文章编号: 1000_0887(2006) 07_0789_10


# 强相关数据回归函数的小波估计*

李林元[1], 肖益民[2]

(1. 美国新罕布什尔大学 数理统计系, 美国;
2. 美国密执安州立大学 统计和概率系, 美国)

(周哲玮推荐)

摘要: 讨论了在强相关数据情形下对回归函数的小波估计, 并且给出了估计量的均方误差的一个渐近展开表示式• 对研究估计量的优劣, 所推导的近似表示式显得非常重要• 对一般的回归函数核估计, 如果回归函数不是充分光滑, 这个均方误差表示式并不成立• 但对小波估计, 即使回归函数间断连续, 这个均方误差表示式仍然成立• 因此, 小波估计的收敛速度要比核估计来得快, 从而小波估计在某种程度上改进了现有的核估计•

关 键 词: 均方误差; 非参数回归; 小波估计; 收敛率
中图分类号: O212.7 文献标识码: A

## 引 言

设 $Y_1, Y_2, \cdots, Y_n$ 是在固定点 $x_1, x_2, \cdots, x_n$ 的 $n$ 个观察值, 适合模型

$$Y_k = g(x_k) + \varepsilon_k, \qquad k = 1, 2, \cdots, n, \tag{1}$$

其中 $g(x)$ 是 $[0, 1]$ 上要估计的未知函数, $x_k = (k/n) \in [0, 1]$, $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是随机误差序列• 在通常情形下, 假定 $\{\varepsilon_k\}$ 是独立、同分布或者弱相相关 ARMA 随机过程[1~3]• 但是在农业、经济、环境科学、地理等学科中, 这些随机误差序列通常并不相互独立, 而是强相关的高斯平稳过程• 确切地说, 这些 $\{\varepsilon_k\}$ 满足如下条件

$$r(j) = E(\varepsilon_1 \varepsilon_{1+j}) \sim C_0 j^{-\alpha}, \tag{2}$$

其中, $\alpha \in (0, 1)$, $C_0$ 是一个常数, $a_j \sim b_j$ 表示当 $j \to \infty$, $a_j / b_j \to 1$•

至今, 对强相关数据已经有很多研究, 这里只列出其中一小部分• 文献[4]系统地介绍了处理这些强相关数据的方法• 通常强相关数据估计量的性质不同于独立数据或者弱相关数据估计量的性质• 文献[5]给出了在强相关数据下核估计的最优收敛率, 这个收敛率不同于弱相关数据的收敛速度• 文献[6]讨论了强相关数据下线性回归模型的参数估计• 文献[7]和文献[8]讨论了强相关数据下回归函数的核估计, 他们证明了估计量的渐近正态分布和均方误差的渐近展开式• 但是以上这些研究都假定回归函数是非常光滑的函数•

非参数回归函数的小波估计的研究开始于 90 年代, 至今已有很多成果• 文献[9]系统地介绍了小波及它在非参数统计中的应用• 由于小波的优良特性, 小波估计其有误差小, 收敛速度快等优点• 文献[10] ~ 文献[15]研究了概率密度函数的小波估计和非参数回归函数的小波

* 收稿日期: 2005_01_17; 修订日期: 2006_04_09
作者简介: 李林元(1965—), 男, 江苏人, 副教授, 博士(联系人. E_mail: linyuan@math.unh.edu)•

估计• 他们证明了估计量达到最优收敛率, 但是这些研究都假定随机误差是相互独立的•

对强相关函数, 文献[16] ~ 文献[18] 研究了非参数回归函数的小波估计• 在回归函数属于一个光滑函数空间的假定下, 他们证明了估计量达到最优收敛率• 本文考虑回归函数是一个固定的, 但是间断光滑函数, 提出了它的小波估计量, 并且得到了它的均方误差的近似展开表示式• 这个结果推广了文献[13] 的结果, 即从原来的独立观察值推广到强相关观察值•

# 1 记号和估计量

这节给出一些有关小波的记号和结果• $\phi(x)$ 和 $\psi(x)$ 表示通常的父亲和母亲小波, 满足如下基本条件 $\int\phi^2\mathrm{d}y = \int\psi^2\mathrm{d}y = 1$• 对 $0 \leqslant k \leqslant r-1$, $v_k = \int y^k\psi(y)\mathrm{d}y = 0$, 但是 $v_r = \int y^r\psi(y)\mathrm{d}y \neq 0$• 对任意 $p > 0$, $-\infty < j < \infty$, $p_i = p2^i$, $i \geqslant 0$, 记

$$\phi_j(x) = p^{1/2}\phi(px - j),\quad \psi_{ij}(x) = p_i^{1/2}\psi(p_ix - j),\qquad x \in \mathrm{R},$$

那么 $\{\psi_{ij}(x), i, j \in \mathrm{Z}\}$ 是空间 $L^2(\mathrm{R})$ 的一个正交基, 并且满足

$$\int\phi_{j_1}\phi_{j_2} = \delta_{j_1j_2},\quad \int\psi_{i_1j_1}\psi_{i_2j_2} = \delta_{i_1i_2}\delta_{j_1j_2},\quad \int\phi_{j_1}\psi_{ij_2} = 0,$$

其中 $\delta_{ij}$ 表示 Kronecker 函数, 即如果 $i = j$, $\delta_{ij} = 1$, 如果 $i \neq j$, $\delta_{ij} = 0$. 文献[19]和文献[9]提供了有关小波的更详细介绍•

