应用概率统计 第二十七卷
第一期 2011年2月

Chinese Journal of Applied Probability
and Statistics Vol.27 No.1 Feb. 2011



# Local Linear Estimation of Conditional Third Central Moment *

Fang Yun Yu Zhou⋆

(*School of Finance and Statistics, East China Normal University, Shanghai, 200241*)

**Abstract**

In this paper, we propose a local linear estimator for conditional third central moment. The asymptotic bias and variance are derived. General cross validation (GCV) is recommended for bandwidth selection. A simple simulation study is carried out to illustrate the usefulness of the proposed method.

**Keywords:** Asymptotic properties, bandwidth selection, conditional third central moment, local linear estimator.

**AMS Subject Classification:** 62G05.

## §1. Introduction

Kernel smoothing is a commonly used method in nonparametric regression. Nadaraya-Watson (Nadaraya (1964) and Watson (1964)) estimator and the local polynomial kernel (Fan and Gijbels (1996)) estimator are the most popular methods. Beside estimating conditional mean, other fields, such as, the estimation of variance function and covariance matrix (Smith and Kohn (2002), Ledoit and Wolf (2004)), attracted more and more attention in the past decade. Many statisticians carefully studied the covariance and correlation function for longitudinal data (Yao, Müller and Wang (2005a, b); Fan, Huang and Li (2007)). Among others, Yin, Geng, Li and Wang (2010) developed a nonparametric model for conditional covariance matrix, extending existing models for conditional variance. Nevertheless, little literature investigated the higher order conditional moments. People put much attention on the covariance, correlation function and conditional covariance (matrix), but ignored higher order conditional moments, which are also important to uncover properties of conditional distribution. By taking the conditional third central moment into consideration, we in this paper try to fill this gap. A local linear kernel

*This work was supported by National Social Science Foundation (08CTJ001).

⋆Corresponding author, E-mail: yz19830224@gmail.com.

Received December 16, 2009. Revised June 22, 2010.

estimator is developed. The asymptotic bias and variance of the proposed estimator are investigated. Consequently, mean squared error (MSE) and mean of integrated squared error are described with the corresponding optimal bandwidth. A general cross-validation (GCV) score is provided to select the bandwidth. We illustrate the estimation procedure with a simulation study.

The rest of the paper is organized as follows. Section 2 describes the conditional third central moment model and the estimator. In Section 3, we study the asymptotic properties of the local linear estimator. We describe the bandwidth selection in Section 4. A simulation study is put in Section 5. All the proofs are relegated in the Appendix.

## §2. Model and Estimator

Let $X$ be a random variable and $U$ be an associated index random variable. Assume that, conditional on $U=u$, $X$ follows a distribution with mean $m(u)$, variance $\sigma^2(u)$, and the third central moment $\gamma_3(u)$. Assume that $\{X_i, U_i, i=1,\cdots,n\}$ are randomly sampled from the distribution of $(X,U)$. Our goal is to estimate $\gamma_3(u)$. For simplicity, we only consider the scalar case of $X$ in this paper. The extension to multivariate $X$ is straightforward.

Firstly, we need to estimate $m(u)$. Assume that $m(u)$ has up to $(p+1)$ continuous derivatives within some interval of interest. Then for each $U_i$, $m(U_i)$ can be approximated by a $p$-th degree polynomial,

$$m(U_i)\approx m(u)+m'(u)(U_i-u)+\cdots+\frac{m^{(p)}(u)}{p!}(U_i-u)^p. \tag{2.1}$$

Then $m(u)$ can be estimated by minimizing

$$\sum_{j=1}^{n}\Big(X_i-\sum_{l=0}^{p}\frac{m^{(l)}(u)}{l!}(U_i-u)^l\Big)^2K_h(U_i-u), \tag{2.2}$$

where $K(\cdot)$ is a symmetric kernel function and $h$ is the bandwidth. $K_h(\cdot)=K(\cdot/h)/h$. We define the notations

$$S_{n,j}=\sum_{i=1}^{n}K_h(U_i-u)(U_i-u)^j,\qquad V_{n,j}=\sum_{i=1}^{n}K_h(U_i-u)(U_i-u)^jX_i.$$

