第28卷第2期 重庆工商大学学报(自然科学版) 2011年4月
Vol. 28 NO. 2 J Chongqing Technol Business Univ. (Nat Sci Ed) Apr. 2010

文章编号:1672－058X(2011)02－0133－02

非线性半参数回归模型中参数的经验欧氏似然置信域*

方连娣

(铜陵学院 文学与艺术传媒系,安徽 铜陵 244000)

摘　要: 考虑非线性半参数回归模型,构造了模型中未知参数的经验欧氏似然比统计量,并证明了所提出的统计量具有渐近 χ^2 分布,由此,可构造未知参数的置信域.

关键词: 非线性半参数回归模型;经验欧氏似然;χ^2 分布

中图分类号: O212.7　　　　**文献标志码:** A

考虑非线性半参数回归模型:

$$y_i = f(x_i,\beta) + g(t_i) + \varepsilon_i;\ i = 1,\cdots,n \tag{1}$$

其中 $f(.,.)$ 为已知可测函数,$g(.)$ 是定义在 $[0,1]$ 上的未知函数,(x_i,t_i) 是在 $R^d \times R$ 上取值的可观测随机点列,ε_i 是 $i.i.d.$ 随机误差,其均值为零,方差 $\sigma^2 < \infty$,β 为待估计的 p 维参数向量. 在模型(1)中,当 $f(x,\beta) = x^\tau\beta$ 时,该模型即为部分线性回归模型;当 $g(t) = 0$ 时,该模型即为非线性回归模型;当 $f(x,\beta) = x^\tau\beta$ 且 $g(t) = 0$ 时,该模型即为线性回归模型. 因此,模型(1)是一类非常广泛的统计模型.

经验似然是由 Owen[1,2] 提出的一种非参数统计方法,在构造置信域方面有很多突出的优点,Qin 与 Lawless[3] 将该方法引入到半参数模型,并提出可用欧氏距离代替距离. 罗旭[4] 对半参数模型构造经验欧氏似然函数,并讨论了得到的参数估计的大样本性质. 由此,利用经验欧氏似然方法构造了模型(1)中未知参数的经验欧氏似然比统计量,在一定条件下,证明了所提出的统计量具有渐近 χ^2 分布,并利用所得结果,构造了参数的渐近置信域.

1 主要结果

首先,定义 $g(.)$ 的估计为 $\bar{g}_n(t) \hat{=} \bar{g}_n(t,\beta) \hat{=} \sum_{j=1}^n W_{nj}(t)[y_j - f(x_j,\beta)]$,其中权函数 $W_{nj}(t) = \dfrac{K\left(\dfrac{t-t_i}{h}\right)}{\sum_{i=1}^n K\left(\dfrac{t-t_i}{h}\right)}$,$K(.)$ 是核函数,窗宽 h 是一个收敛于 0 的常数列.

记 $\hat{y}_i = y_i - \sum_{j=1}^n W_{nj}(t_i)y_j$,$\hat{f}(x_i,\beta) = f(x_i,\beta) - \sum_{j=1}^n W_{nj}(t_i)f(x_j,\beta)$,$\hat{f}^{(1)}(x_i,\beta) \hat{=} \dfrac{\partial \hat{f}(x_i,\beta)}{\partial\beta}$,引入一个辅助随机向量:$Z_i(\beta) = \hat{f}^{(1)}(x_i,\beta)[y_i - f(x_i,\beta) - \hat{g}_n(t_i)] = \hat{f}^{(1)}(x_i,\beta)(\hat{y}_i - \hat{f}(x_i,\beta))$,$i = 1,\cdots,n$,当 β 是真参数时,$EZ_i(\beta) = 0$. 定义经验欧氏似然比统计量为:

$$l(\beta) = \sup\left\{-\frac{1}{2}\sum_{i=1}^n (np_i - 1)^2 \,\middle|\, p_i \geqslant 0, 1 \leqslant i \leqslant n, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i Z_i(\beta) = 0\right\} \tag{2}$$

利用 Lagrange 乘子法可求得满足式(2)的 $p_i = \dfrac{1}{n} + \dfrac{1}{n}(\bar{Z}(\beta))^\tau s^{-1}\bar{Z}(\beta)$,$i = 1,\cdots,n$,其中 $\bar{Z}(\beta) =$

收稿日期:2010－05－21;修回日期:2010－06－30.

* 基金项目:铜陵学院自然科学研究项目基金(2008tlxykj006).