对于本文的回归模型, 假定回归函数 $g$ 定义在固定的区间[0, 1] 上, 在[0, 1] 外的值定义为0• 根据文献[20], 我们可以找到定义在[0, 1] 上的小波 $\psi(x)$, 满足如下条件

$$|\psi(x) - \psi(y)| \leqslant C|x - y|^{1/2},\qquad \text{对所有 } x, y \in [0, 1], \tag{3}$$

对任意回归函数 $g \in L^2[0, 1]$, 其有如下小波展开

$$g(x) = \sum_{j=0}^{p-1}b_j\phi_j(x) + \sum_{i=0}^{\infty}\sum_{j=0}^{p_i-1}b_{ij}\psi_{ij}(x),\qquad x \in [0, 1], \tag{4}$$

其中系数满足

$$b_j = \int g\phi_j,\quad b_{ij} = \int g\psi_{ij}•$$

回归系数 $g(x)$ 的非线性小波估计定义为

$$\hat{g}(x) = \sum_{j=0}^{p-1}\hat{b}_j\phi_j(x) + \sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}\hat{b}_{ij}I(|\hat{b}_{ij}| > \delta_i)\psi_{ij}(x), \tag{5}$$

其中 $\delta_i$、$p$、$q$ 是参数, 系数 $\hat{b}_j$、$\hat{b}_{ij}$ 定义为

$$\hat{b}_j = \frac{1}{n}\sum_{k=1}^{n}Y_k\phi_j(x_k),\quad \hat{b}_{ij} = \frac{1}{n}\sum_{k=1}^{n}Y_k\psi_{ij}(x_k)• \tag{6}$$

本文 $C$ 表示一般的有限正数, 它其体的值在不同的地方可能不一样, 至于 $C_0$、$C_1$、$C_2$ 等表示固定的常数•

# 2 主 要 结 论

在估计量 $\hat{g}$ 中, $p$、$q$ 是 $\delta_i$ 样本容量 $n$ 的函数, 假定满足如下条件

$$(\mathrm{SP}):\quad \begin{cases} p \to \infty,\ q \to \infty,\ p_i\delta_i^2 \to 0,\ p_i^{2r+1}\delta_i^2 \to \infty, \\ \delta_i \geqslant \sqrt{4C_1\ln n/(n^{\alpha}p_i^{1-\alpha})}, \end{cases}\qquad i = 1, 2, \cdots, q-1,$$

其中 $$C_1 = C_0\iint|x - y|^{-\alpha}\psi(x)\psi(y)\mathrm{d}x\mathrm{d}y•$$

**定理 2.1** 设基本条件(SP) 成立, 又设回归函数 $g$ 的 $r$ 阶导数, $g^{(r)}$, 连续, 则

$$E\left|\int(\hat{g}-g)^2-\left\{C_2(n^{-1}p)^\alpha+p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int g^{(r)^2}\right\}\right|=$$
$$o((n^{-1}p)^\alpha+p^{-2r}), \tag{7}$$

其中 $\kappa=(r!)^{-1}\nu_r,\ C_2=C_0\iint|x-y|^{-\alpha}\phi(x)\phi(y)\mathrm{d}x\mathrm{d}y$•

注 2.1 根据 Fourier 变换和 Plancherel 定理

$$\iint|x-y|^{-\alpha}\psi(x)\psi(y)\mathrm{d}x\mathrm{d}y=C_\alpha\int_{\mathbb{R}}\frac{1}{|u|^{1-\alpha}}|\hat{\psi}(u)|^2\mathrm{d}u>0,$$

其中 $C_\alpha>0$ 是一个常数, $\hat{\psi}$ 是 $\psi$ 的 Fourier 变换, 因此常数 $C_1$ 为一个有限的正数• 同样, $C_2$ 也是一个有限的正数•

注 2.2 根据估计量 $\hat{g}$ 的定义, 参数 $\delta_i$ 依赖于强相关参数 $\alpha$• 因此我们必须先估计参数 $\alpha$, 然后才能估计 $g$• 文献[21] 和文献[22] 介绍了如何用小波方法来估计参数 $\alpha$, 他们得到了估计量的无偏性、一致性和渐进正态性•

在定理 2.1, 我们假定回归函数 $g$ 其有 $r$ 阶连续导数, $g^{(r)}$• 实际上, 如果 $g^{(r)}$ 只是间断连续, 上述结论仍然成立• 这就是以下定理

定理 2.2 假定 $g^{(r)}$ 只是间断连续, 即存在 $x_0=0<x_1<x_2<\cdots<x_N<1=x_{N+1}$, 当 $x\in(x_i,x_{i+1}),0\leqslant i\leqslant N$, $g^{(r)}$ 连续, 有界, 并且存在左, 右极限, 又设 $p_q^{2r+\alpha}n^{-2\alpha r}\to\infty$, 那么定理 2.1 结论仍然成立•

注 2.3 文献[3] 在误差序列服从平稳时间序列情形下给出了回归函数的小波估计, 并且给出了均方差的渐近展开式, 他们的均方误差收敛率不同于本文的结果, 因为文[3] 的随机误差属于弱相关•

注 2.4 文献[5] 考虑了相类似的回归函数核估计, 但是在光滑函数的假定下, 得到均方误差的渐进表示, 即

$$E\int(\hat{g}-g)^2\sim C_2(n^{-1}p)^\alpha+p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int g^{(r)^2},\qquad n\to\infty,$$

但是, 即使在间断连续这个情形下, 我们证明了上述结论仍然成立•

## 3 定 理 证 明

定理 2.1 和 2.2 的证明非常类似于文献[13] 定理 2.1 和 2.2 的证明, 不同之处在于文献[13] 考虑独立的误差序列, 而我们考虑的是强相关误差序列•

