Supplementary Materials for "Monotone Nonparametric Regression for Functional/Longitudinal Data"

Ziqi Chen[1,2], Qibing Gao[3], Bo Fu[4] and Hongtu Zhu[1,5]

[1] *The University of Texas MD Anderson Cancer Center,*

[2] *Central South University,* [3] *Nanjing Normal University,* [4] *Fudan University*

[5] *The University of North Carolina at Chapel Hill*

1. Appendix

1.1 Assumptions

We present all the assumptions as follows.

(A) Kernel function.

$K_r(\cdot)$ is assumed to be a symmetric probability density function on $[-1, 1]$ and K_r is twice continuously differentiable on its support such that

$$\kappa_2(K_r) < \infty, \qquad \int K_r^2(u)du < \infty.$$

The assumptions on K_d are the same as those on K_r.

(B) Time points and true functions

(B1) $\{s_{ij} : i = 1, \ldots, n; j = 1, \ldots, n_i\}$ are i.i.d. copies of a random variable S defined on $[0, 1]$. The density $f(\cdot)$ of S is bounded from below and above with $0 < m_f \le \min_{s\in[0,1]} f(s) \le \max_{s\in[0,1]} f(s) \le M_f < \infty$ and $\ddot{f}(s)$, the second derivative of $f(\cdot)$, is continuous on $[0, 1]$.

(B2) $\ddot{m}(s)$, the second derivative of $m(s)$, is continuous on $[0, 1]$.

(B3) $\ddot{\sigma}(s)$, the second derivative of $\sigma(\cdot)$, is continuous on $[0, 1]$.

(B4) $\{\eta_i(\cdot)\}_i$ are i.i.d. copies of $\eta(\cdot)$ and $\{\varepsilon_{ij}\}_{ij}$ are i.i.d. copies of ε. Furthermore, $E(\varepsilon) = 0$, $E(\varepsilon^2) = 1$.

(B5) X is independent of S and ε is independent of S and η.

(B6) $\partial^2\gamma(s, t)/\partial s^2$, $\partial^2\gamma(s, t)/\partial s\partial t$ and $\partial^2\gamma(s, t)/\partial t^2$ are continuous on $[0, 1]^2$.

(C) Bandwidths and moments

(C1) $h_r \to 0$, $h_d \to 0$, $h_r^2/h_d \to 0$, $h_r/h_d \to \infty$, $h_d/\log(n)^2 = O(1)$,

$h_d^2 h_r^{-8} \max\Big\{ \sum_{i=1}^n n_i\omega_i^2/h_r, \sum_{i=1}^n \omega_i^2 n_i(n_i - 1)\Big\} \to \infty$,

$\log(n)^2 h_r^{-2} h_d^{-1} \max\{\sum_{i=1}^n n_i\omega_i^2/h_r, \sum_{i=1}^n \omega_i^2 n_i(n_i - 1)\} \to 0$,

$\sum_{j=1}^n \omega_j^4(n_j^4 + n_j^3/h_d + n_j^2/h_d^2 + n_j/h_d^3)\{\sum_{j=1}^n \omega_j^2 n_j/h_r + \sum_{j=1}^n \omega_j^2 n_j(n_j - 1)\}^{-2} \to 0$.

(C2) $E(\varepsilon^5) < \infty$, $E\sup_{s\in[0,1]} \eta^5(s) < \infty$, and $E\eta^5(s)$ is continuous on $[0, 1]$.

(C3)

$$n\left\{\sum_{i=1}^{n} n_i\omega_i^2 h_r + \sum_{i=1}^{n} n_i(n_i-1)\omega_i^2 h_r^2\right\}\{\log(n)/n\}^{-3/5} \to \infty.$$

(C4) $\sup_n(n \max_i n_i\omega_i) \leq B < \infty$.