Denote by $S_n$ a $(p+1)\times(p+1)$ matrix with the $(k,l)$-th entry $S_{n,k+l-2}$ $(k,l=1,\cdots,p+1)$, and $V_n$ a $(p+1)$ dimensional vector with the $k$-th entry $V_{n,k-1}$ $(k=1,\cdots,p+1)$. Consequently, the estimator $\widehat{m}(u)$ can be explicitly expressed as

$$\widehat{m}(u)=\boldsymbol{\xi}_{p+1}^{\tau}S_n^{-1}V_n, \tag{2.3}$$

where $\boldsymbol{\xi}_{p+1}$ is $(p+1)\times 1$ vector with the first element 1 and the others 0.

Secondly, assume that $\gamma_3(u)$ has up to $(p+1)$ continuous derivatives. Similar to (2.2), we construct the following function for estimating $\gamma_3(u)$.

$$\sum_{j=1}^{n}\Big\{[X_i-\widehat{m}(u)]^3-\sum_{l=0}^{p}\frac{\gamma_3^{(l)}(u)}{l!}(U_i-u)^l\Big\}^2K_h(U_i-u). \tag{2.4}$$

Introduce the notation $\Gamma_{n,j}$ as

$$\Gamma_{n,j}=\sum_{i=1}^{n}K_h(U_i-u)(U_i-u)^j[X_i-\widehat{m}(U_i)]^3. \tag{2.5}$$

Thus the estimator $\widehat{\gamma}_3(u)$ can be expressed as

$$\widehat{\gamma}_3(u)=\boldsymbol{\xi}_{p+1}^{\tau}S_n^{-1}\Gamma_n, \tag{2.6}$$

where $\Gamma_n$ is a $p+1$ dimensional vector with the $k$-th element $\Gamma_{n,k-1}$ $(k=1,\cdots,p+1)$.

Since local linear estimator has several nice properties and was recommended by Fan and Gijbels (1996) in practice, we choose the local linear estimator $(p=1)$ to estimate $m(u)$ and $\gamma_3(u)$ in this paper. The asymptotical properties of local linear estimator $\widehat{\gamma}_3(u)$ are studied in next section. We want to point out that the local linear estimator can not only estimate the first and third central moments, but also can be extended to estimate higher-order conditional moments. For higher order conditional central moments, we can give the estimators by minimizing a similar function to (2.4). We only need to change the power of the $[X_i-\widehat{m}(u)]$ in (2.4) from 3 to a corresponding value.

## §3. Asymptotic Properties

The local linear estimator $\widehat{m}(u)$ has been extensively studied in the literature. In this section, we concentrate on the asymptotic properties of $\widehat{\gamma}_3(u)$ $(p=1)$. Let $h_1$ denote the bandwidth for $\widehat{m}(u)$ and $h_2$ for $\widehat{\gamma}_3(u)$. The moments of $K$ and $K^2$ are denoted respectively by $\mu_j=\displaystyle\int t^jK(t)\mathrm{d}t$ and $\nu_j=\displaystyle\int t^jK^2(t)\mathrm{d}t$. We list some regular conditions which are necessary in the proofs.

(C1) The index random variable $U$ has a compact support and a probability density $f(U)$, which is bounded away from 0 and has bounded and continuous second derivative.

(C2) There exists a constant $\delta\in[0,1)$ such that $\sup\limits_u \mathsf{E}[X^3]^{2+\delta}<\infty$.

(C3) The conditional mean $m(u)$ and the conditional third central moment $\gamma_3(u)$ has continuous third derivative.

(C4) $\mathsf{E}[X^k|U=u]$ has continuous second derivative in $u$ for $0 \le k \le 6$.

(C5) $n \to \infty$, $nh_1^{5/3} \to \infty$, $nh_1^5 \to 0$ and $h_2 = O_p(n^{-1/5})$.

(C6) $K(u)$ is a bounded probability density function symmetric about 0 so that $\mu_2 < \infty$ and $\nu_0 < \infty$.

