文章编号:1672 - 058X(2012)02 - 0039 - 04

NOD 样本下线性模型回归系数估计的强相合性

李树生,刘　锐
(安徽大学 数学科学学院,合肥 230039)

摘　要:线性模型是一类非常重要的数学模型,有着广泛的应运,其回归系数的估计是很多学者的研究对象,也是确定模型的关键,在文献中有不少研究. 此处研究了在 NOD 样本下,线性模型回归系数的估计,并证明了估计的相合性.

关键词:线性模型;回归分位数;NOD 序列
中图分类号:O212. 1　　**文献标志码**: A

1　主要结果

定义 1[1]　称随机变量 $X_1, X_2, \cdots, X_n$ 是 NUOD,如果对于所有的 $x_1, x_2, \cdots, x_n$,都有:

$$P(X_i > x_i, i = 1,2,\cdots,n) \leqslant \prod_{i=1}^{n} P(X_i > x_i) \tag{1}$$

称随机变量 $X_1, X_2, \cdots, X_n$ 是 NLOD,如果对于所有的 $x_1, x_2, \cdots, x_n$,都有:

$$P(X_i \leqslant x_i, i = 1,2,\cdots,n) \leqslant \prod_{i=1}^{n} P(X_i \leqslant x_i) \tag{2}$$

称随机变量 $X_1, X_2, \cdots, X_n$ 是 NOD 序列,如果它们既是 NUOD 又是 NLOD.

定义 2　设线性系统模型为:

$$y = x'\beta_0 + x'r_0 e \tag{3}$$

其中,x 为 p 维已知向量,β_0 为 p 维未知回归系数,e 为一维不可观察随机变量,则称 $\beta(u) = \beta_0 + r_0 Q_e(u)$ 为式(3)的回归分位数. 此处 $Q_e(u)$ 表示 e 的 u 分位数.

对于模型(3),$x_1, x_2, \cdots, x_n$ 是 R^p 中的设计点列,$y_1, y_1, \cdots, y_n$ 为对应的 y 的观察值,则得到的相应模型:

$$y_i = x'_i\beta_0 + x'r_0 e_i, i = 1,2,\cdots,n \tag{4}$$

于是可构造 $\beta(u)$ 的一个估计 $\beta_n(u)$,使其满足:

$$\sum_{i=1}^{n} \rho_u(y_i - x'_i\beta_n(u)) = \inf_{\beta \in R^p} \sum_{i=1}^{n} \rho_u(y_i - x'_i\beta) \tag{5}$$

其中,

$$\rho_u(t) = \begin{cases} ut, t \geqslant 0 \\ (u-1)t, t < 0 \end{cases} \tag{6}$$

收稿日期:2011 - 01 - 11;**修回日期**:2011 - 06 - 02.

作者简介:李树生(1983-),男,安徽肥东人,硕士研究生,从事统计推断研究.

对于 $0<u<1$,用式(4)(5)去估计回归的 u 分位数. 此法最初由 Koenker 和 Baseett 在文献[2]中提出,吴耀华在文献[3]中讨论了在独立样本条件下 $\beta_n(u)$ 的弱相合性,讨论了样本为强 φ 混合序列时 $\beta_n(u)$ 的强相合性,汪名杰等在文献[4]中证明了样本为 NA 序列时,回归系数一类估计的强相合性. 此处则证明了在样本为 NOD 序列时,有下面结果:

定理 1 设 $0<u<1, e_0, e_1, e_2, \cdots$ 为有共同分布 F 的 NOD 序列,且满足:

1) 存在 $\delta>0$, F 在 $[Q_e(u)-\delta, Q_e(u)+\delta]$ 中有导数 f, f 在 $Q_e(u)$ 处连续,且:

$$f(Q_e(u)) > 0 \tag{7}$$

2) 记 $S_n = \sum_1^n x_i x'$,存在自然数 m,当 $n > m$ 时, $S_n > 0$,且:

$$\lim_{n\to\infty} d_n^2 n^{1/2} \log n = 0 \tag{8}$$

其中, $d_n^2 = \max\limits_{1\leqslant x\leqslant n} \| S_n^{-1/2} x_i \|^2$,又设 $0 < a \leqslant x'_i r_0 \leqslant b < \infty$,则由式(4)所确定的 $\beta_n(u)$ 强收敛于 $\beta(u)$. 记:

$$x_{n_i} = S_n^{-1/2} x_i, y_{n_i} = y_i, e_{n_i} = e_i, i = 1,2,\cdots,n$$

$$\beta_{n0} = S_n^{1/2}\beta_0, \beta_{n0}(u) = S_n^{1/2}\beta_n(u), r_{n0} = S_n^{1/2} r_0$$

则模型(4)可写为:

$$y_{n_i} = x_{n'_i}\beta_{n0} + x_{n'_i} r_{n0} e_{ni}, i = 1,2,\cdots,n \tag{9}$$

其中, $\sum_{i=1}^n x_{n_i} x'_{n_i} = I_p$, $S_n^{1/2}$ 为 S_n 的正定对称平方根.

