Supplementary Material for “Multi-round smoothed composite quantile regression for distributed data”

Fengrui Di\textsuperscript{1}, Lei Wang\textsuperscript{1}\textsuperscript{*}

\textsuperscript{1}School of Statistics and Data Science & LPMC, Nankai University
No. 94 Weijin Road, Tianjin 300071, China.

Proof of Proposition 1

Let

\[ C = P_N - P = \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N} x_i \left\{ H\left( \frac{y_i - \tilde{b}_{r_k} - x_i^T \tilde{\beta}}{h} \right) - I\{y_i - b_{r_k} - x_i^T \beta \geq 0\} \right\} + \frac{y_i - \tilde{b}_{r_k} - x_i^T \beta}{h} H'\left( \frac{y_i - \tilde{b}_{r_k} - x_i^T \tilde{\beta}}{h} \right). \]

Note that \( \|C\|_2 = \sup_{v \in \mathbb{R}^p} |v^T C| \). Let \( S_{p-1}^{1/2} \) be a 1/2 net of the unit sphere \( S_{p-1} \) in the Euclidean distance in \( \mathbb{R}^p \). According to the proof of Lemma 3 in Cai et al. (2010),

\[ d_p := \text{Card}(S_{p-1}^{1/2}) \leq 5^p. \]

Let \( v_1, \ldots, v_{d_p} \) be the centers of the \( d_p \) elements in the net. Therefore, for any \( v \) in \( S_{p-1} \), it can be shown that \( \|v - v_j\|_2 \leq 1/2 \) for some \( j \) and \( \|C\|_2 \leq \sup_{j \leq d_p} |v_j^T C| + \|C\|_2/2. \)

Above all lead to

\[ \|C\|_2 \leq 2 \sup_{j \leq d_p} |v_j^T C|. \]

For \( \alpha \in \mathbb{R}^p \), denote \( C_j(\alpha) = \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N} v_j^T x_i H(\alpha) \) and

\[ H(\alpha) = H\left( \frac{y_i - \tilde{b}_{r_k} - x_i^T \alpha}{h} \right) - I\{y_i - b_{r_k} - x_i^T \beta \geq 0\} + \frac{y_i - \tilde{b}_{r_k} - x_i^T \beta}{h} H'\left( \frac{y_i - \tilde{b}_{r_k} - x_i^T \alpha}{h} \right). \]

\textsuperscript{*}Address for correspondence: Lei Wang, School of Statistics and Data Science & LPMC, Nankai University. E-mail: lwangstat@nankai.edu.cn.
When \( \| \tilde{\beta} - \beta \|_2 \leq a_n \), it leads to

\[
\| \mathcal{C} \|_2 \leq 2 \sup_{j \leq d_p} \sup_{\alpha - \beta \leq a_n} |\mathcal{C}_j(\alpha)|. \tag{1}
\]

For every \( i \), we divide the interval \( [\beta_i - a_n, \beta_i + a_n] \) into \( n^M \) small subintervals and each has length \( 2a_n/n^M \), where \( M \) is a large positive number. Therefore, there exist a set of points \( \{ \alpha_l, 1 \leq l \leq n^{Mp} \} \) in \( \mathbb{R}^p \) such that for any \( \alpha \) in the ball \( \| \alpha - \beta \|_2 \leq a_n \),

\[
\| \alpha - \alpha_l \|_2 \leq 2\sqrt{p}a_n/n^M,
\]

for some \( 1 \leq l \leq n^{Mp} \). Let \( \Delta(\alpha) = \alpha - \beta \) and it can be proved that

\[
\begin{align*}
\frac{y_i - \tilde{b}_{r_k} - x_i^T \beta}{h} & H'(\frac{y_i - \tilde{b}_{r_k} - x_i^T \alpha}{h}) - \frac{y_i - \tilde{b}_{r_k} - x_i^T \beta}{h} H'(\frac{y_i - \tilde{b}_{r_k} - x_i^T \alpha_l}{h}) \\
& \leq C h^{-1} \| x_i \|_2 \| \alpha - \alpha_l \|_2 + C h^{-2} \| x_i \|_2^2 \| \Delta(\alpha) \|_2 \| \alpha - \alpha_l \|_2 \\
& \leq C h^{-1} \sqrt{p}a_n n^{-M} \| x_i \|_2 + C h^{-2} \sqrt{p}a_n^2 n^{-M} \| x_i \|_2^2.
\end{align*}
\]