记 $d_j=E(\hat{b}_j)=\dfrac{1}{n}\sum\limits_{k=1}^{n}g(x_k)\phi_j(x_k),\ d_{ij}=E(\hat{b}_{ij})=\dfrac{1}{n}\sum\limits_{k=1}^{n}g(x_k)\psi_{ij}(x_k)$•

引理 3.1 假设 $g$ 在[0, 1] 是连续可微函数, 并且小波 $\phi$ 和 $\psi$ 满足条件(3), 则

$$\sup_j|d_j-b_j|=O(n^{-1/2}) \tag{8}$$

$$\sup_j|d_{ij}-b_{ij}|=O(n^{-1/2})• \tag{9}$$

证明 这里只证明(8), 因为(9) 的证明很相似• 记

$$d_j=\frac{p^{1/2}}{n}\sum_{k=1}^{n}g\left(\frac{k}{n}\right)\phi\left(\frac{pk}{n}-j\right)• \tag{10}$$

对固定的 $n$、$p$ 和 $j$, 注意到

$$0\leqslant\frac{pk}{n}-j\leqslant1 \text{ 等同于} \frac{nj}{p}\leqslant k\leqslant\frac{n(j+1)}{p},$$

记 $m_j=\lceil nj/p\rfloor$, 这里 $\lceil x\rfloor$ 表示整数, 它的值至少是 $x$, 因为 $\phi$ 定义在[0, 1], (10) 的和是从 $m_j$ 到 $m_{j+1}-1$• 为方便起见, 以下不分 $\lceil x\rfloor$ 和 $x$• 因此

$$d_j=\begin{cases}\dfrac{p^{1/2}}{n}\displaystyle\sum_{k=m_j}^{m_{j+1}-1}g\left(\frac{k}{n}\right)\phi\left(\frac{pk}{n}-j\right) & (\text{当 } k=m_j+l),\\ \dfrac{p^{1/2}}{n}\displaystyle\sum_{l=0}^{n/p-1}g\left(\frac{l}{n}+\frac{j}{p}\right)\phi\left(\frac{pl}{n}\right) & (\text{当 } t_l=\frac{pl}{n}),\\ \dfrac{l}{p^{1/2}}\displaystyle\sum_{l=0}^{n/p-1}g\left(\frac{t_l+j}{p}\right)\phi(t_l)\frac{p}{n}.\end{cases}\tag{11}$$

类似地

$$b_j=\begin{cases}\displaystyle\int_0^1 g(x)\phi_j(x)\mathrm{d}x,\\ p^{1/2}\displaystyle\int_{j/p}^{(j+1)/p}g(x)\phi(px-j)\mathrm{d}x & (\text{当 } t=px-j),\\ \dfrac{1}{p^{1/2}}\displaystyle\int_0^1 g\left(\frac{t+j}{p}\right)\phi(t)\mathrm{d}t.\end{cases}\tag{12}$$

结合(11)和(12),我们得到

$$|d_j-b_j|=\frac{1}{p^{1/2}}\sum_{l=0}^{n/p-1}\int_{pl/n}^{p(l+1)/n}\left[g\left(\frac{t_l+j}{p}\right)\phi(t_l)-g\left(\frac{t+j}{p}\right)\phi(t)\right]\mathrm{d}t=J_1+J_2.\tag{13}$$

其中 $$J_1=\frac{1}{p^{1/2}}\sum_{l=0}^{n/p-1}\int_{pl/n}^{p(l+1)/n}\left[g\left(\frac{t_l+j}{p}\right)-g\left(\frac{t+j}{p}\right)\right]\phi(t_l)\mathrm{d}t$$

和 $$J_2=\frac{1}{p^{1/2}}\sum_{l=0}^{n/p-1}\int_{pl/n}^{p(l+1)/n}g\left(\frac{t+j}{p}\right)[\phi(t_l)-\phi(t)]\mathrm{d}t.$$

为了估计 $J_1$, 根据 $g$ 的连续可微性, $\phi$ 的有界性, 得到

$$J_1\leqslant(1/p^{1/2})(C/n),\tag{14}$$

其中 $C>0$, 是一个有限常数. 对于 $J_2$, 根据条件(3)

$$J_2\leqslant(1/p^{1/2})\cdot C\left(p/n\right)^{1/2}=Cn^{-1/2}.\tag{15}$$

因此,结合(14)和(15),引理证毕.

引理 3.2 在定理 2.1 的假设下

$$S_1\equiv E\left|\sum_{j=0}^{p-1}(\hat{b}_j-b_j)^2-C_2(n^{-1}p)^{\alpha}\right|=o((n^{-1}p)^{\alpha}).$$

证明 记 $Q_n=\sum_{j=0}^{p-1}(\hat{b}_j-d_j)^2$和 $\mu_n==E(Q_n)$.

注意到

$$\mu_n=\frac{1}{n^2}\sum_{j=0}^{p-1}\sum_{k=1}^{n}\sum_{l=1}^{n}E(\varepsilon_k\varepsilon_l)\phi_j(x_k)\phi_j(x_l)=$$

$$\frac{p}{n^2}\sum_{j=0}^{p-1}\sum_{k=1}^{n}\sum_{l=1}^{n}r(k-l)\phi(px_k-j)\phi(px_l-j).$$

对于固定的 $j=0,1,\dots,p-1$, 类似于(11),得到

$$\frac{p}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n}r(k-l)\phi(px_k-j)\phi(px_l-j)=$$

$$\frac{p}{n^2}\sum_{k=1}^{n/p-1}\sum_{l=1}^{n/p-1}r(k-l)\phi\left(\frac{pk}{n}\right)\phi\left(\frac{pl}{n}\right)=$$

$$p^{-1}C_0(n^{-1}p)^{\alpha}\left[\iint|x-y|^{-\alpha}\phi(x)\phi(y)\mathrm{d}x\mathrm{d}y+o(1)\right],\tag{16}$$