1.2 Lemmas and Proofs

Define $m_N^{-1}(t) := N^{-1}\int_{-\infty}^{t}\sum_{i=1}^{N} K_{d,h_d}\left(m\left(i/N\right)-u\right)du$, and let $m_N(t)$ denote its inverse. Let $\dot{K}_{d,h}(v) := h^{-2}\dot{K}_d(v/h)$ and $\ddot{K}_{d,h}(v) := h^{-3}\ddot{K}_d(v/h)$ with $\dot{K}_d(\cdot)$ and $\ddot{K}_d(\cdot)$ being the first and second derivatives of $K_d(\cdot)$, respectively. Let $(\ddot{m^{-1}})(t) := d^2m^{-1}(t)/dt^2$.

The following two lemmas are essentially Lemmas 2.1 and 2.2 in Dette et al. (2006), respectively.

Lemma 1. *If function $m(\cdot)$ is strictly increasing, continuous and twice continuously differentiable, then for each $t \in [m(0), m(1)]$ with $\dot{m}(m^{-1}(t)) > 0$, we have*

$$m_N^{-1}(t) = m^{-1}(t) + \kappa_2(K_d)h_d^2(\ddot{m^{-1}})(t) + o(h_d^2) + O\left\{(Nh_d)^{-1}\right\}.$$

Lemma 2. *If function $m(\cdot)$ is strictly increasing, continuous and twice continuously differentiable, then for each $s \in [0,1]$ with $\dot{m}(s) > 0$, we have*

$$m_N(s) = m(s) + \kappa_2(K_d)h_d^2\frac{\ddot{m}(s)}{\dot{m}^2(s)} + o(h_d^2) + O\left\{(Nh_d)^{-1}\right\}.$$

The next lemma gives the uniform convergence rate of $\widehat{m}(s)$.

Lemma 3. *Assuming that Conditions (A)–(C) hold, then we have*

$$\sup_{s\in[0,1]} |\widehat{m}(s) - m(s)| = O\left(h_r^2 + \left[\log(n) \left\{ \sum_{i=1}^{n} n_i\omega_i^2/h_r + \sum_{i=1}^{n} n_i(n_i-1)\omega_i^2 \right\} \right]^{\frac{1}{2}} \right) \; a.s.$$

This assertion can be proven following the same argument for the proof of Theorem 3.1 in Li & Hsing (2010).

The next lemma gives the consistency properties of $\widehat{m}_I^{-1}(s)$ and $\widehat{m}_I(s)$.

Lemma 4. *Assume that Conditions (A)–(C) hold. If function $m(\cdot)$ is strictly increasing,*

$$\widehat{m}_I^{-1}(t) = m_N^{-1}(t) + o_p(1)$$

holds uniformly for $t \in [m(0), m(1)]$ with $\dot{m}(m^{-1}(t)) > 0$ and

$$\widehat{m}_I(s) = m_N(s) + o_p(1),$$

holds uniformly for $s \in [0,1]$ with $\dot{m}(s) > 0$.

Proof. There exists a constant M_d such that $\sup_u K_d(u) < M_d$ by the assumptions on the kernel function. We have the decomposition

$$\widehat{m}_I^{-1}(t) = N^{-1} \int_{-\infty}^{t} \sum_{i=1}^{N} K_{d,h_d}\left(\widehat{m}\left(i/N\right) - u\right) du \;\; = \;\; m_N^{-1}(t) + \Delta_N(t),$$

where

$$
\begin{aligned}
|\Delta_N(t)| &= \left| N^{-1}\sum_{i=1}^{N}\int_{-\infty}^{t}\left\{K_{d,h_d}\left(\widehat{m}\left(i/N\right)-u\right)-K_{d,h_d}\left(m\left(i/N\right)-u\right)\right\}du\right| \\
&= \left| N^{-1}\sum_{i=1}^{N}\int_{-\infty}^{t}\dot{K}_{d,h_d}\left(\xi_i-u\right)\left\{\widehat{m}\left(i/N\right)-m\left(i/N\right)\right\}du\right| \\
&= \left| -N^{-1}\sum_{i=1}^{N}K_{d,h_d}\left(\xi_i-t\right)\left\{\widehat{m}\left(i/N\right)-m\left(i/N\right)\right\}\right| \\
&\leq N^{-1}\sum_{i=1}^{N}\left|K_{d,h_d}\left(\xi_i-t\right)\right|\sup_{s\in[0,1]}\left|\widehat{m}\left(s\right)-m\left(s\right)\right| \\
&\leq M_d h_d^{-1}\sup_{s\in[0,1]}\left|\widehat{m}\left(s\right)-m\left(s\right)\right|
\end{aligned}
$$

