

Supplementary Materials for “Distributed quantile regression for big data with missing covariates”

Ye Fan¹, Hua Zhou^{2,3}, Jin Zhou^{3,4,5}, Tiejun Tong⁶, Nan Lin^{7*}

¹School of Statistics and Data Science, Capital University of Economics and Business, Beijing, 100070, China.

²Department of Computational Medicine, University of California, Los Angeles, CA 90095, USA.

³Department of Biostatistics, University of California, Los Angeles, CA 90095, USA.

⁴Department of Medicine, University of California, Los Angeles, CA 90095, USA.

⁵Department of Epidemiology and Biostatistics, University of Arizona, Tucson, AZ 85721, USA.

⁶Department of Mathematics, Hong Kong Baptist University, Hong Kong, 999077, China.

^{7*}Department of Statistics and Data Science, Washington University in St. Louis, St. Louis, MO 63130, USA.

*Corresponding author(s). E-mail(s): nlin@wustl.edu;

In this file, we give the technical proofs for Theorems 1 and 2.

1 Proof of Theorem 1

Denote $\mathbf{e}_{m1}^k = \boldsymbol{\beta}_m^k - \boldsymbol{\beta}^k$, $\mathbf{e}_{m2}^k = \mathbf{y}_m^* - X_m^* \boldsymbol{\beta}_m^k - \mathbf{r}_m^k$ ($m = 1, 2, \dots, M$) and

$$D^k = \sum_{m=1}^M \left[\eta^{-1} \|\mathbf{u}_m^k - \mathbf{u}_m^*\|_2^2 + \eta^{-1} \|\mathbf{v}_m^k - \mathbf{v}_m^*\|_2^2 + \eta \|\boldsymbol{\beta}_m^k - \boldsymbol{\beta}_m^*\|_2^2 + \eta \|X_m^* (\boldsymbol{\beta}_m^k - \boldsymbol{\beta}_m^*)\|_2^2 \right].$$

Next, we prove that the sequence $\{D^k, k = 1, 2, \dots\}$ is decreasing.

Firstly, according to the assumption imposed in Theorem 1, $(\boldsymbol{\beta}^*, \mathbf{r}_m^*, \boldsymbol{\beta}_m^*, \mathbf{u}_m^*, \mathbf{v}_m^*)$ is a saddle point of $\mathcal{L}_0(\boldsymbol{\beta}, \mathbf{r}_m, \boldsymbol{\beta}_m, \mathbf{u}_m, \mathbf{v}_m)$, and the following inequality holds at any iteration k ,

$$\mathcal{L}_0(\boldsymbol{\beta}^*, \mathbf{r}_m^*, \boldsymbol{\beta}_m^*, \mathbf{u}_m^*, \mathbf{v}_m^*) \leq \mathcal{L}_0(\boldsymbol{\beta}^{k+1}, \mathbf{r}_m^{k+1}, \boldsymbol{\beta}_m^{k+1}, \mathbf{u}_m^*, \mathbf{v}_m^*).$$

By using the facts that $\boldsymbol{\beta}_m^* = \boldsymbol{\beta}^*$ and $\mathbf{r}_m^* = \mathbf{y}_m^* - X_m^* \boldsymbol{\beta}_m^*$, this inequality can be rewritten as

$$\sum_{m=1}^M \rho_\tau(\mathbf{r}_m^*) \leq \sum_{m=1}^M \rho_\tau(\mathbf{r}_m^{k+1}) + \sum_{m=1}^M \mathbf{u}_m^{*T} \mathbf{e}_{m1}^{k+1} + \sum_{m=1}^M \mathbf{v}_m^{*T} \mathbf{e}_{m2}^{k+1}. \quad (1)$$

Secondly, from the optimality of $\boldsymbol{\beta}^{k+1}$ and $\boldsymbol{\beta}_m^{k+1}$ ($m = 1, 2, \dots, M$) (see the subproblems (11) and (13) in the main article), we can get

$$\sum_{m=1}^M [\mathbf{u}_m^k + \eta(\boldsymbol{\beta}_m^k - \boldsymbol{\beta}^{k+1})] = \mathbf{0}_p, \quad (2)$$

and

$$\mathbf{u}_m^k + \eta(\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}^{k+1}) - X_m^{*T} \mathbf{v}_m^k - \eta X_m^{*T} (\mathbf{y}_m^* - X_m^* \boldsymbol{\beta}_m^{k+1} - \mathbf{r}_m^{k+1}) = \mathbf{0}_p. \quad (3)$$

