第25卷 第1期 湖 南 工 业 大 学 学 报 Vol.25 No.1
2011年1月 Journal of Hunan University of Technology Jan. 2011

$\widetilde{\rho}$混合误差下半变系数模型的相合估计

李　强，陈志彬

（湖南工业大学 理学院，湖南 株洲 412007）

摘　要：讨论了误差序列是$\widetilde{\rho}$混合序列的变系数模型。给出变系数模型系数函数的局部多项式估计，并证明了该模型系数函数估计是弱相合估计。

关键词：半变系数模型；$\widetilde{\rho}$混合；局部多项式估计；弱相合

中图分类号：O212.7　　**文献标志码**：A　　**文章编号**：1673-9833(2011)01-0050-03

Consistent Estimation of Semi-Varying-Coefficient Models under $\widetilde{\rho}$-Mixing Sequences

Li Qiang，Chen Zhibin

（School of Science, Hunan University of Technology, Zhuzhou Hunan 412007, China）

Abstract: Discusses the semi-varying-coefficient models whose error sequence is $\widetilde{\rho}$-mixing. Gives the local polyomial estimation of coefficient functions of semi-varying-coefficient models. Proves that the estimation of the coefficient functions of this model is weak consistency.

Keywords: semi-varying-coefficient models ; $\widetilde{\rho}$-mixing ; local polynomial method ; weak consistency

T. J. Hastie和R. J. Tibshirani[1]于1993年提出了变系数模型。J. Fan 等人[2]在检验函数系数是否真的变化时，提出了半变系数模型，其定义如下：

$$Y=\sum_{i=1}^{p}\alpha_i(U)X_i+\sum_{i=1}^{q}\beta_j Z_j+\varepsilon, \tag{1}$$

式中：Y为响应变量；$U\in\mathbf{R}$，随机误差ε独立于协变量$\left(U, X_1, X_2, \cdots, X_P, Z_1, Z_2,\cdots, Z_q\right)$，且$E(\varepsilon)=0, D(\varepsilon)=\sigma^2$；$\alpha_i(\cdot), i=1, 2, \cdots, p$为未知函数系数；$\beta_j$为常数。

半变系数模型是一类应用较广泛的模型，例如，当$p=1, X_i\equiv 1$时，式（1）就成为部分线性模型。文献[3-7]等都对此进行了讨论。吕士钦等人研究了误差独立时，半变系数模型PLS估计（profile least-squares estimation）的一致收敛速度[7]，本文在此基础上讨论$\widetilde{\rho}$混合误差条件下半变系数模型估计的相合性。

1　$\widetilde{\rho}$混合误差的定义

设$\{X_i\mid i=1, 2, \cdots, n, \cdots\}$是概率空间$(\Omega, \mathfrak{B}, P)$上的随机变量序列，$\mathfrak{F}_S=\sigma\left(X_i \mid i\in S\subset\mathbf{N}\right)$为$\sigma-$域，在$\mathfrak{B}$中给定$\sigma-$域$\mathfrak{F}, R$，令

$$\rho(\mathfrak{F}, R)=\sup\{|\mathrm{cor}(X,Y)|\,|X\in L_2(F), Y\in L_2(R)\},$$

其中，$\mathrm{cor}(X,Y)=\dfrac{E(XY)-E(X)E(Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$为相关系数。

Bradley[8]引入如下相依系数：对$k\geq 0$令

$\widetilde{\rho}(k)=\sup\{\rho(\mathfrak{F}_S, \mathfrak{F}_T)|$有限子集$S,T\subset\mathbf{N}$,且$\mathrm{dist}(S,T)\geq k\}$，

显然$0\leq\widetilde{\rho}(k+1)\leq\widetilde{\rho}(k)\leq 1$，且$\widetilde{\rho}(0)=1$。

定义 1　对随机序列$\{X_i|i=1, 2, \cdots, n, \cdots\}$，如果存在$k\in\mathbf{N}$使$\widetilde{\rho}(k)\leq 1$，则称$\{X_i\}$是$\widetilde{\rho}$混合序列。