因此

$$\mu_n = \sum_{j=0}^{p-1} C_0 (n^{-1}p)^{\alpha} p^{-1} \iint | x - y |^{-\alpha} \phi(x) \phi(y) \mathrm{d}x \mathrm{d}y + o((n^{-1}p)^{\alpha}) = C_2 (n^{-1}p)^{\alpha} + o((n^{-1}p)^{\alpha}). \tag{17}$$

根据三角不等式和 Cauchy_Schwartz 不等式, 得到

$$S_1 \leqslant E^{1/2}(Q_n - \mu_n)^2 + \sum_{j=0}^{p-1}(d_j - b_j)^2 + 2\Big[\sum_{j=0}^{p-1} E(\hat{b}_j - d_j)^2 \sum_{j=0}^{p-1}(d_j - b_j)^2\Big]^{1/2} + o((n^{-1}p)^{\alpha}) = S_{11} + S_{12} + S_{13} + o((n^{-1}p)^{\alpha}). \tag{18}$$

根据引理 3.1 的(8)

$$\sup_j | d_j - b_j | = O(n^{-1/2}),$$

$$\begin{cases} S_{12} = O(n^{-1}p) = o((n^{-1}p)^{\alpha}), \\ S_{13} = O((n^{-1}p)^{1/2}\mu_n^{1/2}) = o((n^{-1}p)^{\alpha}). \end{cases} \tag{19}$$

至于第 1 项 $S_{11}$, 记

$$a_{kl} = p n^{-2} \sum_{j=0}^{p-1} \phi(px_k - j) \phi(px_l - j)$$

则 $$Q_n - \mu_n = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl} \varepsilon_k \varepsilon_l - \mu_n.$$

因为 $Q_n$ 是强相关高斯随机变量的一个二次式, 我们得利用 Wiener_Itô_Dobrushin 积分来估计上述 $Q_n$ 的方差. 文献[23] 和文献[24] 提供了简便的方法来计算上述高斯随机变量乘积的期望, 因为$\left\{ \varepsilon_k, k \geqslant 1 \right\}$是一个高斯平稳过程, 根据 Bochner 定理, 它的协方差函数 $r(k)$ 有如下表达式

$$r(k) = \int_{[-\pi,\pi]} \mathrm{e}^{\mathrm{i}kx} G(\mathrm{d}x), \qquad k \in \mathrm{Z},$$

其中 $G$ 是在$[-\pi,\pi]$ 上的一个 Borel 测度. 记 $Z_G$ 是对应的随机谱测度, 那么对所有的 Borel 集 $A \subset [-\pi,\pi]$, $E(Z_G(A))^2 = G(A)$, 所以对所有 $k \geqslant 1$,

$$\varepsilon_k = \int_{[-\pi,\pi]} \mathrm{e}^{\mathrm{i}kx} \mathrm{d}Z_G(x).$$

根据文献[23] 的引理 3.5, 当 $m = 1$, 得到

$$Q_n - \mu_n = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl} K_1(k, l),$$

其中 $K_1(k, l)$ 是 Wiener_Itô_Dobrushin 二重积分, 表示式为

$$K_1(k, l) = \int''_{[-\pi,\pi]^2} \mathrm{e}^{\mathrm{i}kx_1 + \mathrm{i}lx_2} \mathrm{d}Z_G(x_1) \mathrm{d}Z_G(x_2).$$

进一步, 根据文献[23] 的引理 3.6, 当 $m = 1$, 得到对任何整数 $k$、$s$、$l$、$t$

$$E[K_1(k, k+s) K_1(l, l+t)] = r(k-l+s) r(k-l-t) + r(k-l) r(k-l+s-t).$$

因此

$$E(Q_n - \mu_n)^2 = \sum_{k_1} \sum_{k_2} \sum_{l_1} \sum_{l_2} a_{k_1 l_1} a_{k_2 l_2} E[K_1(k_1, l_1) K_1(k_2, l_2)] = \sum_{k_1} \sum_{k_2} \sum_{l_1} \sum_{l_2} a_{k_1 l_1} a_{k_2 l_2} [r(l_1 - k_2) r(k_1 - l_2) + r(k_1 - k_2) r(l_1 - l_2)] =:$$

$A_{n1}+A_{n2}$.

我们先来估计第 1 项 $A_{n1}$. 根据类似于(16)和(17)的推导和 $r(k)\sim C_0k^{-\alpha}, k\to\infty$, 我们得到

$$\begin{aligned}A_{n1}=&\ p^2n^{-4}\sum_{k_1}\sum_{k_2}\sum_{l_1}\sum_{l_2}\sum_{j_1}\sum_{j_2}r(l_1-k_2)r(k_1-l_2)\times\\&\phi(px_{k_1}-j_1)\phi(px_{l_1}-j_1)\phi(px_{k_2}-j_2)\phi(px_{l_2}-j_2)=\\&C_0^2p^{2\alpha-2}n^{-2\alpha}\sum_{j_1}\sum_{j_2}\left[\iint|x-y+j_1-j_2|^{-\alpha}\phi(x)\phi(y)\mathrm{d}x\mathrm{d}y\right]^2+\\&o((n^{-1}p)^{2\alpha})=:A_{n1}(1)+o((n^{-1}p)^{2\alpha}).\end{aligned}$$