By substituting the equations (14)-(15) in the main article into (2)-(3), we obtain

$$\sum_{m=1}^M [\mathbf{u}_m^{k+1} - \eta (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)] = \mathbf{0}_p \quad \text{and} \quad \mathbf{u}_m^{k+1} - \mathbf{X}_m^{wT} \mathbf{v}_m^{k+1} = \mathbf{0}_p. \quad (4)$$

Besides, by the convexity of the check loss function $\rho_\tau(\cdot)$ and the optimality of $\mathbf{r}_m^{k+1}$, we have

$$\mathbf{0}_{n_m^*} \in \partial \rho_\tau(\mathbf{r}_m^{k+1}) - \mathbf{v}_m^k - \eta (\mathbf{y}_m^* - \mathbf{X}_m^* \boldsymbol{\beta}_m^k - \mathbf{r}_m^{k+1}). \quad (5)$$

From (5), we can see that optimizing the subproblem (12) in the main article is equivalent to solving

$$\mathbf{r}_m^{k+1} = \arg \min_{\mathbf{r}_m} \left\{ \rho_\tau(\mathbf{r}_m) - [\mathbf{v}_m^{k+1} + \eta \mathbf{X}_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T \mathbf{r}_m \right\}.$$

Then we have

$$\begin{aligned} & \sum_{m=1}^M \rho_\tau(\mathbf{r}_m^{k+1}) - \sum_{m=1}^M [\mathbf{v}_m^{k+1} + \eta \mathbf{X}_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T \mathbf{r}_m^{k+1} \\ & \leq \sum_{m=1}^M \rho_\tau(\mathbf{r}_m^*) - \sum_{m=1}^M [\mathbf{v}_m^{k+1} + \eta \mathbf{X}_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T \mathbf{r}_m^*. \end{aligned} \quad (6)$$

Multiplying the two equations in (4) by $(\boldsymbol{\beta}^{k+1} - \boldsymbol{\beta}^*)$ and $(\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)$, respectively, we can get

$$\sum_{m=1}^M [\mathbf{u}_m^{k+1} - \eta (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T (\boldsymbol{\beta}^{k+1} - \boldsymbol{\beta}^*) = 0, \quad (7)$$

and

$$\sum_{m=1}^M (\mathbf{u}_m^{k+1} - \mathbf{X}_m^{*T} \mathbf{v}_m^{k+1})^T (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*) = 0. \quad (8)$$

Through adding (7) and (8) to the inequality (6), we obtain

$$\begin{aligned}
\sum_{m=1}^M \rho_{\tau}(\mathbf{r}_m^{k+1}) &\leq \sum_{m=1}^M \rho_{\tau}(\mathbf{r}_m^*) - \sum_{m=1}^M [\mathbf{v}_m^{k+1} + \eta X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T \mathbf{r}_m^* \\
&\quad + \sum_{m=1}^M [\mathbf{v}_m^{k+1} + \eta X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T \mathbf{r}_m^{k+1} \\
&\quad + \sum_{m=1}^M [\mathbf{u}_m^{k+1} - \eta (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}^*) \\
&\quad - \sum_{m=1}^M (\mathbf{u}_m^{k+1} - X_m^{*T} \mathbf{v}_m^{k+1})^T (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}^*).
\end{aligned} \tag{9}$$

Rearranging (9), we have

$$\begin{aligned}
\sum_{m=1}^M \rho_{\tau}(\mathbf{r}_m^{k+1}) &\leq \sum_{m=1}^M \rho_{\tau}(\mathbf{r}_m^*) - \sum_{m=1}^M (\mathbf{u}_m^{k+1})^T [(\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}^{k+1}) - (\boldsymbol{\beta}_m^* - \boldsymbol{\beta}^*)] \\
&\quad - \sum_{m=1}^M (\mathbf{v}_m^{k+1})^T [(X_m^* \boldsymbol{\beta}_m^* + \mathbf{r}_m^*) - (X_m^* \boldsymbol{\beta}_m^{k+1} + \mathbf{r}_m^{k+1})] \\
&\quad + \sum_{m=1}^M \eta [X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T (\mathbf{r}_m^{k+1} - \mathbf{r}_m^*) \\
&\quad - \sum_{m=1}^M \eta (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)^T (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}^*).
\end{aligned} \tag{10}$$