收稿日期：2010-12-09

基金项目：湖南省教育厅科学研究基金资助项目（10C0656），湖南工业大学教学改革基金资助项目（09E66）

作者简介：李　强（1975-），男，湖南湘潭人，湖南工业大学讲师，博士，主要从事非参数估计和高维数据分析方面的研究，E-mail：lqdr2008@126.com

独立序列是$\tilde{\rho}$混合序列的特殊情形。$\tilde{\rho}$混合序列要求对某距离为k的任意有限子集之间的相关度不大，这是一种很弱的相依性要求，因此$\tilde{\rho}$混合序列是一类应用广泛且很有研究价值的随机变量序列。自文献[8]引入$\tilde{\rho}$混合概念以来，它引起了许多学者的兴趣。近年来，$\tilde{\rho}$混合序列极限理论的研究已取得不少成果[9-12]。

2　系数的 PLS 估计

假设

$$\left\{\left(U_k, X_{k1}, X_{k2}, \cdots, X_{kp}, Z_{k1}, Z_{k2}, \cdots, Z_{kp}, Y_k\right),\ k=1,2,\cdots,n\right\}$$

是模型（1）的随机样本，对任意给定的$\boldsymbol{\beta}$，模型（1）可记作：$Y_k^* = \sum_{i=1}^{p} \alpha_i(U_k) X_{ki} + \varepsilon_k, k=1,2,\cdots,n$。　　（2）

其中，$Y_k^* = Y_k - \sum_{i=1}^{q} \beta_j Z_{kj}$。若$\alpha_i(u)(i=1,2,\cdots,p)$有连续的二阶导数，则在$u_0$附近可记为

$\alpha_i(u) \approx \alpha_i(u_0) + \alpha_i'(u_0)(u-u_0) \equiv \alpha_i + b_i(u-u_0)$，$i=1,2,\cdots,p$。

根据局部线性方法的思想，求解

$$\min_{\substack{a_i, b_i \\ i=1,2,\cdots,p}} \sum_{k=1}^{n}\left[Y_k^* - \sum_{i=1}^{p} f_{ki} X_{ki}\right]^2 K_h(U_k - u_0), \tag{3}$$

其中，$f_{ki} = a_i + b_i(U_k - u_0)$；$K_h(\cdot)$是核函数，$h = h_n$是正数序列，称为窗宽，且当$n \to \infty$时，$h \to 0$，$nh \to \infty$。

令$\tilde{\boldsymbol{Y}} = (Y_1, Y_2, \cdots, Y_n)^{\mathrm{T}}$，$\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)^{\mathrm{T}}$，$\tilde{\boldsymbol{X}} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n)^{\mathrm{T}}$，

$\tilde{\boldsymbol{Z}} = (\boldsymbol{Z}_1, \boldsymbol{Z}_2, \cdots, \boldsymbol{Z}_n)^{\mathrm{T}}$，$\boldsymbol{Z}_i = (Z_{i1}, Z_{i2}, \cdots, Z_{iq})^{\mathrm{T}}$，

$\boldsymbol{W}_u = \operatorname{diag}\left(\boldsymbol{K}_h(U_1 - u), \boldsymbol{K}_h(U_2 - u), \cdots, \boldsymbol{K}_h(U_n - u)\right)$，

$\boldsymbol{\gamma}(u) = \left(a_1(u), a_2(u), \cdots, a_p(u), hb_1(u), hb_2(u), \cdots, hb_p(u)\right)^{\mathrm{T}}$，

且$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{\alpha}^{\mathrm{T}}(U_1)\boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{\alpha}^{\mathrm{T}}(U_n)\boldsymbol{X}_n \end{bmatrix}$，$\boldsymbol{D}_n = \begin{bmatrix} \boldsymbol{X}_1^{\mathrm{T}} & \dfrac{U_1 - u}{h}\boldsymbol{X}_1^{\mathrm{T}} \\ \vdots & \vdots \\ \boldsymbol{X}_n^{\mathrm{T}} & \dfrac{U_1 - u}{h}\boldsymbol{X}_n^{\mathrm{T}} \end{bmatrix}$，