记 $k=j_1-j_2$, 并交换和的次序, 得到

$$\begin{aligned}A_{n1}(1)=&\ C_0^2p^{2\alpha-2}n^{-2\alpha}\left\{\sum_{k=0}^{p-1}(p-k)\left[\iint|x-y+k|^{-\alpha}\phi(x)\phi(y)\mathrm{d}x\mathrm{d}y\right]^2+\right.\\&\left.\sum_{k=-(p-1)}^{-1}(p+k)\left[\iint|x-y+k|^{-\alpha}\phi(x)\phi(y)\mathrm{d}x\mathrm{d}y\right]^2\right\}.\end{aligned}$$

因为 $\phi(x)$ 有界, 它的定义域有限

当 $k\to\infty$

$$\left|\iint|x-y+k|^{-\alpha}\phi(x)\phi(y)\mathrm{d}x\mathrm{d}y\right|=O(|k|^{-\alpha}),\qquad k\to\infty.$$

所以 $\quad A_{n1}(1)=O(p^{2\alpha-2}n^{-2\alpha}p^{2-2\alpha})=O((n^{-1}p)^{2\alpha}p^{-2\alpha})=o((n^{-1}p)^{2\alpha})$.

因此, $A_{n1}=o((n^{-1}p)^{2\alpha})$. 类似地我们可以证明 $A_{n2}=o((n^{-1}p)^{2\alpha})$. 从而 $S_{11}=o((n^{-1}p)^{\alpha})$. 结合(18)和(19), 引理证毕.

引理 3.3 在定理 2.1 的假设下

$$S_2\equiv\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}E\left\{(\hat b_{ij}-b_{ij})^2I(|\hat b_{ij}|>\delta_i)\right\}=o((n^{-2\alpha r/(2r+\alpha)}).$$

证明 引理的证明类似于文献[13]定理 2.1 的证明. 记 $\lambda$和$\beta$是任意正数满足$\lambda+\beta=1$, 记

$$S_{21}'=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}E\left\{(\hat b_{ij}-b_{ij})^2\right\}I(|b_{ij}|>\lambda\delta_i),$$

$$S_{22}'=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}E\left\{(\hat b_{ij}-b_{ij})^2I(|\hat b_{ij}-b_{ij}|>\beta\delta_i)\right\}.$$

根据三角不等式, 得到 $S_2\leqslant S_{21}'+S_{22}'$, 记

$$S_{21}=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}E\left\{(\hat b_{ij}-d_{ij})^2\right\}I(|b_{ij}|>\lambda\delta_i),$$

$$S_{22}=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}E\left\{(\hat b_{ij}-d_{ij})^2I(|\hat b_{ij}-d_{ij}|>\beta\delta_i)\right\}.$$

根据引理 3.1 的(8), 和 $n\delta_i^2\to\infty$以及类似于(18)的证明, 我们可以证明 $S_{21}$ 和 $S_{22}$ 是 $S_{21}'$ 和 $S_{22}'$ 的主要项. 即 $S_{21}'=O(S_{21})$ 和 $S_{22}'=O(S_{22})$. 因此为了证明引理 3.3, 我们只要证明

$$S_{21}=o((n^{-2\alpha r/(2r+\alpha)})\ \text{和}\ S_{22}=o((n^{-2\alpha r/(2r+\alpha)}).$$

因为我们假定回归函数 $g$ 是$r$ 阶连续可微, 那么根据Taylor 展开, 以及小波 $\psi(x)$ 的矩条件

$$b_{ij}=p^{-1/2}\int\psi(y)\left[\sum_{v=0}^{r-1}\frac{1}{v!}\left(\frac{y}{p_i}\right)^vg^{(v)}\left(\frac{j}{p_i}\right)+\right.$$

$$\frac{1}{(r-1)!}\left(\frac{y}{p_i}\right)^r\int_0^1(1-t)^{r-1}g^{(r)}\left(\frac{j+ty}{p_i}\right)\mathrm{d}t\Big]\mathrm{d}y=$$

$$\kappa p_i^{-(r+1/2)}(g_{ij}+\xi_{ij}),\tag{20}$$

其中 $g_{ij}=g^{(r)}\left(\frac{j}{p_i}\right)$ 和 $\sup_{0\leqslant j\leqslant p_i-1;0\leqslant i\leqslant q-1}|\xi_{ij}|\to 0$.

因此 $|b_{ij}|\leqslant Cp_i^{-(r+1/2)}$. 通过类似引理 3.2 $\mu_n$ 的计算, 得到

$$E(\hat{b}_{ij}-d_{ij})^2\leqslant Cp_i^{-1}(n^{-1}p_i)^{\alpha},$$

因此

$$S_{21}\leqslant\sum_{i=0}^{q-1}C(n^{-1}p_i)^{\alpha}I(p_i\leqslant C\delta_i^{-2/(2r+1)})=$$

$$Cn^{-\alpha}p^{\alpha}\sum_{i=0}^{q-1}2^{\alpha i}(p_i\leqslant C\delta_i^{-2/(2r+1)}).\tag{21}$$

注意到 $p_i\leqslant C\delta_i^{-2/(2r+1)}$, 那么 $(p2^i)^{2r+1}\leqslant Cn^{\alpha}p^{1-\alpha}(\ln n)^{-1}2^{(1-\alpha)i}$, 因此

$$2^{(2r+\alpha)i}\leqslant C\frac{n^{\alpha}p^{-(2r+\alpha)}}{\ln n}.$$

实际上只有有限项 $i$ 满足上述不等式, 记最大的 $i$ 值为 $m$, 那么(21)变为

$$S_{21}\leqslant Cn^{-\alpha}p^{\alpha}\sum_{i=0}^{m}2^{\alpha i}\leqslant Cn^{-\alpha}p^{\alpha}2^{\alpha m}\leqslant Cn^{-\alpha}p^{\alpha}\left[\frac{n^{\alpha}p^{-(2r+\alpha)}}{\ln n}\right]^{\alpha/(2r+\alpha)}=$$

$$C(n^{2r}\ln n)^{-\alpha/(2r+\alpha)}=o(n^{-2r\alpha/(2r+\alpha)}).\tag{22}$$

为了估计 $S_{22}$, 根据 Cauchy_Schwartz 不等式

$$S_{22}\leqslant\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}E^{1/2}[(\hat{b}_{ij}-d_{ij})^4]P^{1/2}(|\hat{b}_{ij}-d_{ij}|>\beta\delta_i).$$