By applying the facts that $\boldsymbol{\beta}_m^* = \boldsymbol{\beta}^*$ and $\mathbf{r}_m^* = \mathbf{y}_m^* - X_m^* \boldsymbol{\beta}_m^*$, and the definitions of $\mathbf{e}_{m1}^{k+1}$ and $\mathbf{e}_{m2}^{k+1}$, we can rewrite (10) as

$$\begin{aligned}
\sum_{m=1}^M \rho_{\tau}(\mathbf{r}_m^{k+1}) &\leq \sum_{m=1}^M \rho_{\tau}(\mathbf{r}_m^*) - \sum_{m=1}^M (\mathbf{u}_m^{k+1})^T \mathbf{e}_{m1}^{k+1} - \sum_{m=1}^M (\mathbf{v}_m^{k+1})^T \mathbf{e}_{m2}^{k+1} \\
&\quad + \sum_{m=1}^M \eta [X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T (\mathbf{r}_m^{k+1} - \mathbf{r}_m^*) \\
&\quad - \sum_{m=1}^M \eta (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)^T (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}^*).
\end{aligned} \tag{11}$$

Besides, noting that $(\mathbf{r}_m^{k+1} - \mathbf{r}_m^*)$ and $(\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)$ are respectively equal to $[-\mathbf{e}_{m2}^{k+1} - X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)]$ and $[-\mathbf{e}_{m1}^{k+1} + (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)]$, the inequality (11) can be further rewritten as

$$\begin{aligned} \sum_{m=1}^M \rho_\tau(\mathbf{r}_m^{k+1}) &\leq \sum_{m=1}^M \rho_\tau(\mathbf{r}_m^*) - \sum_{m=1}^M (\mathbf{u}_m^{k+1})^\top \mathbf{e}_{m1}^{k+1} - \sum_{m=1}^M (\mathbf{v}_m^{k+1})^\top \mathbf{e}_{m2}^{k+1} \\ &\quad + \sum_{m=1}^M \eta [X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)]^\top [-\mathbf{e}_{m2}^{k+1} - X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)] \\ &\quad + \sum_{m=1}^M \eta (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)^\top [\mathbf{e}_{m1}^{k+1} - (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)]. \end{aligned} \quad (12)$$

Combining (1) and (12), we get

$$\begin{aligned} &\sum_{m=1}^M (\mathbf{u}_m^{k+1} - \mathbf{u}_m^*)^\top \mathbf{e}_{m1}^{k+1} + \sum_{m=1}^M (\mathbf{v}_m^{k+1} - \mathbf{v}_m^*)^\top \mathbf{e}_{m2}^{k+1} \\ &\leq \sum_{m=1}^M \eta [X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^\top [-\mathbf{e}_{m2}^{k+1} - X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)] \\ &\quad + \sum_{m=1}^M \eta (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)^\top [\mathbf{e}_{m1}^{k+1} - (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)]. \end{aligned} \quad (13)$$

On the other hand, by the definition of D^k , we have

$$\begin{aligned} D^{k+1} - D^k &= -\eta \sum_{m=1}^M \|\mathbf{e}_{m1}^{k+1}\|_2^2 + 2 \sum_{m=1}^M (\mathbf{u}_m^{k+1} - \mathbf{u}_m^*)^\top \mathbf{e}_{m1}^{k+1} \\ &\quad - \eta \sum_{m=1}^M \|\mathbf{e}_{m2}^{k+1}\|_2^2 + 2 \sum_{m=1}^M (\mathbf{v}_m^{k+1} - \mathbf{v}_m^*)^\top \mathbf{e}_{m2}^{k+1} \\ &\quad - \eta \sum_{m=1}^M \|\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k\|_2^2 + 2 \sum_{m=1}^M (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)^\top (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*) \\ &\quad - \eta \sum_{m=1}^M \|X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)\|_2^2 + 2 \sum_{m=1}^M [X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^\top [X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^*)]. \end{aligned}$$