则式（2）可记作：$\boldsymbol{Y} - \tilde{\boldsymbol{Z}}\boldsymbol{\beta} = \boldsymbol{M} + \boldsymbol{\varepsilon}$，　　（4）

可得式（3）的解为：$\hat{\boldsymbol{\gamma}}(u) = \left\{\boldsymbol{D}_u^{\mathrm{T}}\boldsymbol{W}_u\boldsymbol{D}_u\right\}^{-1}\boldsymbol{D}_u^{\mathrm{T}}\boldsymbol{W}_u\left(\boldsymbol{Y} - \tilde{\boldsymbol{Z}}\boldsymbol{\beta}\right)$。

$\boldsymbol{M}$的估计为：

$$\hat{\boldsymbol{M}} = \begin{bmatrix} \left[\boldsymbol{X}_1^{\mathrm{T}}\ \ 0\right]\left\{\boldsymbol{D}_{u_1}^{\mathrm{T}}\boldsymbol{W}_{u_1}\boldsymbol{D}_{u_1}\right\}^{-1}\boldsymbol{D}_{u_1}^{\mathrm{T}}\boldsymbol{W}_{u_1}\left(\boldsymbol{Y} - \tilde{\boldsymbol{Z}}\boldsymbol{\beta}\right) \\ \vdots \\ \left[\boldsymbol{X}_n^{\mathrm{T}}\ \ 0\right]\left\{\boldsymbol{D}_{u_n}^{\mathrm{T}}\boldsymbol{W}_{u_n}\boldsymbol{D}_{u_n}\right\}^{-1}\boldsymbol{D}_{u_n}^{\mathrm{T}}\boldsymbol{W}_{u_n}\left(\boldsymbol{Y} - \tilde{\boldsymbol{Z}}\boldsymbol{\beta}\right) \end{bmatrix}\left(\boldsymbol{Y} - \tilde{\boldsymbol{Z}}\boldsymbol{\beta}\right) \triangleq$$

$$\boldsymbol{S}(\boldsymbol{Y} - \tilde{\boldsymbol{Z}}\boldsymbol{\beta}), \tag{5}$$

将$\hat{\boldsymbol{M}}$代入式（4）得：$(\boldsymbol{I} - \boldsymbol{S})\boldsymbol{Y} = (\boldsymbol{I} - \boldsymbol{S})\tilde{\boldsymbol{Z}}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$，　　（6）

对线性模型（6）应用最小二乘法，可得：

$\hat{\boldsymbol{\beta}} = \left\{\tilde{\boldsymbol{Z}}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{S})^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{S})\tilde{\boldsymbol{Z}}^{\mathrm{T}}\right\}^{-1}\tilde{\boldsymbol{Z}}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{S})^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{S})\boldsymbol{Y}$，

于是$\hat{\boldsymbol{\gamma}}(u)$可修正为：$\hat{\boldsymbol{\gamma}}(u) = \left\{\boldsymbol{D}_u^{\mathrm{T}}\boldsymbol{W}_u\boldsymbol{D}_u\right\}^{-1}\boldsymbol{D}_u^{\mathrm{T}}\boldsymbol{W}_u\left(\boldsymbol{Y} - \tilde{\boldsymbol{Z}}\hat{\boldsymbol{\beta}}\right)$，

从而得：$\hat{\boldsymbol{\alpha}}_i(u) = \boldsymbol{e}_{i,2p}^{\mathrm{T}}\left\{\boldsymbol{D}_u^{\mathrm{T}}\boldsymbol{W}_u\boldsymbol{D}_u\right\}^{-1}\boldsymbol{D}_u\boldsymbol{W}_u\left(\boldsymbol{Y} - \tilde{\boldsymbol{Z}}\hat{\boldsymbol{\beta}}\right)$。