因为 $\hat{b}_{ij}-d_{ij}$ 是中心化的正态随机变量, 它的二阶矩为

$$E[(\hat{b}_{ij}-d_{ij})^2]=C_1n^{-\alpha}p_i^{-1+\alpha}+o(n^{-\alpha}p_i^{-1+\alpha}),$$

可以得到

$$E[(\hat{b}_{ij}-d_{ij})^4]\leqslant CE^2[(\hat{b}_{ij}-d_{ij})^2]\leqslant C(n^{-1}p_i)^{2\alpha}p_i^{-2},$$

根据正态随机变量的概率不等式, 有

$$P(|\hat{b}_{ij}-b_{ij}|>\beta\delta_i)\leqslant P(|Z|>(1-\epsilon)\beta\sqrt{4\ln n})\leqslant$$

$$\left\{C/[(1-\epsilon)\beta\sqrt{4\ln n})]\right\}\mathrm{e}^{-2(1-\epsilon)^2\beta^2\ln n}=O(n^{-2(1-\epsilon)^2\beta^2}).$$

其中 $Z$ 是标准正态随机变量, 因此我们得到

$$S_{22}\leqslant C\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}(n^{-1}p_i)^{\alpha}p_i^{-1}n^{-(1-\epsilon)^2\beta^2}=Cn^{-(1-\epsilon)^2\beta^2}(n^{-1}p_q)^{\alpha}.$$

根据(SP)中的假设 $p_i\delta_i^2\to 0$, 得到 $n^{-1}p_q\to 0$. 选择 $\beta<1$, 但接近 1, $\varepsilon$ 接近 0, 那么 $S_{22}=o(n^{-2\alpha r/(2r+\alpha)})$, 结合(22), 引理证毕.

引理 3.4 在定理 2.1 的假设下

$$S_3\equiv E\left|\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}b_{ij}^2I(|\hat{b}_{ij}|\leqslant\delta_i)-p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int g^{(r)^2}\right|=o(p^{-2r}).$$

证明 记 $\varepsilon$ 是任意小的正数 $\varepsilon>0$, 和

$$S_{30}=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}b_{ij}^2I(|\hat{b}_{ij}|\leqslant\delta_i),\quad S_{31}=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}b_{ij}^2I(|b_{ij}|\leqslant(1+\epsilon)\delta_i),$$

$$S_{32}=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}b_{ij}^2I(|b_{ij}|\leqslant(1-\epsilon)\delta_i),\quad\Delta=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}b_{ij}^2I(|\hat{b}_{ij}-b_{ij}|\geqslant\epsilon\delta_i).$$

根据三角不等式，得到

$$S_{32}-\Delta\leqslant S_{30}\leqslant S_{31}+\Delta. \tag{23}$$

根据(20)和(SP)中的假设 $p_i^{2r+1}\delta_i^2\to\infty$，得到，当 $n$ 充分大时

$$I\left\{|b_{ij}|\leqslant(1+\epsilon)\delta_i\right\}=I\left\{|b_{ij}|\leqslant(1-\epsilon)\delta_i\right\}=1,$$

以及 $$S_{31}=S_{32}=p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int g^{(r)^2}+o(p^{-2r}). \tag{24}$$

至于证明细节，可以参考文献[13](920页)。同时，类似于 $S_{22}$ 的证明，得到

$$E\Delta=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}b_{ij}^2P(|\hat b_{ij}-b_{ij}|\geqslant\epsilon\delta_i)\leqslant Cn^{-\epsilon^2}\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}b_{ij}^2=o(p^{-2r}).$$

结合(23)和(24)，引理证毕.

**引理 3.5** 在定理 2.1 的假设下

$$S_4\equiv\sum_{i=q}^{\infty}\sum_{j=0}^{p_i-1}b_{ij}^2=o(p^{-2r}).$$

证明 根据(20) $b_{ij}$ 的Taylor 展开和 $q\to\infty$ 的假设，很容易得到上述结果.

**定理 2.1 的证明** 根据引理 3.2 到引理 3.5 以及下面的不等式

$$E\left|\int(\hat g-g)^2-\left\{C_1(n^{-1}p)^{\alpha}+p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int g^{(r)^2}\right\}\right|\leqslant$$
$$S_1+S_2+S_3+S_4.$$

很容易推得定理 2.1 的结论.