in which we have used $|\xi_i - m(i/N)| < |\widehat{m}(i/N) - m(i/N)|$ for $i = 1, \ldots, N$. Define $l_n = h_r^2 + [\log(n)\{\sum_{i=1}^{n} n_i\omega_i^2/h_r + \sum_{i=1}^{n} n_i(n_i-1)\omega_i^2\}]^{\frac{1}{2}}$. It follows from Lemma 3 and Condition (C1) that $\sup_t |\Delta_n(t)| = O_p(l_n/h_d) = o_p(1)$ holds. This finishes the proof of the first equation of Lemma 4.

By using Taylor expansion of Lemma 2.2 in Dette et al. (2006) and similar arguments, we have

$$\widehat{m}_I(s) - m_N(s) = O_p(l_n/h_d)$$

holding uniformly for $s \in [0,1]$ with $\dot{m}(s) > 0$, which implies the second equation of Lemma 4. □

Lemma 5. *Suppose that all assumptions of Theorem 1 hold. Then, for a* *fixed interior point* $t \in (m(0), m(1))$ *satisfying* $\dot{m}(m^{-1}(t)) > 0$, *we have*

$$\Gamma^{-1/2}\left\{\widehat{m}_I^{-1}(t) - m_N^{-1}(t) + h_r^2\kappa_2(K_r)\left(\frac{\ddot{m}}{\dot{m}}\right)(m^{-1}(t))\right\} \xrightarrow{d} N(0,1), \quad (1.1)$$

where $\Gamma = \Gamma_0^A \sum_{j=1}^n \omega_j^2 n_j/h_r + \Gamma_0^B \sum_{j=1}^n \omega_j^2 n_j(n_j-1)$.

Proof. Without loss of generality, we assume that the number of design points N equals the sample size n. We use the decomposition

$$\widehat{m}_I^{-1}(t) = n^{-1}\int_{-\infty}^{t}\sum_{i=1}^{N} K_{d,h_d}(\widehat{m}(i/n) - u)\, du \quad = \quad m_n^{-1}(t) + \Delta_n(t),$$

where

$$\Delta_n(t) \quad = \quad n^{-1}\sum_{i=1}^{n}\int_{-\infty}^{t}\{K_{d,h_d}(\widehat{m}(i/n) - u) - K_{d,h_d}(m(i/n) - u)\}\, du.$$

For $\Delta_n(t)$, by Taylor expansion, it follows that

$$\Delta_n(t) = \Delta_n^{(1)}(t) + \Delta_n^{(2)}(t)/2$$

with $\Delta_n^{(1)}(t) = n^{-1}\sum_{i=1}^n \int_{-\infty}^t \dot{K}_{d,h_d}(m(i/n) - u)\{\widehat{m}(i/n) - m(i/n)\}\, du$ and

$$\begin{aligned}\Delta_n^{(2)}(t) \quad &= \quad n^{-1}\sum_{i=1}^{n}\int_{-\infty}^{t}\ddot{K}_{d,h_d}(\xi_i - u)\{\widehat{m}(i/n) - m(i/n)\}^2\, du \\ &= \quad -n^{-1}\sum_{i=1}^{n}\dot{K}_{d,h_d}(\xi_i - t)\{\widehat{m}(i/n) - m(i/n)\}^2,\end{aligned}$$

where $|\xi_i - m(i/n)| < |\widehat{m}(i/n) - m(i/n)|, i = 1, \ldots, n$. A straight calculation shows that

$$\begin{aligned}|\Delta_n^{(2)}(t)| \quad &= \quad \left|\frac{1}{n}\sum_{i=1}^{n}\dot{K}_{d,h_d}(\xi_i - t)\{\widehat{m}(i/n) - m(i/n)\}^2\right| \\ &= \quad \left|\int_0^1 \dot{K}_{d,h_d}(m(x) - t)\{\widehat{m}(x) - m(x)\}^2 dx\right|\{1 + o_p(1)\}.\end{aligned}$$