Let $[X_i - m(U_i)]^3 = \gamma_3(U_i) + e_i$, where $e_i$ satisfies $\mathsf{E}(e_i|U_i) = 0$ and $\mathsf{Var}\,(e_i|U_i) = \omega(U_i)$. We demonstrate the asymptotic bias and variance of $\widehat{\gamma}_3(u)$ in Theorem 3.1.

**Theorem 3.1** Under the conditions (C1) to (C6), we have

$$\begin{aligned}\mathsf{E}(\widehat{\gamma}_3(u) - \gamma_3(u)|\mathcal{U}) &= \frac{h_2^2\mu_2}{2}\gamma_3^{(2)}(u) + o_p(h_2^2),\\ \mathsf{Var}\,(\widehat{\gamma}_3(u) - \gamma_3(u)|\mathcal{U}) &= \frac{\omega(u)\nu_0}{nh_2 f(u)} + o_p[(nh_2)^{-1}],\end{aligned} \tag{3.1}$$

where $\mathcal{U} = \{U_1, \cdots, U_n\}$ represents the observation of the index variable $U$.

**Remark 1** From (A.6) and (A.8) in the Appendix, we can find that two terms of $O_p[(nh_1)^{-1}]$ and $O_p(h_1^2)$ exists in the bias of $\widehat{\gamma}_3(u)$. To eliminate the influence of $\widehat{m}(u)$ on the bias of $\widehat{\gamma}_3(u)$ compared to $O_p(h_2^2)$, we assume that $nh_1^{5/3} \to \infty$ and $nh_1^5 \to 0$ in (C5). Obviously, we need undersmoothing in estimating $m(u)$, but the curve $\widehat{m}(u)$ can not be very rugged due to the restriction $nh_1^{5/3} \to \infty$. Yin et al. (2010) did not assume the undersmoothing for $\widehat{m}(u)$ because it estimated the conditional covariance matrix. The estimated moments themselves decided the different bandwidths between our paper and Yin et al. (2010). More explanation will be given in the Appendix.

By Theorem 3.1, it is easy to derive MSE of $\widehat{\gamma}_3(u)$.

$$\mathrm{MSE}(\widehat{\gamma}_3(u)) = \frac{h_2^4\mu_2^2}{4}[\gamma_3^{(2)}(u)]^2 + \frac{\omega(u)\nu_0}{nh_2 f(u)} + o_p[h_2^4 + (nh_2)^{-1}]. \tag{3.2}$$

Consequently, the MISE of $\widehat{\gamma}_3(u)$ (with weight function $w(u) \equiv 1$) is

$$\mathrm{MISE}(\widehat{\gamma}_3(u)) = \int \frac{h_2^4\mu_2^2}{4}[\gamma_3^{(2)}(u)]^2 \mathrm{d}u + \frac{1}{nh_2}\int \frac{\omega(u)\nu_0}{f(u)}\mathrm{d}u + o_p[h_2^4 + (nh_2)^{-1}]. \tag{3.3}$$

So the optimal bandwidth $h_2$ to minimize MISE is

$$h_{2,\mathrm{opt}} = n^{-1/5}\left[\frac{\nu_0 \displaystyle\int \omega(u)/f(u)\mathrm{d}u}{\mu_2^2 \displaystyle\int [\gamma_3^{(2)}(u)]^2 \mathrm{d}u}\right]^2.$$

The optimal bandwidth of $h_2$ is the order of $n^{-1/5}$. But unfortunately, we cannot directly use such an asymptotic bandwidth because of some unknowns. However, it shows that for such conditional third central moment problem, the order of the optimal bandwidth is exactly the same as that in common regression model. A general cross validation to selection bandwidth is given in next section.