**定理 2.2 的证明** 我们仍然使用文献[13]的记号，根据小波 $\phi$ 和 $\psi$ 的正交性，得到

$$\int(\hat g-g)^2=I_q(\mathrm{Z},\mathrm{Z},\cdots),$$

其中 Z 表示对应的整数集，例如 $\Psi_i=\left\{0,1,\cdots,p_{i-1}\right\}$，那么

$$\begin{aligned}I_q(\Psi,\Psi_0,\Psi_1,\cdots)=&\sum_{j\in\Psi}(\hat b_j-b_j)^2+\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}(\hat b_{ij}-b_{ij})^2I(|\hat b_{ij}|>\delta_i)+\\&\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}b_{ij}^2I(|\hat b_{ij}|\leqslant\delta_i)+\sum_{i=q}^{\infty}\sum_{j\in\Psi_i}b_{ij}^2=\\&\sum_{j\in\Psi}(\hat b_j-d_j)^2+\sum_{j\in\Psi}(d_j-b_j)^2+2\sum_{j\in\Psi}(\hat b_j-d_j)(d_j-b_j)+\\&\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}(\hat b_{ij}-d_{ij})^2I(|\hat b_{ij}|>\delta_i)+\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}(d_{ij}-b_{ij})^2I(|\hat b_{ij}|>\delta_i)+\\&2\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}(\hat b_{ij}-d_{ij})(d_{ij}-b_{ij})I(|\hat b_{ij}|>\delta_i)+\\&\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}b_{ij}^2I(|\hat b_{ij}|\leqslant\delta_i)+\sum_{i=q}^{\infty}\sum_{j\in\Psi_i}b_{ij}^2=:\\&I_1+I_2+I_3+I_4+I_5+I_6+I_7+I_8.\end{aligned}$$

根据引理 3.1，很容易得到 $E(I_2)=o((n^{-1}p)^{\alpha})$ 和 $E(I_5)=O(E(I_4))$. 下面我们将证明，不管回归系数 $g$ 是连续可微或间断连续可微，都有 $E(I_1)=O((n^{-1}p)^{-\alpha})$ 和 $E(I_4)=o(n^{-2\alpha r/(2r+\alpha)})$. 因此，根据 Cauchy_Schwartz 不等式，可以证明 $E(I_3)=o((n^{-1}p)^{\alpha})$ 和 $E(I_6)=o(n^{-2\alpha r/(2r+\alpha)})$，与 $(n^{-1}p)^{\alpha}$ 相比，这个阶是可忽略的.

因此我们只要考虑 $I_1$、$I_4$、$I_7$ 和 $I_8$ 4 项，当回归函数 $g$ 只是间断光滑，记 $\Pi$ 是一个集合包含 $g^{(s)}$ 的间断点，$0\leqslant s\leqslant r$. 又设，对有些 $x\in\Pi$，

$$\mathrm{K}=\left\{k: k \in(p x-v, p x+v)\right\}, \quad \text{当 } x \in \Pi,$$
$$\mathrm{K}_i=\left\{k: k \in(p_i x-v, p_i x+v)\right\}, \quad \text{当 } x \in \Pi,$$

$\mathrm{K}^c$ 和 $\mathrm{K}_i^c$ 表示K和 $\mathrm{K}_i$ 的补集• 那么, 除非 $j \in \mathrm{K}_i$, $b_{ij}$ 中的 $g$ 是 $r$ 阶连续可微• 同样, 当 $j \notin \mathrm{K}_i$, $\hat{b}_{ij}$ 是 $n$ 个独立观察值的平均值, 而这些观察值定义在区间上, $g^{(r)}$ 是连续可微的• 类似地, 除非 $j \in \mathrm{K}$, $b_i$ 和 $\hat{b}_j$ 定义在同样区域上, 因此

$$\begin{aligned} I_q(\Psi, \Psi_0, \Psi_1, \dots) = {} & I_1(\mathrm{K})+I_2+I_3+I_4(\mathrm{K}_0, \mathrm{K}_1, \mathrm{K}_2, \dots)+I_5+I_6+ \\ & I_7(\mathrm{K}_0, \mathrm{K}_1, \mathrm{K}_2, \dots)+I_8(\mathrm{K}_0, \mathrm{K}_1, \mathrm{K}_2, \dots)+I_1(\mathrm{K}^c)+I_4(\mathrm{K}_0^c, \mathrm{K}_1^c, \mathrm{K}_2^c, \dots)+ \\ & I_7(\mathrm{K}_0^c, \mathrm{K}_1^c, \mathrm{K}_2^c, \dots)+I_8(\mathrm{K}_0^c, \mathrm{K}_1^c, \mathrm{K}_2^c, \dots), \end{aligned} \tag{25}$$

其中

$$I_1(\mathrm{K})=\sum_{j \in \mathrm{K}}(\hat{b}_j-d_j)^2, \quad I_1(\mathrm{K}^c)=\sum_{j \in \mathrm{K}^c}(\hat{b}_j-d_j)^2,$$

$$I_4(\mathrm{K}_0, \mathrm{K}_1, \mathrm{K}_2, \dots)=\sum_{i=0}^{q-1} \sum_{j \in \mathrm{K}_i}(\hat{b}_{ij}-d_{ij})^2 I(|\hat{b}_{ij}|>\delta_i),$$

$$I_4(\mathrm{K}_0^c, \mathrm{K}_1^c, \mathrm{K}_2^c, \dots) \sum_{i=0}^{q-1} \sum_{j \in \mathrm{K}_i^c}(\hat{b}_{ij}-d_{ij})^2 I(|\hat{b}_{ij}|>\delta_i),$$

其他项可以相应地推得• 然而, 由于小波 $\phi$ 和 $\psi$ 定义在有限区间上, 对每一个 $i$, K和 $\mathrm{K}_i$ 至多有 $(2v+1)(\# \Pi)$ 个元素• 考虑到 $q=O(\ln n)$, 我们可以证明 $I_1(\mathrm{K})$, $I_4(\mathrm{K}_0, \mathrm{K}_1, \mathrm{K}_2, \dots)$ 和 $I_7(\mathrm{K}_0, \mathrm{K}_1, \mathrm{K}_2, \dots)$ 至多 $o((n^{-1}p)^{\alpha})$• 因此这些项是可以忽略的• 虽然当 $g$ 不是 $r$ 阶连续可微时, $b_{ij}=O(p_i^{-1/2})$, 但是根据定理2.2的补充假设 $p_q^{2r+\alpha_n-2\alpha r} \to \infty$, 我们可以证得 $I_8(\mathrm{K}_0, \mathrm{K}_1, \mathrm{K}_2, \dots)=o(n^{-2\alpha r/(2r+\alpha)})$• 根据整个定理2.1的证明, 我们可以看到等式(25)右边的其他项正好等于定理2.2中 $\int(\hat{g}-g)^2$ 的主要项• 因此定理证毕•

## [参 考 文 献]


[1] Hart J D. Kernel regression estimation with time series errors[J]. J Roy Statist Soc, Ser B, 1991, **53**: 173—187.