Similarly, we can prove that

\[
\left| H\left(\frac{y_i - \tilde{b}_{r_k} - x_i^T \alpha}{h}\right) - H\left(\frac{y_i - \tilde{b}_{r_k} - x_i^T \alpha_l}{h}\right) \right| \leq C h^{-1} \sqrt{p}a_n n^{-M} \| x_i \|_2.
\]

Therefore, it leads to

\[
\sup_j \sup_{\| \alpha - \beta \|_2 \leq a_n} |\mathcal{C}_j(\alpha)| - \sup_l |\mathcal{C}_j(\alpha_l)| \leq C \sqrt{p}a_n K \sum_{i=1}^{N} \| x_i \|_2 + C \sqrt{p}a_n^2 K \sum_{i=1}^{N} \| x_i \|_2^3.
\]

Since \( \max_{i,j} E|x_{ij}|^3 < \infty \), by letting \( M \) large enough, we have

\[
\sup_j \sup_{\| \alpha - \beta \|_2 \leq a_n} |\mathcal{C}_j(\alpha)| - \sup_l |\mathcal{C}_j(\alpha_l)| = O_p(n^{-\gamma}), \tag{3}
\]

for any \( \gamma > 0 \). It is enough to show that \( \sup_j \sup_l |\mathcal{C}_j(\alpha_l)| \) satisfies the bound in the Proposition 1.
We first prove the proposition under (C3)*. Let \( \hat{x}_{ij} = x_{ij}I\{ |x_{ij}| \leq N^k \} \), \( \hat{x}_i = (\hat{x}_{i1}, \ldots, \hat{x}_{ip})^T \), \( \epsilon_{ik} = y_i - b_{rk} - x_i^T \beta \), \( \Delta(b_k) = \hat{b}_{rk} - b_{rk} \) and \( |\hat{b}_{rk} - b_{rk}| \leq b_n \) for \( i = 1, \ldots, N \) and \( k = 1, \ldots, K \). Denote \( \tilde{C}_j(\alpha) = \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N} v_j^T \hat{x}_i \tilde{H}(\alpha) \) and

\[
\tilde{H}(\alpha) = H \left( \frac{\epsilon_{ik} - \Delta(b_k) - \hat{x}_i^T \Delta(\alpha)}{h} \right) - I\{ \epsilon_{ik} \geq 0 \} + \frac{\epsilon_{ik} - \Delta(b_k)}{h} H' \left( \frac{\epsilon_{ik} - \Delta(b_k) - \hat{x}_i^T \Delta(\alpha)}{h} \right).
\]

Using Chebyshev’s inequality, we have

\[
P \left( \sup_j \sup_l |C_j(\alpha_l)| \neq \sup_j \sup_l |\tilde{C}_j(\alpha_l)| \right) \leq \sum_{i=1}^{N} \sum_{j=1}^{p} P(|x_{ij}| \geq N^k) = O(p/N) = o(1).
\]

Denote \( f_1(\cdot|\mathbf{x}) \), \( \ldots, f_K(\cdot|\mathbf{x}) \) as the corresponding conditional density functions of \( \epsilon_1, \ldots, \epsilon_K \) given \( \mathbf{x} \), which are local-scale functions for \( f(\cdot|\mathbf{x}) \) in Condition (C1). Since the conditional density functions \( f_1(\cdot|\mathbf{x}), \ldots, f_K(\cdot|\mathbf{x}) \) are Lipschitz continuous, they are also bounded. By Condition (C2) that \( H(\mathbf{x}) \) is bounded,

\[
E \left[ H \left( \frac{\epsilon_{ik} - \Delta(b_k) - \hat{x}_i^T \Delta(\alpha)}{h} \right) - I\{ \epsilon_{ik} \geq 0 \} \right]^2 \leq C h \left( \frac{|\hat{x}_i^T \Delta(\alpha)/h + \Delta(b_k)/h|}{|\hat{x}_i^T \Delta(\alpha)/h|^2 + 1} \right).
\]