By Lemma 3, we have $\Delta_n^{(2)}(t) = O_p(l_n^2/h_d)$, where $l_n = h_r^2 + [\log(n)\{\sum_{i=1}^n n_i\omega_i^2/h_r + \sum_{i=1}^n n_i(n_i-1)\omega_i^2\}]^{\frac{1}{2}}$. It follows from Condition (C1) that $\Delta_n^{(2)}(t)$ is of an order smaller than the orders of the bias term in (1.2) and the square root of the variance term in (1.4), respectively.

To finish the proof of Lemma 5, we only need to prove

$$\Gamma^{-1/2}\left\{\Delta_n^{(1)}(t) + h_r^2\kappa_2(K_r)\left(\frac{\ddot{m}}{\dot{m}}\right)(m^{-1}(t))\right\} \xrightarrow{d} N(0,1).$$

We have, by simple calculations,

$$\widehat{m}(s) = \frac{R_0S_2 - R_1S_1}{S_0S_2 - S_1^2},$$

where for $k = 0, 1, 2,$

$$S_k = \sum_{i=1}^n \omega_i \sum_{j=1}^{n_i} K_{r,h_r}(s_{ij}-s)(\frac{s_{ij}-s}{h_r})^k, \quad R_k = \sum_{i=1}^n \omega_i \sum_{j=1}^{n_i} K_{r,h_r}(s_{ij}-s)(\frac{s_{ij}-s}{h_r})^k y_{ij}.$$

Thus, we can have

$$\Delta_n^{(1)}(t) = \left(\Delta_n^{(1.1)}(t) + \Delta_n^{(1.2)}(t)\right)\{1 + o_p(1)\},$$

where

$$\begin{aligned}
\Delta_n^{(1.1)}(t) &= -n^{-1}\sum_{i,j=1}^n K_{d,h_d}(m(i/n)-t)\,\omega_j \sum_{k=1}^{n_j} K_{r,h_r}(s_{jk}-i/n)\,\frac{m(s_{jk}) - m(i/n)}{f(i/n)} \\
&\quad + n^{-1}\sum_{i,j=1}^n K_{d,h_d}(m(i/n)-t)\,\omega_j \sum_{k=1}^{n_j} K_{r,h_r}(s_{jk}-i/n) \\
&\qquad \times \frac{f'(i/n)(s_{jk}-i/n)\{m(s_{jk}) - m(i/n)\}}{f^2(i/n)},
\end{aligned}$$

$$\Delta_n^{(1.2)}(t) = -n^{-1} \sum_{i,j=1}^{n} K_{d,h_d}\left(m\left(i/n\right) - t\right) \omega_j \sum_{k=1}^{n_j} K_{r,h_r}\left(s_{jk} - i/n\right) \frac{\eta_j(s_{jk}) + \sigma(s_{jk})\varepsilon_{jk}}{f\left(i/n\right)}.$$