作者简介:方连娣(1982－),女,安徽枞阳人,助教,从事非参数统计研究.

$\frac{1}{n}\sum_{i=1}^{n} Z_i(\beta)$. 故 $l(\beta) = -\frac{1}{2}\sum_{i=1}^{n}(np_i - 1)^2 = -\frac{n}{2}(\bar{Z}(\beta))^{\tau} S^{-1}\bar{Z}(\beta)$.

其中 $S = \frac{1}{n}\sum_{i=1}^{n}(Z_i(\beta) - \bar{Z}(\beta))(Z_i(\beta) - \bar{Z}(\beta))^{\tau} = \frac{1}{n}\sum_{i=1}^{n} Z_i(\beta)(Z_i(\beta))^{\tau} - \bar{Z}(\beta)(\bar{Z}(\beta))^{\tau}$.

记 $\frac{\partial f(x_i,\beta)}{\partial \beta_j} = h_j(t_i,\beta) + u_{ij}(\beta)$,其中 $h_j(t_i,\beta) = E\left(\frac{\partial f(x_i,\beta)}{\partial \beta_j}\middle| t_i\right)$,$u_i = (u_{i1},\cdots,u_{ip})^{\tau}$,为得到 $l(\beta)$ 的渐近分布,需要以下条件:

C1 t_1 在 $[0,1]$ 有连续密度 $r(t)$,且 $0 < \inf_{0\leqslant t\leqslant 1} r(t) \leqslant \sup_{0\leqslant t\leqslant 1} r(t) < \infty$.

C2 对每一个 x,$f(x,\beta)$ 关于 β 有连续的二阶导数.

C3 $E\varepsilon_1 < \infty$,$Eu_i^4 < \infty$.

C4 $nh^2\log n \to \infty$,$nh^4 \to 0$.

C5 存在正常数 $M_1, M_2 > 0$,及 $\rho > 0$,使得 $M_1 I(|u| \leqslant p) \leqslant K(u) \leqslant M_2 I(|u| \leqslant p)$,且 $K(\cdot)$ 在 $[-\rho,\rho]$ 上有界变差.

C6 $B \hat{=} \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} u_i u_i^{\tau}$ 存在且正定,且 $\lim_{n\to\infty}\frac{1}{a_n}\sum_{i=1}^{n} u_i = 0$,$u_i = o(n^{\frac{1}{4}})$,这里,$a_n = n^{\frac{1}{2}}\log n$.

C7 对于 $t \in [0,1]$,$g(t)$ 和 $h_j(t,\beta)$ 满足一阶 Lipschitz 条件,$1 \leqslant j \leqslant p$.

注:条件 C1,C7 是研究非参数所需的基本条件,C2,C6 是研究非线性回归模型的正则条件.

定理1 假设条件 C1-C7 成立,如果 β 为参数真值,则:

$$-2l(\beta) \xrightarrow{L} \chi_p^2, n \to \infty \tag{3}$$

基于定理1,可以定义参数 β 的置信域,即对任给的 $0 < \alpha < 1$,存在 c_{α} 使得 $P(\chi_p^2 > c_{\alpha}) = \alpha$,则 β 的具有渐近置信水平 $1-\alpha$ 的置信域为 $I_a(\beta) = \{\beta \in R^p \mid -2l(\beta) \leqslant c_{\alpha}\}$.

2 定理的证明

为证明定理,先给出下面两个引理:

引理1 在定理1的条件下,当 β 是参数真值时,有:

$$\max_{1\leqslant i,j\leqslant n} |W_{ni}(t_j)| = o((nh)^{-1}), \text{a. s.} \tag{4}$$

$$\max_{1\leqslant i\leqslant n}\left|\sum_{j=1}^{n} W_{nj}(t_i)\varepsilon_j\right| = o(n^{-\frac{1}{4}}), \text{a. s.} \tag{5}$$

$$\max_{1\leqslant i\leqslant n}\left|\sum_{j=1}^{n} W_{nj}(t_i) u_{jk}(\beta)\right| = o(n^{-\frac{1}{2}}), \text{a. s.} \tag{6}$$

对所有 $k = 1,2,\cdots,p$ 均成立.

证明 类似于文献[6],由条件 C1 和 C5 及密度核估计的一致相合性可证式(4)成立,在利用 Bernstein 不等式可证式(5)和式(6)成立.