3　主要假设

首先给出证明需要的一些假设：

A1　函数系数$\alpha_i(\cdot)(i=1,2,\cdots,p)$有连续的二阶导数；

A2　随机变量U有有界支撑Ω，其密度函数$p(\cdot)$ Lipschitz 连续；

A3　核函数$K(\cdot)$是一有界对称的，且为有界紧支撑的概率密度函数；

A4　$E\left(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}} \middle| U=u\right)$对任意的$u \in \Omega$非奇异，$E\left(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}} \middle| U=u\right)$，$E\left(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}} \middle| U=u\right)^{-1}$和$E\left(\boldsymbol{X}\boldsymbol{Z}^{\mathrm{T}} \middle| U=u\right)$都是 Lipschitz 连续的；

A5　存在$s>2$，满足$E\|\boldsymbol{X}\|^{2s} < \infty$和$E\|\boldsymbol{Z}\|^{2s} < \infty$；

A6　当$n \to \infty$时，$h \to 0$，$nh \to \infty$，$h=n^{-l}$，$0<l<1/2$；

A7　设$\left\{\varepsilon_i \middle| i=1,2,\cdots,n\right\}$为$\tilde{\rho}$混合同分布序列，$E(\varepsilon_i)=0$，$E(\varepsilon_i^2)=\sigma^2$，

$\max_{1\le i\le n}\left|n^{-1}\boldsymbol{X}_i\boldsymbol{K}_h(U_i - u)\left(\dfrac{U_i - u}{h}\right)\right| \le \dfrac{c}{nh}$，$c$为常数。

4　结论及证明

引理1 [12]　设$\left\{X_i \middle| i=1,2,\cdots,n\right\}$为$\tilde{\rho}$混合随机变量序列，$E(X_i)=0$，$E\left(\left|X_i\right|^{\frac{2}{\alpha}}\right)<\infty$，$0<\alpha\le 1$，$|a_{ni}| \le cn^{-\alpha/2-\delta}$，$0<\delta\le 2/\alpha$，$\sum_{i=1}^{n} a_{ni}^2 \le cn^{-\theta}$，$\theta>0$，$c$为常数，则$\sum_{i=1}^{n} a_{ni}\boldsymbol{X}_i \xrightarrow{\text{a.s.}} 0$。

证明　参见文献[12]中定理2的证明。

引理2　在假设A1~A6成立的条件下，有

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \xrightarrow{p} 0。$$

证明　由文献[13]中定理4.1易得。

引理3　在假设A1~A7成立的条件下，有

$$\boldsymbol{C}_{n,\lambda} \xrightarrow{p} 0，$$

其中，$\boldsymbol{C}_{n,\lambda} = n^{-1}\sum_{i=1}^{n}\left(\boldsymbol{X}_i\boldsymbol{K}_h(U_i - u)(\varepsilon_i + \sum_{j=1}^{q} Z_{ij}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_j\right))\left(\dfrac{U_i - u}{h}\right)^{\lambda}\right)$，$\lambda=0,1$。

证明　令$\boldsymbol{C}_{n,\lambda} = \boldsymbol{C}_{n,\lambda}^{(1)} + \boldsymbol{C}_{n,\lambda}^{(2)}$，其中

$$\boldsymbol{C}_{n,\lambda}^{(1)} = n^{-1}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{K}_h(U_i - u)\left(\sum_{j=1}^{q} Z_{ij}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_j\right)\left(\frac{U_i - u}{h}\right)^{\lambda}\right),$$

$$\boldsymbol{C}_{n,\lambda}^{(2)} = n^{-1}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{K}_h(U_i - u)\left(\frac{U_i - u}{h}\right)^{\lambda}\varepsilon_i,$$