Based on (4), (5) and the definition of \( \tilde{H}(\alpha) \), it leads to

\[
E \left( v_j^T \hat{x}_i \tilde{H}(\alpha) \right)^2 \leq 2 E \left( v_j^T \hat{x}_i \right)^2 \left[ H \left( \frac{\epsilon_{ik} - \hat{x}_i^T \Delta(\alpha) - \Delta(b_k)}{h} \right) - I\{ \epsilon_{ik} \geq 0 \} \right]^2 + 2 E \left( v_j^T \hat{x}_i \right)^2 \left[ \frac{\epsilon_{ik} - \Delta(b_k)}{h} H' \left( \frac{\epsilon_{ik} - \Delta(b_k) - \hat{x}_i^T \Delta(\alpha)}{h} \right) \right]^2 \leq C h \left( \sup_{\| \theta \|_2 = 1} E \theta^T \hat{x}_i \right)^2 + \sup_{\| \theta \|_2 = 1} E \| \theta^T \hat{x}_i \|^3 \| \alpha - \beta \|_2 / h \left\| \theta \right\|_2 + \sup_{\| \theta \|_2 = 1} E \| \theta^T \hat{x}_i \|^4 \| \alpha - \beta \|_2^2 / h^2 \leq C h \left( 1 + \| \alpha - \beta \|_2 / h + \| \Delta(b_k) \| / h + \| \alpha - \beta \|_2^2 / h^2 \right),
\]
where we use the inequalities that

\[
\sup_{\|\theta\|_2=1} E(\theta^T \hat{x}_i)^4 \leq 8 \sup_{\|\theta\|_2=1} E(\theta^T x_i)^4 + 8 \sup_{\|\theta\|_2=1} (E \sum_{j=1}^p |\theta_j x_{ij}| I\{\|x_{ij}\| \geq N^\kappa\})^4
\]

\[
\leq 8 \sup_{\|\theta\|_2=1} E(\theta^T x_i)^4 + 8p \sum_{j=1}^p E(x_{ij})^4 I\{\|x_{ij}\| \geq N^\kappa\} \leq C,
\]

\[
\sup_{\|v\|=1,\|u\|=1} E[(v^T \hat{x}_i)^2|u^T \hat{x}_i|] \leq \sup_{\|\theta\|=1} E|\theta^T \hat{x}_i|^3 \leq C,
\]

\[
\sup_{\|v\|=1,\|u\|=1} E[(v^T \hat{x}_i)^2(u^T \hat{x}_i)^2] \leq \sup_{\|\theta\|=1} E|\theta^T \hat{x}_i|^4 \leq C.
\]

According to (2) and noting that \(H(x), H'(x)\) and \(xH'(x)\) are bounded, we can obtain that

\[
|v_j^T \hat{x}_i \hat{H}(\alpha)| \leq C\|\hat{x}_i\|_2(1 + \|\hat{x}_i\|_2\|\alpha - \beta\|_2/h) \leq C\sqrt{pN^\kappa} + CpN^{2\kappa}\|\alpha - \beta\|_2/h.
\]

Under Condition (C3)*,

\[
p \log(KN) = o\left(\frac{\sqrt{Nph \log(kN)}}{pN^{2\kappa}}\right).
\]

Using Benstein’s inequality, for any \(\gamma > 0\), there exists a constant \(C\) such that

\[
\sup_j\sup_l P\left(\left|\hat{C}_j(\alpha_l) - E\hat{C}_j(\alpha_l)\right| \geq C\sqrt{\frac{p \log(KN)}{N}}\right) = O\left((kN)^{-\gamma p}\right).
\]

This leads to

\[
\sup_j\sup_l |C_j(\alpha_l) - EC_j(\alpha_l)| = Op\left(\sqrt{\frac{p \log(KN)}{N}}\right).
\]

(6)