引理2 在定理1的条件下,如果 β 是参数真值,有 $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i(\beta) \xrightarrow{L} N(0,\sigma^2 B)$,$\sum_{i=1}^{n}\frac{1}{n}Z_i(\beta)(Z_i(\beta))^{\tau} - \sigma^2 B = o_p(1)$,$\bar{Z}(\beta)(\bar{Z}(\beta))^{\tau} = o_p(1)$.

证明 利用定理条件和引理1,类似于文献[6]即证.

定理1的证明 由引理2知:

$$\begin{aligned} -2l(\beta) &= (\sqrt{n}\bar{Z}(\beta))^{\tau}((\sigma^2\beta)^{-1} + o_p(1))(\sqrt{n}\bar{Z}(\beta)) = \\ &(\sigma^2)^{-1}(\sqrt{n}\bar{Z}(\beta))^{\tau} B^{-1}(\sqrt{n}\bar{Z}(\beta)) + o_p(1) = \\ &\left(\frac{1}{\sigma}\sqrt{n}B^{-\frac{1}{2}}\bar{Z}(\beta)\right)^{\tau}\left(\frac{1}{\sigma}\sqrt{n}B^{-\frac{1}{2}}\bar{Z}(\beta)\right) + o_p(1) \end{aligned} \tag{7}$$

又由引理2得 $\frac{1}{\sigma}\sqrt{n}B^{-\frac{1}{2}}\bar{Z}(\beta) \xrightarrow{L} N(0,I_p)$,故:

$$\left(\frac{1}{\sigma}\sqrt{n}B^{-\frac{1}{2}}\bar{Z}(\beta)\right)^{\tau}\left(\frac{1}{\sigma}\sqrt{n}B^{-\frac{1}{2}}\bar{Z}(\beta)\right) \xrightarrow{L} \chi_p^2 \tag{8}$$

由式(7)和式(8)知定理1成立. (下转第144页)

[4] 熊金城,拓扑传递系统中的混沌 [J]. 中国科学 A辑·数学,2005.35(3):302-311
[5] 陈绥阳,褚蕾蕾. 动力系统基础及其方法 [M]. 北京:科学出版社,2002
[6] 关鹏. 关于几乎周期点的讨论 [J]. 河南科技大学学报,2007,28(5):76-78
[7] 范钦杰. 混沌与拓扑混合 [J]. 大学数学,2004,20(6):68-72
[8] 熊金城. 线段自映射的动力系统:非游荡集、拓扑熵及混乱 [J]. 数学进展,1988.17(1):1-11
[9] 周作领. 一维动力系统 [J]. 数学月刊,1988,3(3):43-65

Several Properties of Asymptotic Periodic Point

ZOU Cheng[1], ZHAO Qing-jun[2]

(1. Department of Basic Science, Sichuan College of Chemical Technology, Sichuan Luzhou 646005, China;
2. School of Mathematics and Statistics, Chongqing University of Arts and Sciences,
Chongqing 402160, China)

Abstract: f is a continuous self-mapping on unit segment I, $APer(f)$ is its asymptotic periodic point set, and some of its properties are discussed by comparison and examples.

Key words: unit interval; point set; asymptotic periodic point; property

责任编辑:李翠薇

(上接第134页)

参考文献:

[1] OWEN A B. Empirical likelihood ratio confidence intervials for a single function [J]. Riometrika, 1988, 75: 237-249
[2] OWNG A B. Empirical likelihood confidence regions [J]. Ann Statist, 1990, 18(1): 90-120
[3] QIN J, LAWLESS J F. Empirical likelihood and general estimating equations [J]. Ann Statist, 1994, 22: 300-325
[4] 罗旭. 半参数模型的经验欧氏似然估计的大样本性质 [J]. 应用概率统计,1994,10(4):344-352
[5] 冯三营. 非线性半参数回归模型中参数的经验似然置信域 [J]. 数学物理学报,2009,29A(5):1338-1349

Empirical Euclidean Likelihood Regions of the Parameters in Nonlinear Semiparametric Regression Models

FANG Lian-di

(Literature, Art and Media Department, Tongling University, Anhui Tongling 244000, China)

Abstract: In this paper, we consider the nonlinear semiparametric regression model and construct empirical Euclidean likelihood ratio statistics for the unkown parameters. It is shown that the proposed statistics has the asymptotic standard chi-square distribution, and hence it can be used to construct the confidence region of the unknown parameter.

Key words: nonlinear semiparametric regression model; empirical Euclidean likelihood; chi-square distribution

责任编辑:李翠薇