## §4. Bandwidth Selection

For $\widehat{\gamma}_3(u)$, there are several popular methods for selecting the bandwidth, including the plug-in method (Ruppert, Sheather and Wand (1995)), the pre-asymptotic substitution method (Fan and Gijbels (1996)), the empirical bias bandwidth selection method (Ruppert (1997)), the cross-validation method, and the GCV method. Here we use the plug-in method to select an optimal bandwidth $b$ for $\widehat{m}(u)$ firstly. Since undersmooth for $\widehat{m}(u)$ is needed, we then use $h_1 = b \times n^{-1/5}$ for implementation. On the other hand, the estimator $\widehat{\gamma}_3(u)$ in (2.6) is a local linear kernel estimator with $[X_i - \widehat{m}(U_i)]^3$ as the response variable and $U_i$ as the predictor. So we define a GCV score as follows,

$$\mathrm{GCV}(h_2) = \frac{\sum\limits_{i=1}^{n}\{[X_i - \widehat{m}(U_i)]^3 - \widehat{\gamma}_3(U_i)\}^2}{[1 - \mathrm{tr}(A_n/n)]^2}, \tag{4.1}$$

where $A_n$ is the hat matrix

$$A_n = S_n^{-1}\begin{pmatrix} \sum\limits_{i=1}^{n} K_h(U_i - u) \\ \sum\limits_{i=1}^{n}(U_i - u)K_h(U_i - u) \end{pmatrix}.$$

The bandwidth $h_2$ can be obtained by minimizing $\mathrm{GCV}(h_2)$. We will use this method to select bandwidth in next section of simulation study.

## §5. Simulation Study

To illustrate our estimation method, we carried out a simple simulation study. The index variable $U$ was generated from uniform distribution $\mathscr{U}[0,1]$. The sample size of observation was set as $n = 200$, 400 and 800. There were two cases for $X$,

(1) $X$ was simulated from an exponential distribution $\mathrm{Exp}(1/\sin(U))$;

(2) $X$ was simulated from a normal distribution $N(\sin(U), 0.1 * \cos(U))$.

The bandwidths $h_1$ and $h_2$ were selected as described in Section 4. Epanechnikov kernel $K(z) = (3/4)(1 - z^2)I(|z| \leq 1)$ was employed through the simulation study. We chose the integrated bias, variance and MSE (IBS, IVAR, IMSE) to evaluate the performance of the estimators. 100 $\widetilde{U}_i$'s equally spaced from 0 to 1 were used to approximate integrated mean squared error of $\widehat{\gamma}_3(u)$. The IMSE was approximated by IMSE $\approx M^{-1}\sum\limits_{i=1}^{M}\sum\limits_{j=1}^{100}[\gamma_3(\widetilde{U}_j)-\widehat{\gamma}_3(\widetilde{U}_j)]^2\Delta_j$, where $\Delta_j$ was the grid space between $\widetilde{U}_j$ and $\widetilde{U}_{j+1}$. In our cases, all the $\Delta_j$ was 0.01. $M$ represented the total number of the simulation runs. We used $M=500$. The integrated bias and variance were approximated by similar approach. We report the results in Table 1.

Table 1 The integrated bias square, variance and MSE (IBS, IVAR, IMSE) of $\widehat{\gamma}_3(u)$

| | Case I | | | Case II | | |
|---|---|---|---|---|---|---|
| $n$ | IBS | IVAR | IMSE | IBS | IVAR | IMSE |
| 200 | 0.0295 | 0.2037 | 0.2328 | 1.01e-11 | 1.90e-08 | 1.90e-08 |
| 400 | 0.0189 | 0.0593 | 0.0781 | 6.02e-12 | 8.90e-09 | 8.90e-09 |
| 800 | 0.0089 | 0.0442 | 0.0530 | 9.83e-12 | 4.37e-09 | 4.37e-09 |

From Table 1, we can see that all the values are very small. Our estimator performs very well. The true value $\gamma_3(U_i)$ in case II keeps 0 so that all the values in case II are much smaller than those in case I. In both two cases, integrated bias, variance and MSE decrease as sample size $n$ arises, which verifies the consistency of the estimator $\widehat{\gamma}_3(u)$.

