ALLERTON PRESS, INC., USA