It remains to give a bound for \(EC_j(\alpha)\). Let \(F_k(\cdot|x)\) be the conditional distribution of \(\varepsilon_k\) given \(x\). Denote \(t = hx + \hat{x}_i^T \Delta(\alpha) + \Delta(b_k)\) and we have

\[
EH\left(\frac{\varepsilon_{ik} - \Delta(b_k) - \hat{x}_i^T \Delta(\alpha)}{h}\Big| x_i\right) = \int_{-\infty}^{\infty} H\left(\frac{t - \Delta(b_k) - \hat{x}_i^T \Delta(\alpha)}{h}\right) f_k(t|x_i) dt
\]

\[
= \int_{-\infty}^{\infty} H(x) f_k(t|x_i) dt = -\int_{-\infty}^{\infty} H(x) d(1 - F_k(t|x_i))
\]

\[
= \int_{-\infty}^{\infty} (1 - F_k(t|x_i)) H'(x) dx = 1 - \tau_k - f_k(0|x_i) \int_{-\infty}^{\infty} (t) H'(x) dx + O(1) \int_{-\infty}^{\infty} (t)^2 |H'(x)| dx
\]

\[
= 1 - \tau_k - f_k(0|x_i) \left(\hat{x}_i^T \Delta(\alpha) + \Delta(b_k) + h \int_{-\infty}^{\infty} xH'(x) dx\right) + O\left(h^2 + \left(\hat{x}_i^T \Delta(\alpha) + \Delta(b_k)\right)^2\right).
\]
Similarly,

\[
E\left[\frac{\epsilon_{ik} - \Delta(b_k)}{h} H'(x) \left(\frac{\epsilon_{ik} - \Delta(b_k) - \hat{x}_i^T \Delta(\alpha)}{h}\right) x_i\right] = \int_{-\infty}^{\infty} (t - \Delta(b_k)/h) f_k(t|x_i) dx
\]

\[
= f_k(0|x_i) \left(\hat{x}_i^T \Delta(\alpha) + h \int_{-\infty}^{\infty} x H'(x) dx \right) + O(h) \int_{-\infty}^{\infty} (h x + \hat{x}_i^T \Delta(\alpha) + \Delta(b_k))^2 |H'(x)| dx
\]

These two equalities imply that uniformly in \(\alpha\) and \(j\),

\[
|EC_j(\alpha)| \leq C(K_1 h^2 + K_2 \|\Delta(\alpha)\|_2^2 + K_3 \Delta(b_1)^2).
\]

Hence

\[
\sup_i \sup_t |EC_j(\alpha_t)| = O_p(h^2 + a_n^2 + b_n^2). \tag{7}
\]

Combining that with (1), (3) and (6), this completes the proof of the proposition under Condition (C3)*.

We now prove the proposition under condition (C3). To bound \(C_j(\alpha_t) - EC_j(\alpha_t)\) under condition (C3), we introduce the following exponential inequality from Cai and Liu (2011).

**LEMMA 1** Let \(S_1, \ldots, S_N\) be independent random variables with mean zero. Suppose that there exist some \(t > 0\) and \(B_N\) such that \(\sum_{k=1}^{N} ES_k^2 e^{|S_k|} \leq B_N^2\). Then for \(0 < x < B_N\)

\[
P\left(\sum_{k=1}^{N} S_k > C_t B_N x \right) \leq \exp(-x^2),
\]

where \(C_t = t + t^{-1}\).

According to \(a_n = O(h)\) and (2), denote \(|\sum_{k=1}^{K} v_j^T x_i H(\alpha)| \leq C |v_j^T x_i| (1 + |(\alpha - \beta)^T x_i|)/\|\alpha - \beta\|_2 := S_{ij}\), which implies that

\[
E(|\sum_{k=1}^{K} v_j^T x_i H(\alpha)|^2 e^{n |\sum_{k=1}^{K} v_j^T x_i H(\alpha)|}) \leq E(|\sum_{k=1}^{K} v_j^T x_i H(\alpha)|^2 e^{n S_{ij}}) \leq Ch.
\]

Let \(\tilde{B}_N = C \sqrt{Nh}\) and \(x = \sqrt{p \log(KN)}\). By Lemma 1 and the fact that \(\sqrt{p \log(KN)} = o(\sqrt{Nh})\), for sufficiently large \(C\), it can be shown that

\[
\sup_j \sup_t P\left(|C_j(\alpha_t) - EC_j(\alpha_t)| \geq C \sqrt{\frac{p \log(KN)}{N}}\right) = O\left(\left(\frac{NK}{N}\right)^{-\gamma p}\right).
\]

Combining that with (1), (3) and (7), we complete the proof of the proposition under condition (C3).
Proof of Proposition 2

Based on the proof of Lemma 3 in Cai et al. (2010), we have \(|Q_N - Q| \leq 5 \sup_{j \leq b_p} |v_j^T (Q_N - Q) v_j|\), where \(\{v_i, 1 \leq i \leq b_p\}\) are some non-random vectors with \(\|v_i\|_2 = 1\) and \(b_p \leq 5^p\).