## Appendix

**Proof of Theorem 3.1** Note that for any random variable $R$ with first two moments, we have $R=\mathsf{E}(R)+O_p[\mathsf{Var}^{1/2}(R)]$. Thus for $r=0$, 1 and 2, we have

$$\sum_{i=1}^{n}K_h(U_i-u)(U_i-u)^r=nh^rf(u)\{1+O_p[(nh)^{-1/2}]\}.$$

Therefore, it is obvious that

$$\begin{aligned}&\widehat{m}(u)-m(u)\\=&\frac{1}{nh_1f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_1}\Big)[X_i-m(u)-m'(u)(U_i-u)][1+O_p((nh_1)^{-1/2})]\\=&\Big\{\frac{1}{nh_1f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_1}\Big)[X_i-m(U_i)]+\frac{h_1^2}{2}m^{(2)}(u)\mu_2\Big\}\{1+O_p[(nh_1)^{-1/2}]\}.\end{aligned}\tag{A.1}$$

Let $[X_i-m(U_i)]^3=\gamma_3(U_i)+e_i$, where $e_i$ satisfies $\mathsf{E}(e_i|U_i)=0$ and $\mathsf{Var}\,(e_i|U_i)=\omega(U_i)$.

We represent $\widehat{\gamma}_3(u)-\gamma_3(u)$ as

$$
\begin{aligned}
&\widehat{\gamma}_3(u)-\gamma_3(u)\\
=\;& \frac{1}{nh_2f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_2}\Big)\{[X_i-\widehat{m}(U_i)]^3-\gamma_3(u)-\gamma_3'(u)(U_i-u)\}\{1+O_p[(nh_2)^{-1/2}]\}\\
=\;& \frac{1}{nh_2f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_2}\Big)\{[X_i-m(U_i)]^3-\gamma_3(u)-(U_i-u)\gamma_3'(u)\}\\
&+\frac{1}{nh_2f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_2}\Big)e_i-3\frac{1}{nh_2f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_2}\Big)[X_i-m(U_i)]^2[\widehat{m}(U_i)-m(U_i)]\\
&+3\frac{1}{nh_2f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_2}\Big)[X_i-m(U_i)][\widehat{m}(U_i)-m(U_i)]^2\\
&-\frac{1}{nh_2f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_2}\Big)[\widehat{m}(U_i)-m(U_i)]^3+R_{n,1}\\
=:\;& I_1+I_2-3I_3+3I_4-I_5+R_{n,1}, \qquad\qquad (\mathrm{A}.2)
\end{aligned}
$$

where $R_{n,1}=O_p[(nh_2)^{-1/2}][I_1+I_2-3I_3+3I_4-I_5]$. Let $\mathcal{U}=\{U_1,\cdots,U_n\}$ represent the observation of the index variable $U$. It is easy to derive that

$$\mathsf{E}(I_1|\mathcal{U})=\frac{h_2^2}{2}\mu_2\gamma_3^{(2)}(u)+o_p(h_2^2). \qquad (\mathrm{A}.3)$$

Note that $\mathsf{E}(I_2|\mathcal{U})=0$ and by standard arguments, it follows that

$$
\begin{aligned}
\mathsf{Var}\,(I_2|\mathcal{U}) &= \frac{1}{nh_2^2f^2(u)}\mathsf{E}\Big[K^2\Big(\frac{U_i-u}{h_2}\Big)e_i^2\Big]\\
&= \frac{\omega(u)\nu_0}{nh_2f(u)}+o_p[(nh_2)^{-1}]. \qquad (\mathrm{A}.4)
\end{aligned}
$$