We have

$$\begin{aligned}
\Delta_n^{(1.1)}(t) \quad = \quad & -\sum_{j=1}^{n} \omega_j \sum_{k=1}^{n_j} \int_0^1 K_{d,h_d}\left(m(x) - t\right) K_{r,h_r}\left(s_{jk} - x\right) \frac{m(s_{jk}) - m(x)}{f(x)} dx\{1 + o_p(1)\} \\
& + \sum_{j=1}^{n} \omega_j \sum_{k=1}^{n_j} \int_0^1 K_{d,h_d}\left(m(x) - t\right) K_{r,h_r}\left(s_{jk} - x\right) \frac{f'(x)(s_{jk} - x)\{m(s_{jk}) - m(x)\}}{f^2(x)} dx \\
& \quad \times \{1 + o_p(1)\}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
& E\left\{\Delta_n^{(1.1)}(t)\right\} \\
= \quad & -\int_0^1 \int_0^1 K_{d,h_d}\left(m(x) - t\right) K_{r,h_r}\left(y - x\right) \frac{f(y)\{m(y) - m(x)\}}{f(x)} dydx\{1 + o(1)\} \\
& + \int_0^1 \int_0^1 K_{d,h_d}\left(m(x) - t\right) K_{r,h_r}\left(y - x\right) \frac{f(y)f'(x)(y - x)\{m(y) - m(x)\}}{f^2(x)} dydx\{1 + o(1)\} \\
= \quad & -h_r^2 \kappa_2(K_r) \left(\frac{\ddot{m} f + 2\dot{m}\dot{f}}{f\dot{m}}\right) (m^{-1}(t))\{1 + o(1)\} + h_r^2 \kappa_2(K_r) \left(\frac{2\dot{m}\dot{f}}{f\dot{m}}\right) (m^{-1}(t))\{1 + o(1)\} \\
= \quad & -h_r^2 \kappa_2(K_r) \left(\frac{\ddot{m}}{\dot{m}}\right) (m^{-1}(t))\{1 + o(1)\}. \qquad (1.2)
\end{aligned}$$

(1.2) is the dominant bias term. We next calculate the variance, i.e., $\mathrm{var}\{\Delta_n^{(1.2)}(t)\}$. Note that $E\{\Delta_n^{(1.2)}(t)\} = 0$. Define

$$\xi_{jk}(t) := \int_0^1 \frac{h_d h_r K_{d,h_d}\left(m(x) - t\right) K_{r,h_r}\left(s_{jk} - x\right)}{f(x)} dx, \quad \theta(t, s_{jk}) = \ \xi_{jk}(t)\{\eta_j(s_{jk}) + \sigma(s_{jk})\varepsilon_{jk}\},$$

for $k = 1, \ldots, n_j,\ j = 1, \ldots, n$. It follows that

$$\Delta_n^{(1.2)}(t) = -(h_d h_r)^{-1} \sum_{j=1}^{n} \omega_j \sum_{k=1}^{n_j} \theta(t, s_{jk})\{1 + o(1)\}.$$

We use a change of variables

$$u = \frac{x-z}{h_r},\ v = \frac{m(z)-t}{h_d},\ w = \frac{m(y)-t}{h_d}, \tag{1.3}$$

because of $h_r/h_d \to \infty$, we obtain,

$$E\left\{\theta^2(t, s_{jk})\right\} = h_d^2 h_r \frac{\gamma(m^{-1}(t), m^{-1}(t)) + \sigma^2(m^{-1}(t))}{\{m'(m^{-1}(t))\}^2 f(m^{-1}(t))} \int K_r^2(u)du.$$