Now let

\[
Q_j(\alpha) = \frac{1}{Nh} \sum_{k=1}^{K} \sum_{i=1}^{N} (v_j^T x_i x_i^T v_j) H'\left(\frac{y_i - \hat{b}_{rk} - x_i^T \alpha}{h}\right) = \frac{1}{Nh} \sum_{k=1}^{K} \sum_{i=1}^{N} (v_j^T x_i)^2 H'\left(\frac{y_i - \hat{b}_{rk} - x_i^T \alpha}{h}\right).
\]

Therefore, when \(\|\hat{\beta} - \beta\|_2 \leq a_n\),

\[
\|Q_N - Q\| \leq 5 \sup_{j \leq b_p} \sup_{|\alpha - \beta| \leq a_n} |Q_j(\alpha) - v_j^T Q v_j|.
\]

(8)

Because \(H'(x)\) is Lipschitz continuous, we get

\[
\left| \frac{1}{h} H'\left(\frac{y_i - \hat{b}_{rk} - x_i^T \alpha}{h}\right) - \frac{1}{h} H'\left(\frac{y_i - \hat{b}_{rk} - x_i^T \alpha_l}{h}\right) \right| \leq C h^{-2} |x_i^T (\alpha - \alpha_l)|.
\]

Therefore,

\[
\sup_{j} \sup_{\|\alpha - \beta\|_2 \leq a_n} |Q_j(\alpha) - v_j^T Q v_j| - \sup_{j} \sup_{l} |Q_j(\alpha_l) - v_j^T Q v_j| \leq \frac{C \sqrt{p} K a_n}{nM Nh^2} \sum_{i=1}^{N} \|x_i\|_2^3.
\]

Since \(\max_{ij} E |x_{ij}|^3 < \infty\), by letting \(M\) large enough, we have for any \(\gamma > 0\),

\[
\sup_{j} \sup_{\|\alpha - \beta\|_2 \leq a_n} |Q_j(\alpha) - v_j^T Q v_j| - \sup_{j} \sup_{l} |Q_j(\alpha_l) - v_j^T Q v_j| = O_P(n^{-\gamma}).
\]

(9)

We now prove the proposition under under (C3)\(^*\). Similarly to the proof of Proposition 1, denote \(x_{ij} = x_{ij} \mathbb{1}\{|x_{ij}| \leq N^\kappa\}\) and

\[
\bar{Q}_j(\alpha) = \frac{1}{Nh} \sum_{k=1}^{K} \sum_{i=1}^{N} (v_j^T \hat{x}_i)^2 H'\left(\frac{\epsilon_{ik} - \Delta(b_k) - \hat{x}_i^T \Delta(\alpha)}{h}\right).
\]

We can prove that

\[
P\left(\sup_{j} \sup_{l} |Q_j(\alpha_l) - v_j^T Q v_j| \neq \sup_{j} \sup_{l} |\bar{Q}_j(\alpha_l) - v_j^T Q v_j|\right)
\leq \sum_{i=1}^{N} \sum_{j=1}^{p} P(|x_{ij}| \geq N^\kappa) = o(1).
\]

(10)

As the proof of (5), we have

\[
E\left[H'\left(\frac{\epsilon_{ik} - \Delta(b_k) - \hat{x}_i^T \Delta(\alpha)}{h} \right) | x_i \right]^2 = O(h).
\]

6
Using Bernstein’s inequality, for any \( \gamma > 0 \), there exists a constant \( C \) such that

\[
\sup_j \sup_l P \left( |\hat{Q}_j(\alpha_l) - E(\hat{Q}_j(\alpha_l))| \geq C \sqrt{\frac{p \log(KN)}{Nh}} \right) = O((KN)^{-\gamma p}).
\]

(11)