$\boldsymbol{C}_{n,\lambda}^{(1)} \xrightarrow{p} 0$的证明见文献[7]，下面只需证

$\boldsymbol{C}_{n,\lambda}^{(2)} \xrightarrow{p} 0$。令$v_{ni} = n^{-1}\boldsymbol{X}_i\boldsymbol{K}_h(U_i-u)\left(\dfrac{U_i-u}{h}\right)$，则 $\boldsymbol{C}_{n,1}^{(2)} = \sum_{i=1}^{n} v_{ni}\varepsilon_i$，在引理1中，令$\alpha=1$，$0<\delta=1/2-l<1/2<2/\alpha=2$，由假设A6和A7可知 $\boldsymbol{C}_{n,1}^{(2)} = \sum_{i=1}^{n} v_{ni}\varepsilon_i$满足引理1的条件，故有：

$\boldsymbol{C}_{n,1}^{(2)} = \sum_{i=1}^{n} v_{ni}\varepsilon_i \xrightarrow{\text{a.s.}} 0$。从而，$\boldsymbol{C}_{n,1}^{(2)} = \sum_{i=1}^{n} v_{ni}\varepsilon_i \xrightarrow{p} 0$。

同理可证，$\boldsymbol{C}_{n,0}^{(2)} = \sum_{i=1}^{n} v_{ni}\varepsilon_i \xrightarrow{p} 0$。

引理4 在假设A1 ~ A6成立的条件下，有

$$B_{n,\lambda} \xrightarrow{p} 0,\ A_{n,\lambda} - p(u)\Gamma(u)\mu_\lambda \xrightarrow{p} 0,$$

其中，

$$B_{n,\lambda} = n^{-1}\sum_{i=1}^{n}\left(\left(\frac{U_i-u}{h}\right)^{\lambda}\boldsymbol{X}_i\boldsymbol{K}_h(U_i-u)\cdot \sum_{j=1}^{p}\left(\alpha_j(U_i)-\alpha_j(u)-\alpha_j'(u)(U_i-u)\right)X_{ij}\right),$$

$$A_{n,\lambda} = n^{-1}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{X}_i^{\mathrm{T}}\left(\frac{U_i-u}{h}\right)^{\lambda}\boldsymbol{K}_h(U_i-u),$$

$\Gamma(u)=E\left(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\middle|U=u\right)$，$\mu_\lambda=\int u^\lambda K(u)\mathrm{d}u$。

证明 类似文献[13]中引理7.2的证明。

定理1 在假设A1 ~ A7成立的条件下，

$$\hat{\boldsymbol{\gamma}}(u)-\boldsymbol{\gamma}(u) \xrightarrow{p} 0,$$

从而$\hat{\alpha}_i(u)-\alpha_i(u) \xrightarrow{p} 0$。$i=1,2\cdots,p$。

证明 $\hat{\boldsymbol{\gamma}}(u)-\boldsymbol{\gamma}(u)=$

$$\left\{\boldsymbol{D}_u^{\mathrm{T}}\boldsymbol{W}_u\boldsymbol{D}_u\right\}^{-1}\boldsymbol{D}_u^{\mathrm{T}}\boldsymbol{W}_u\left(\boldsymbol{Y}-\tilde{\boldsymbol{Z}}\hat{\boldsymbol{\beta}}-\boldsymbol{D}_u\boldsymbol{\gamma}\right)=$$

$$\left\{\boldsymbol{D}_u^{\mathrm{T}}\boldsymbol{W}_u\boldsymbol{D}_u\right\}^{-1}\boldsymbol{D}_u^{\mathrm{T}}\boldsymbol{W}_u\left(\boldsymbol{M}-\boldsymbol{D}_u\boldsymbol{\gamma}\right)+\left(\hat{\boldsymbol{Z}}\boldsymbol{\beta}-\boldsymbol{Z}\hat{\boldsymbol{\beta}}+\boldsymbol{\varepsilon}\right)=$$

$$\begin{pmatrix} A_{n,0} & A_{n,1} \\ A_{n,1} & A_{n,2}\end{pmatrix}^{-1}\begin{pmatrix} B_{n,0}+C_{n,0} \\ B_{n,1}+C_{n,1}\end{pmatrix}。$$

由引理2 ~ 4可知，$A_{n,\lambda}$有界，且$C_{n,\lambda} \xrightarrow{p} 0$，$B_{n,\lambda}=O_p(h^2)\longrightarrow 0, \lambda=0,1$。定理得证。

参考文献：

[1] Hastie T J, Tibshirani R J. Varying-Coefficient Models[J]. J. Roy. Statist. Soc. Ser. B, 1993, 55(4)：757-796.