Rewrite $X_i=m(U_i)+\sigma(U_i)\epsilon_i$, where $\sigma(U_i)>0$, $\mathsf{E}(\epsilon_i|U_i)=0$ and $\mathsf{Var}\,(\epsilon_i|U_i)=1$. The term $I_3$ can be decomposed as $I_3=I_{31}+I_{32}+I_{33}$. Firstly, under condition (C5), we have

$$
\begin{aligned}
I_{31} &= \frac{1}{nh_2f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_2}\Big)\sigma^2(U_i)\epsilon_i^2\frac{1}{nh_1f(U_i)}\sum_{j=1}^{n}K\Big(\frac{U_i-U_j}{h_1}\Big)\sigma(U_j)\epsilon_j\\
&= \frac{1}{n^2h_1h_2f(u)}\sum_{i\neq j}\psi_{ij}+\frac{1}{n^2h_1h_2f(u)}\sum_{i=1}^{n}\psi_{ii}, \qquad (\mathrm{A}.5)
\end{aligned}
$$

with

$$\psi_{ij}=K\Big(\frac{U_i-u}{h_2}\Big)K\Big(\frac{U_j-U_i}{h_1}\Big)\sigma^2(U_i)\sigma(U_j)f^{-1}(U_i)\epsilon_i^2\epsilon_j.$$

It is obvious that $\mathsf{E}\psi_{ij}=0$. Assume that $\mathsf{E}(\epsilon_i^4|U_i)=\gamma_\epsilon^4$. We can derive that $\mathsf{E}(\psi_{ij}^2)=h_1h_2\nu_0^2\sigma^6(u)\gamma_\epsilon^4+o_p(h_1h_2)$. In consequence, $\mathsf{E}\big(\sum_{i\neq j}\psi_{ij}\big)^2\leq O(n(n-1)h_1h_2)$. Then for an arbitrary $\alpha>0$, by Markov's Inequality, we have

$$
\begin{aligned}
\mathsf{P}\Big\{\Big(\frac{1}{n^2h_1h_2}\Big|\sum_{i\neq j}\psi_{ij}\Big|\Big)/h_2^2>\alpha\Big\} &< \frac{1}{n^4h_1^2h_2^6\alpha^2}\mathsf{E}\Big(\sum_{i\neq j}\psi_{ij}\Big)^2\\
&\leq O[(n^2h_1h_2^5)^{-1}].
\end{aligned}
$$

Then by condition (C5),

$$\mathsf{P}\Big\{\Big(\frac{1}{n^2h_1h_2}\Big|\sum_{i\neq j}\psi_{ij}\Big|\Big)\Big/h_2^2>\alpha\Big\}\to 0.$$

It yields that

$$[n^2h_1h_2f(u)]^{-1}\sum_{i\neq j}\psi_{ij}=o_p(h_2^2).$$

In addition, we can obtain

$$[n^2h_1h_2f(u)]^{-1}\sum_{i=1}^{n}\psi_{ii}=O_p[(nh_1)^{-1}]. \tag{A.6}$$

Then by condition (C5), $I_{31}=o_p(h_2^2)$. On the other hand,

$$I_{32}=\frac{\mu_2h_1^2}{2nh_2f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_2}\Big)\sigma^2(U_i)\epsilon_i^2m^{(2)}(U_i). \tag{A.7}$$

By $\mathsf{E}(\epsilon_i^2|\mathcal{U})=1$, we can easily obtain that

$$\mathsf{E}(I_{32}|\mathcal{U})=\frac{\mu_2h_1^2}{2}\sigma^2(u)m^{(2)}(u)+o_p(h_1^2). \tag{A.8}$$

Then we can see the right side of (A.8) existed in the bias of $\widehat{\gamma}_3(u)$. Thus the estimation of $m(u)$ affects $\widehat{\gamma}_3(u)$. However, under (C5), $\mathsf{E}(I_{32}|\mathcal{U})=o_p(h_2^2)$. Then the influence of $\widehat{m}(u)$ on $\widehat{\gamma}_3(u)$ can be removed compared to $O_p(h_2^2)$ in (A.3). While if the conditional covariance matrix is estimated just as in Yin et al. (2010), then the term $I_{32}$ should contain $\epsilon_i$ but not $\epsilon_i^2$. As $\mathsf{E}(\epsilon_i|\mathcal{U})=0$, the term $I_{32}$ is naturally $o_p(h_2^2)$. So Yin et al. (2010) needed no undersmoothing for $\widehat{m}(u)$, but our paper needs. The undersmoothing for $\widehat{m}(u)$ required in our paper is determined by the estimated function $\gamma_3(u)$ itself.