以 $\beta_{n0}(u)$ 记模型(9)中回归 u 分位数 β_{n0} 的估计,则对应的 d_n^2 可相应改写为:

$$d_n^2 = \max_{1\leqslant x\leqslant n} \| x_{n_i} \|^2 \tag{10}$$

其中, $\|.\|$ 表示欧式模.

不失一般性,以下设 $\beta_{n0}=0$.

2 定理的证明

为了证明定理 1,引入以下引理.

引理 1[5] 设 $X_1, X_2, \cdots, X_n$ 为 NOD 序列, $f_1, f_2, \cdots, f_n$ 是单调上升(或单调下降)的函数,那么随机变量 $f_1(X_1), f_2(X_2), \cdots, f_n(X_n)$ 是 NOD 序列.

引理 2[5] 设 $X_1, X_2, \cdots, X_n$ 为 NOD 序列,则:

$$E\left(\prod_{j=1}^n X_j\right) \leqslant \prod_{j=1}^n E(X_j) \tag{11}$$

引理 3 设 $X_1, X_2, \cdots, X_n$ 为 NOD 序列, $EX_i = 0$, $|X_i| \leqslant b$, a. s 成立. 记 $S_n = \sum_{i=1}^n X_i$,则对任意 $\varepsilon>0$,当 n 充分大时,有:

$$P\left(\frac{1}{n}|S_n| \geqslant \varepsilon\right) = c\exp\left\{-\frac{1}{b}n^{1/2}\varepsilon\right\} \tag{12}$$

其中 c 是与 η, ε, b 无关正的常数(下文中的 c 均为正的常数,不同的地方可能有不同的值).

证明 当 $0<t\leqslant\frac{1}{b}$ 时,有 $|tX_i|\leqslant 1$, a. s 成立. 接下来证明 $Ee^{tX_i}\leqslant e^{Et^2X_i^2}$.

$$\mathrm{e}^{tX_i} \leqslant 1 + tX_i + \frac{(tX_i)^2}{2!} + \frac{(tX_i)^3}{3!} + \cdots + \frac{(tX_i)^n}{n!} + o((tX_i)^n) \leqslant$$

$$1 + tX_i + \frac{(tX_i)^2}{2!} + \frac{(tX_i)^2}{3!} + \cdots + \frac{(tX_i)^2}{n!} \leqslant 1 + tX_i + (tX_i)^2$$

$$E\mathrm{e}^{tX_i} \leqslant 1 + EtX_i + E(tX_i)^2 = 1 + E(tX_i)^2 \leqslant \mathrm{e}^{E(tX_i)^2} = \mathrm{e}^{Et^2X_i^2}$$

于是对于 $\forall \varepsilon > 0$,由引理 2 可得:

$$P(S_n \geqslant n\varepsilon) \leqslant P(tS_n \geqslant tn\varepsilon) \leqslant P(\mathrm{e}^{tS_n} \geqslant \mathrm{e}^{tn\varepsilon}) \leqslant \mathrm{e}^{-tn\varepsilon} E\mathrm{e}^{tS_n} \leqslant$$

$$\mathrm{e}^{-tn\varepsilon} \prod_{i=1}^{n} E\mathrm{e}^{tX_i} \leqslant \mathrm{e}^{-tn\varepsilon} \prod_{i=1}^{n} \mathrm{e}^{Et^2X_i^2} \leqslant \mathrm{e}^{-tn\varepsilon} \mathrm{e}^{nt^2b^2} \tag{13}$$

令 $t = \dfrac{1}{\sqrt{n}b}$,代入式(13),得 $P(S_n \geqslant n\varepsilon) \leqslant c\mathrm{e}^{-\sqrt{n}\varepsilon/b}$.

同理可证 $P(-S_n \leqslant n\varepsilon) \leqslant c\mathrm{e}^{-\sqrt{n}\varepsilon/b}$,故式(12)成立.

引理 4　对模型(9),设有式(10)(11),且定理对 $e_1, e_2, \cdots$ 的条件仍然成立,则对任何满足:

$$\lim_{n\to\infty} d_n v_n = 0, \lim_{n\to\infty} v_n^2 / n^{1/2} \log n = \infty \tag{14}$$

的正的常数列 $\{v_n\}$,都有:

$$|\boldsymbol{\beta}_{n0}(u)| \leqslant v_n, \text{a. s.} \tag{15}$$

其中 $|.|$ 表示分量的极大模,即若 $X = (x^{(1)}, x^{(2)}, \cdots, x^{(p)})'$,则 $|X| = \max\limits_{1\leqslant i\leqslant n} |x^{(i)}|$.

证明类似于文献[4]中的引理 3,并且需要此处的引理 3.