For $k \neq l$ with $k, l = 1, \ldots, n_j$, we have

$$\begin{aligned}
& E\{\theta(t, s_{jk})\theta(t, s_{jl})\}/(h_d^2 h_r^2) \\
= & \int_0^1 \int_0^1 \left\{\int_0^1 \frac{K_{d,h_d}(m(u)-t)\, K_{r,h_r}(x-u)}{f(u)} du\right\} \left\{\int_0^1 \frac{K_{d,h_d}(m(v)-t)\, K_{r,h_r}(y-v)}{f(v)} dv\right\} \\
& \qquad \times \gamma(x,y) f(x) f(y) dx dy \\
= & \int_0^1 \left\{\int_0^1 \frac{K_{d,h_d}(m(u)-t)\, K_{r,h_r}(x-u)}{f(u)} du\right\} \left\{\int_0^1 K_{d,h_d}(m(v)-t)\, \gamma(x,v) dv\right\} f(x) dx \{1+o(1)\} \\
= & \int_0^1 \int_0^1 K_{d,h_d}(m(u)-t)\, K_{d,h_d}(m(v)-t)\, \gamma(u,v) du dv \{1+o(1)\} \\
= & \frac{\gamma(m^{-1}(t), m^{-1}(t))}{\{\dot{m}(m^{-1}(t))\}^2} \{1+o(1)\}.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\mathrm{var}\{\Delta_n^{(1.2)}(t)\} &= (h_d h_r)^{-2} \sum_{j=1}^n \mathrm{var}\Big\{\omega_j \sum_{k=1}^{n_j} \theta(t, s_{jk})\Big\}\{1+o(1)\} \\
&= (h_d h_r)^{-2} \sum_{j=1}^n \omega_j^2 E\Big\{\sum_{k=1}^{n_j} \theta(t, s_{jk})\Big\}^2 \{1+o(1)\} \\
&= \Gamma_0^A \sum_{j=1}^n \omega_j^2 n_j / h_r + \Gamma_0^B \sum_{j=1}^n \omega_j^2 n_j (n_j - 1)\{1+o(1)\}. \qquad (1.4)
\end{aligned}$$

Because of Condition (C3), all moments of $\{\eta_j(\cdot)+\sigma(\cdot)\varepsilon_{jk}\}$ up to order four are bounded. By similar changes of variables as (1.3), we have

$$E\{\theta(t,s_{jk_1})\theta(t,s_{jk_2})\theta(t,s_{jk_3})\theta(t,s_{jk_4})\} = O(h_d^4 h_r^4),$$

$$E\{\theta^2(t,s_{jk_1})\theta(t,s_{jk_2})\theta(t,s_{jk_3})\} = O(h_d^3 h_r^4),$$

$$E\{\theta^2(t,s_{jk_1})\theta^2(t,s_{jk_2})\} = O(h_d^2 h_r^4),$$

$$E\{\theta^3(t,s_{jk_1})\theta(t,s_{jk_2})\} = O(h_d^2 h_r^4),$$

$$E\{\theta^4(t,s_{jk_1})\} = O(h_d h_r^4),$$

which imply that

$$\sum_{j=1}^{n} E\left\{\omega_j (h_d h_r)^{-1}\sum_{k=1}^{n_j}\theta(t,s_{jk})\right\}^4 = O\left\{\sum_{j=1}^{n}\omega_j^4(n_j^4+n_j^3/h_d+n_j^2/h_d^2+n_j/h_d^3)\right\}.$$

Because we have

$$\sum_{j=1}^{n}\omega_j^4(n_j^4+n_j^3/h_d+n_j^2/h_d^2+n_j/h_d^3)\{\sum_{j=1}^{n}\omega_j^2 n_j/h_r+\sum_{j=1}^{n}\omega_j^2 n_j(n_j-1)\}^{-2}\to 0,$$

the Lyapunov condition is satisfied. Consequently, the asymptotic normality (1.1) follows from the Lyapunov central limit theorem. □

References

Dette, H., Neumeyer, N., & Pilz, K. F. (2006). A simple nonparametric estimator of a strictly monotone regression function. *Bernoulli*, *12(3)*, 469–490.

Li, Y., & Hsing, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Ann. Statist.*, *38*, 3321–3351.

Ziqi Chen

Department of Biostatistics,

The University of Texas MD Anderson Cancer Center, Houston, USA

School of Mathematics and Statistics, Central South University, Changsha, China

E-mail: (chenzq453@gmail.com)

Qibing Gao

School of Mathematical Sciences, Nanjing Normal University, Nanjing, China

E-mail: (gaoqibing@njnu.edu.cn)

Bo Fu

School of Data Science, Fudan University, Shanghai, China

E-mail: (bofu@fudan.edu.cn)

Hongtu Zhu

Department of Biostatistics,

The University of Texas MD Anderson Cancer Center, Houston, USA

The University of North Carolina at Chapel Hill, Chapel Hill, USA

E-mail: (htzhu@email.unc.edu)