[2] Fan J，Zhang W Y. Simultaneous Confidence Bands Are and Hypothesis Testing in Varying Coeffcient Models[J]. Scandinavian Journal of Statistics, 2000, 27(4)：715-731.

[3] 李志强，薛留根. 缺失数据下的半参数变系数模型的借补估计[J]. 应用数学学报，2009，32(3)：422-430.
Li Zhiqiang，Xue Liugen. The Imputation Estimators of Semiparametric Varying-Coefficient Models with Missing Data [J]. Acta Mathematicae Applicatae Sinica，2009，32(3)：422-430.

[4] 吕士钦，张日权，卢准炜. 删失数据下半变系数模型估计的渐近正态性[J]. 应用概率统计，2009，25(6)：641-648.
Lv Shiqin，Zhang Riquan，Lu Zhunwei. The Asymptotic Normality of Estimation on Semivarying Coefficient Models with Censored Data[J]. Chinese Journal of Applied Probability and Statistics，2009，25(6)：641-648.

[5] Lv Shiqin，Zhang Riquan，Lu Zhunwei. Asymptotic Normality of Constrained Profile Least Squares Estimation on Semivarying Coeffcient Models[J]. Chinese Journal of Engineering Mathematics，2009，26 (2)：315-319.

[6] 富伯亭，郑刘锁. 半变系数模型的M估计[J]. 工程数学学报，2009，26(4)：609-617.
Fu Boting，Zheng Liusuo. M-Estimation on Semi-Varying Coefficient Models[J]. Chinese Journal of Engineering Mathematics，2009，26(4)：609-617.

[7] 吕士钦，张日权，蒙家福. 半变系数模型PLS估计的一致收敛速度[J].太原科技大学学报，2006，27(6)：432-434.
Lv Shiqin，Zhang Riquan，Meng Jiafu. The Uniform Convergence Rates of Profile Least-Squares Estimation on Semivarying Coefficient Models[J]. Journal of Taiyuan University of Science and Technology，2006，27(6)：432-434.

[8] Bradley R C. Eqnivalent Mixing Conditions for Random Fields [R]. Chapel Hill: Technical Report，1990.

[9] 吴群英. 混合序列的概率极限理论[M]. 北京：科学出版社，2006：206-247.
Wu Qunying. Limit Theory of Mixing Sequences[M]. Beijing：Science Press，2006：206-247.

[10] 蔡光辉，郭宝才. $\tilde{\rho}$混合序列的对数律和强大数律[J]. 科技通报，2006，22(3)：286-289.
Cai Guanghui，Guo Baocai. Law of Logarithm and Strong Law of Large Numbers for $\tilde{\rho}$ Mixing Sequences[J]. Bulletin of Science and Technology，2006，22(3)：286-289.

[11] Lu Zhaoyang, Xu Wei, Niu Yujun. Marcinkiewicz Strong Laws for Weighted Sums of $\tilde{\rho}$-Mixing Random Variable Sequences [J]. Chinese Journal of Engineering Mathematics，2009，26 (2)：320-330.

[12] 吴群英. 混合序列加权和的完全收敛性和强收敛[J]. 应用数学，2002，15(1)：1-4.
Wu Qunying. Convergence for Weighted Sums of Mixing Random Sequences[J]. Mathematica Applicata，2002，15 (1)：1-4.

[13] Fan J Q，Huang T. Profile Likelihood Inferences on Semi-parametric Varying-Coefficient Partially Linear Models[J]. Bernoulli，2005，11(6)：1031-1057.

（责任编辑：廖友媛）