证明　由引理 4 可得:

$$\|\boldsymbol{\beta}_n(u)\|^2 = \|S_n^{-1/2}\boldsymbol{\beta}_{n0}(u)\|^2 = \boldsymbol{\beta}'_{n0}(u)S_n^{-1}\boldsymbol{\beta}_{n0}(u) = \mathrm{tr}(\boldsymbol{\beta}_{n0}(u)S_n^{-1}\boldsymbol{\beta}'_{n0}(u)) =$$

$$\mathrm{tr}(S_n^{-1}\boldsymbol{\beta}_{n0}(u)\boldsymbol{\beta}'_{n0}(u)) \leqslant [\mathrm{tr}(S_n^{-2})\mathrm{tr}(\boldsymbol{\beta}_{n0}(u)\boldsymbol{\beta}'_{n0}(u))^2]^{1/2} =$$

$$\mathrm{tr}(S_n^{-2})^{1/2} \|\boldsymbol{\beta}_{n0}(u)\|^2 \leqslant \mathrm{tr}(S_n^{-2})^{1/2} p v_n^2, \text{a. s.} \tag{16}$$

固定 m,使 $S_m > 0$,对一切 $n \geqslant m$. 由 S_n 的正定性,记 S_n 的特征根为 $r_{n_1} \geqslant r_{n_2} \geqslant \cdots \geqslant r_{n_p} > 0$. 显然 $0 < r_{n_1}^{-2} \leqslant r_{n_2}^{-2} \leqslant \cdots \leqslant r_{n_p}^{-2}$ 为 S_n^{-2} 的特征根,由此可得:

$$\mathrm{tr}(S_n^{-2}) = \sum_{j=1}^{p} r_{n_j^{-2}} \leqslant p r_{n_p^{-2}} \tag{17}$$

代入式(16)即得:

$$\|\boldsymbol{\beta}_n(u)\|^2 \leqslant p^{1/2} r_{n_p}^{-1} p v_n^2 \leqslant p^{3/2} r_{n_p}^{-1} v_n^2 \tag{18}$$

由文献[7]知:

$$\mathrm{tr}(S_m S_n^{-1}) \geqslant \sum_{j=1}^{p} r_{m_j} r_{n_i}^{-1}, m \leqslant n \tag{19}$$

$$v_n^2 \mathrm{tr}(S_m S_n^{-1}) = v_n^2 \mathrm{tr} \sum_{j=1}^{m} x_j x'_j S_n^{-1} = v_n^2 \mathrm{tr} \sum_{j=1}^{m} x'_j S_n^{-1} x_j \leqslant$$

$$m v_n^2 \max_{1\leqslant j\leqslant n} x'_j S_n^{-1} x_j = m v_n^2 d_n^2 \to 0 \tag{20}$$

又 $r_{m_p} > 0$,否则 S_m 不可能大于 0.

利用式(19)(20),可得 $\lim\limits_{n\to\infty} v_n^2 r_{m_j} r_{n_j}^{-1} = 0, j = 1, 2, \cdots, p$,特别地,取 $j = p$,即得:

$$\lim_{n\to\infty} v_n^2 r_{n_j}^{-1} = 0 \tag{21}$$

由式(18)(21)得 $\lim\limits_{n\to\infty}\beta_n(u)=0$, a. s.

因假设 $\beta_{n0}=0$,显然有 $\beta(u)=0$,由此 $\beta_n(u)$ 为 $\beta(u)$ 的强相合估计. 定理得证.

参考文献:

[1] JOAG-DEV K, PROSCHAN F. Negative association of random variables with applications[J]. Ann Statist,1983(11):286-295

[2] KOEKER R, BASSETT G. Regress on Quantles[J]. Econometrica,1978(8):285-298

[3] 吴耀华. 回归分位数过程的收敛性[J]. 数学年刊(A 辑),1992(5):622-632

[4] NA 样本下线性模型回归系数的一类估计的强相合性[J]. 安徽大学学报,2003,27(4):12-17

[5] BOZORGNIA A, PATTERSON R F, TAYLOR R L. Limited theorems for dependent random variables[J]. World Congress Nonlinear Analysis',1996(92):1639-1650

[6] 苏淳,赵林城,王岳宝. NA 序列的矩不等式与弱收敛[J]. 中国科学(A 辑),1996,26(12):1091-1099

[7] 王松桂,贾忠贞. 矩阵论中不等式[M]. 合肥:安徽教育出版社,1994

Strong Consistency of the Estimation of Regression Coefficient in Linear Models Based on NOD Sequence Sample

LI Shu-sheng, LIU Rui

(School of Mathematical Science, Anhui University, Hefei 230039, China)

Abstract: The linear model is a very important mathematical model and is widely applied. The estimation of its regression coefficients is the research objects of many scholars and is also the key to model determined. There have existed some researches in the literature. In this paper, the estimation of regression coefficients in linear model based on NOD sequence sample is studied and its strong consistency is proved.

Key words: linear model; regression quantile; NOD sequence

责任编辑:李翠薇