Moreover,

\[
E \left[ \frac{1}{h} H' \left( \frac{\epsilon_{ik} - \Delta(b_k) - \hat{\alpha}_l^T \Delta(\alpha)}{h} | \mathbf{x}_i \right) \right] = \int_{-\infty}^{\infty} H'(x) f_k(hx + \hat{\alpha}_l^T \Delta(\alpha) + \Delta(b_k)| \mathbf{x}_i) dx
\]

\[
= f_k(0| \mathbf{x}_i) + O(h + |\hat{\alpha}_l^T \Delta(\alpha) + \Delta(b_k)|)
\]

Therefore,

\[
|E(\hat{Q}_j(\alpha)) - v_j^T \mathbf{Q} v_j| \leq CK \max_i E|x_{ij}| + CK(h + \|\alpha - \beta\|_2 + \Delta(b_1))
\]

\[
\leq CK \max_i \sum_{j=1}^p E x_{ij}^2 I\{|x_{ij}| \geq N^\kappa\}^{1/2} + CK \max_i \sum_{j=1}^p E x_{ij}^2 I\{|x_{ij}| \geq N^\kappa\} + CK(h + \|\alpha - \beta\|_2 + \Delta(b_1)) \leq C_1(h + a_n + b_n).
\]

(12)

Combining (12) with (8), (9), (10) and (11), we can get the desired inequality under (C3)*. We now prove the proposition under condition (C3). Let \( S_{NK} = \sum_{k=1}^K (v_j^T \mathbf{x}_i)^2 H' \left( \frac{\epsilon_{ik} - \mathbf{x}_i^T \Delta(\alpha) - \Delta(b_k)}{h} \right) \) and we have

\[
E(S_{ij})^2 e^{\eta |S_{ij}|} \leq E(S_{ij})^2 e^{C \eta (v_j^T \mathbf{x}_i)^2} = O(h).
\]

Let \( B_N = C \sqrt{Nh} \) and \( x = \sqrt{\gamma p \log(KN)} \). By Lemma 1 again, it leads to

\[
\sup_j \sup_l P \left( \frac{1}{N} S_{NK} \right) \geq C \sqrt{\frac{p \log(KN)}{Nh}} = O((KN)^{-\gamma p}).
\]

(13)

Combining (8), (9), (12) and (13), the desired inequality under (C3) is obtained.

**Proof of Theorems 1-3**

For independent random vectors \( \{\mathbf{x}_i, 1 \leq i \leq N\} \) with \( \sup_{ij} E|x_{ij}|^3 = O(1) \), let

\[
S_{NK} = \sum_{k=1}^K \sum_{i=1}^N \mathbf{x}_i(I\{\epsilon_{ik} \geq 0\} + \tau_k - 1).
\]

Since \( E\|S_{NK}\|^2 = O(Np) \), it leads to

\[
\left\| \frac{1}{N} S_{NK} \right\|_2 = O_P\left( \sqrt{\frac{p}{N}} \right).
\]

(14)
By (14) and Propositions 1-2, it is easy to show that the result holds. For \( q = 1 \), let 
\( a_n = \sqrt{p/n}, \ b_n = \sqrt{1/n} \), since we assume that \( \| \tilde{\beta} - \beta \|_2 = O_P(\sqrt{p/n}) \) and \( |\tilde{b}_{\tau_k} - b_{\tau_k}| = O_P(\sqrt{1/n}) \) for \( k = 1, \ldots, K \). Suppose the theorem holds for \( q = g - 1 \) with some \( g \geq 2 \). Noting that \( p = O(\sqrt{n/(\log(KN))^2}) \), we have 
\[
\sqrt{ph^{(g-1)}(\log(KN))/N} = O(\sqrt{p/N}).
\]

Then for \( q = g \) with initial estimator \( (\tilde{b}_{\tau_1}, \ldots, \tilde{b}_{\tau_K}, \tilde{\beta}) = (\hat{b}^{(g-1)}_{\tau_1}, \ldots, \hat{b}^{(g-1)}_{\tau_K}, \hat{\beta}^{(g-1)}) \), it can be seen that \( a_n = O(h^{(g)}) \) and \( b_n = O(h^{(g)}) \). Hence, we have proved Theorem 1. Theorem 2 follows directly from Theorem 1 by the Lindeberg-Feller central limit theorem. In addition, Theorem 2 is a special situation of Theorem 3, which can be confirmed similarly as Theorem 2.

**References**