[2] Tran L T, Roussas G G, Yakowitz S, et al. Fixed_design regression for linear time series[J]. Ann Statist, 1996, **24**: 975—991.

[3] Truong Y K, Patil P N. Asymptotics for wavelet based estimates of piecewise smooth regression for stationary time series[J]. Ann Inst Statist Math, 2001, **53**: 159—178.

[4] Beran J. Statistics for Long Memory Processes[M]. New York: Chapman and Hall, 1994.

[5] Hall P, Hart J D. Nonparametric regression with long_range dependence[J]. Stochastic Process Appl, 1990, **36**: 339—351.

[6] Robinson P M, Hidalgo F J. Time series regression with long_range dependence[J]. Ann Statist, 1997, **25**: 77—104.

[7] Csorgo S, Mielniczuk J. Nonparametric regression under long_range dependent normal errors[J]. Ann Statist, 1995, **23**: 1000—1014.

[8] Robinson P M. Large_sample inference for nonparametric regression with dependent errors[J]. Ann Statist, 1997, **25**: 2054—2083.

[9] Hardle W, Kerkyacharian G, Picard D, et al. Wavelets, Approximation and Statistical Applications [M]. Lecture Notes in Statistics, 129. New York: Springer_Verlag, 1998.

[10] Donoho D L, Johnstone I M. Minimax estimation via wavelet shrinkage[J]. Ann Statist, 1998, **26**: 879—921.

[11] Donoho D L, Johnstone I M, Kerkyacharian G, et al. Wavelet shrinkage: asymptopia? (with discussion)[J]. J Roy Statist Soc, Ser B, 1995, **57**: 301—369.

[12] Donoho D L, Johnstone I M, Kerkyacharian G, et al. Density estimation by wavelet thresholding[J]. Ann Statist, 1996, **24**: 508—539.

[13] Hall P, Patil P. Formulae for mean integated squared error of non_linear waveletbased density estimators[J]. Ann Statist, 1995, **23**: 905—928.

[14] Hall P, Patil P. On the choice of smoothing parameter, threshold and truncation in nonparametric regression by nonlinear wavelet methods[J]. J Roy Statist Soc, Ser B, 1996, **58**: 361—377.

[15] Hall P, Patil P. Effect of threshold rules on performance of wavelet_based curve estimators[J]. Statistic Sinica, 1996, **6**: 331—345.

[16] Johnstone I M. Wavelet threshold estimators for correlated data and inverse problems: Adaptivity results[J]. Statistica Sinica, 1999, **9**: 51—83.

[17] Johnstone I M, Silverman B W. Wavelet threshold estimators for data with correlated noise[J]. J Roy Statist Soc, Ser B, 1997, **59**: 319—351.

[18] Wang Y. Function estimation via wavelet shrinkage for long_memory data[J]. Ann Statist, 1996, **24**: 466—484.

[19] Daubechies I. Ten Lectures on Wavelets[M]. Philadelphia: SIAM, 1992.

[20] Cohen A, Daubechies I, Vial P. Wavelets on the interval and fast wavelet transforms[J]. Appl Comput Harm Anal, 1993, **1**: 54—82.

[21] Abry P, Veitch D. Wavelet analysis of long_range_dependent traffic[J]. IEEE Trans on Inform Theory, 1998, **44**: 2—15.

[22] Delbeke L, Van Assche Walter. A wavelet based estimator for the parameter of selfsimilarity of fractional Brownian motion[A]. In: 3rd International Conference on Approximation and Optimization in the Caribbean[C]. Puebla: 1995, Soc Mat Mexicana, Ḿexico: Aportaciones Mat. Comun 24, 1998, 65—76.

[23] Fox R, Taqqu M. Noncentral limit theorems for quadratic forms in random variables having long_range dependence[J]. Ann Probab, 1985, **13**: 428—446.

[24] Mojor P. Multiple Wiener_Itô Integrals[M]. Lect Notes in Math 849, New York: Springer_Verlag, 1981.

[25] Durrett R. Probability: Theorey and Examples[M]. Second Edition: Duxbury Press, 1996.

# Wavelet_Based Estimators of the Mean Regression Function With Long Memory Date

LI Lin_yuan, XIAO Yi_min

(1. Department of Mathematics and Statistics, University of New Hampshire, USA;
2. Department of Statistics and Probability, Michigan State Univeristy, USA)

**Abstract:** An asymptotic expansion is provide for the mean integrated squared error(MISE) of nonlinear wavelet_based mean regression function estimators with long memory data. This MISE expansion is shown, when the underlying mean regression function is only piecewise smooth. It is the same with analogous expansion for the kernel estimators. However, for the kernel estimators, this MISE expansion generally fails if the additional smoothness assumption is absent.

**Key words:** mean integrated square error; nonparametric regression; nonlinear wavelet_based estimator; rate of convergence