Besides, $I_{33}=O_p[(nh_1)^{-1/2}](I_{31}+I_{32})$. Obviously, $I_{33}=o_p(h_2^2)$. Now we consider $I_4$.

$$\begin{aligned}I_4 &= \frac{1}{nh_2f(u)}\sum_{i=1}^{n}K\Big(\frac{U_i-u}{h_2}\Big)\sigma(U_i)\epsilon_i\Big\{\frac{1}{nh_1f(u)}\sum_{j=1}^{n}K\Big(\frac{U_j-U_i}{h_1}\Big)\sigma(U_j)\epsilon_j\\&\quad+\frac{\mu_2h_1^2}{2}m^{(2)}(U_i)+o_p[h_1^2+(nh_1)^{-1/2}]\Big\}^2\\&= O_p[(nh_2)^{-1/2}]\{O_p[(nh_1)^{-1/2}]+O_p(h_1^2)\}.\end{aligned} \tag{A.9}$$

Under condition (C5), $I_4=o_p(h_2^2)$. Similarly, $I_5=o_p(h_2^2)$. Consequently, $R_{n,1}=o_p(h_2^2)$.

In summary,

$$\begin{aligned}\mathsf{E}(\widehat{\gamma}_3(u)-\gamma_3(u)|\mathcal{U})&=\frac{h_2^2\mu_2}{2}\gamma_3^{(2)}(u)+o_p(h_2^2),\\\mathsf{Var}\,(\widehat{\gamma}_3(u)-\gamma_3(u)|\mathcal{U})&=\frac{\omega(u)\nu_0}{nh_2f(u)}+o_p[(nh_2)^{-1}].\end{aligned} \tag{A.10}$$

Then the proof for Theorem 3.1 has been completed. □

## References


[1] Fan, J. and Gijbels, I., *Local Polynomial Modeling and its Applications*, London: Chapman and Hall, 1996.

[2] Fan, J., Huang, T. and Li, R., Analysis of longitudinal data with semiparametric estimation of covariance function, *J. Amer. Statist. Assoc.*, **102**(2007), 632–641.

[3] Ledoit, O. and Wolf, M., Honey, I shrunk the sample covariance matrix, *J. Portfolio Management*, **4**(2004), 110–119.

[4] Nadaraya, E.A., On estimating regression, *Theory of Probability and its Applications*, **9**(1964), 141–142.

[5] Ruppert, D., Empirical-bias, bandwidths for local polynomial nonparametric regression and density estimation, *J. Amer. Statist. Assoc.*, **92**(1997), 1049–1062.

[6] Ruppert, D., Sheather, S.J. and Wand, M.P., An effective bandwidth selector for local least squares regression, *J. Amer. Statist. Assoc.*, **90**(1995), 1257–1270.

[7] Smith, M. and Kohn, R., Parsimonious covariance matrix estimation for longitudinal data, *J. Amer. Statist. Assoc.*, **97**(2002), 1141–1153.

[8] Yao, F., Müller, H.G. and Wang, J.-L., Functional data analysis for sparse longitudinal data, *J. Amer. Statist. Assoc.*, **100**(2005a), 577–590.

[9] Yao, F., Müller, H.G. and Wang, J.-L., Functional regression analysis for longitudinal data, *Ann. Statist.*, **33**(2005b), 2873–2903.

[10] Yin, J., Geng, Z., Li, R. and Wang, H., Nonparametric covariance model, Statistica Sinica, **20**(2010), 469–479.

[11] Watson, G.S., Smooth regression analysis, *Sankhya, Series A*, **26**(1964), 359–372.


# 条件三阶中心矩的局部线性估计


房 云 於 州

(华东师范大学金融与统计学院, 上海, 200241)



本文给出了条件三阶中心矩的局部线性估计, 并研究了估计的条件偏差和方差. 本文利用广义交叉核实法(GCV)进行窗宽选择. 我们通过模拟说明了该估计的实用性.



**关键词:** 渐近性质, 窗宽选择, 条件三阶中心矩, 局部线性估计.

**学科分类号:** O212.