By substituting (13) into the above equation, we get

$$\begin{aligned}
D^{k+1} - D^k &\leq -\eta \sum_{m=1}^M [\|\mathbf{e}_{m1}^{k+1}\|_2^2 + \|\mathbf{e}_{m2}^{k+1}\|_2^2 + \|\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k\|_2^2 + \|X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)\|_2^2] \\
&\quad + \sum_{m=1}^M \eta (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)^T \mathbf{e}_{m1}^{k+1} - \sum_{m=1}^M \eta [X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T \mathbf{e}_{m2}^{k+1} \\
&\leq -\eta \sum_{m=1}^M [\|\mathbf{e}_{m1}^{k+1}\|_2^2 + \|\mathbf{e}_{m2}^{k+1}\|_2^2 + \|\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k\|_2^2 + \|X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)\|_2^2],
\end{aligned} \tag{14}$$

where the second inequality is obtained by using (4) and the fact that

$$\begin{aligned}
&\sum_{m=1}^M \eta (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)^T \mathbf{e}_{m1}^{k+1} - \sum_{m=1}^M \eta [X_m^* (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)]^T \mathbf{e}_{m2}^{k+1} \\
&= \sum_{m=1}^M (\boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k)^T [(\mathbf{u}_m^{k+1} - X_m^{*T} \mathbf{v}_m^{k+1}) - (\mathbf{u}_m^k - X_m^{*T} \mathbf{v}_m^k)] = 0.
\end{aligned}$$

The inequality (14) implies that D^k is decreasing with respect to k . Meanwhile, we have $D^k \geq 0$ by its definition. Therefore, the sequence $\{D^k, k = 1, 2, \dots\}$ must be convergent and hence we have $\mathbf{e}_{m1}^{k+1} = \boldsymbol{\beta}_m^{k+1} - \boldsymbol{\beta}_m^k \rightarrow \mathbf{0}_p$, $\mathbf{e}_{m2}^{k+1} = \mathbf{y}_m^* - X_m^* \boldsymbol{\beta}_m^{k+1} - \mathbf{r}_m^{k+1} \rightarrow \mathbf{0}_{n_m^*}$, $\boldsymbol{\beta}_m^{k+1} \rightarrow \boldsymbol{\beta}_m^k$ and $X_m^* \boldsymbol{\beta}_m^{k+1} \rightarrow X_m^* \boldsymbol{\beta}_m^k$. That is to say, there must exist $\tilde{\boldsymbol{\beta}}_m^*$, $\tilde{\boldsymbol{\beta}}^*$ and $\tilde{\mathbf{r}}_m^*$, such that $\boldsymbol{\beta}_m^{k+1} \rightarrow \tilde{\boldsymbol{\beta}}_m^*$, $X_m^* \boldsymbol{\beta}_m^{k+1} \rightarrow X_m^* \tilde{\boldsymbol{\beta}}_m^*$, $\boldsymbol{\beta}_m^{k+1} \rightarrow \tilde{\boldsymbol{\beta}}^*$ and $\mathbf{r}_m^{k+1} \rightarrow \tilde{\mathbf{r}}_m^*$. In addition, $\tilde{\boldsymbol{\beta}}_m^* - \tilde{\boldsymbol{\beta}}^* = \mathbf{0}_p$ and $\mathbf{y}_m^* - X_m^* \tilde{\boldsymbol{\beta}}_m^* - \tilde{\mathbf{r}}_m^* = \mathbf{0}_{n_m^*}$ hold, which means $(\tilde{\boldsymbol{\beta}}^*, \tilde{\mathbf{r}}_m^*, \tilde{\boldsymbol{\beta}}_m^*)$ is a feasible solution to the problem (10) in the main article. Besides, from the fact that $D^{k+1} \leq D^0$, we know that both $\boldsymbol{\beta}_m^{k+1}$ and $X_m^* \boldsymbol{\beta}_m^{k+1}$ are bounded. Then from the inequality (12), we finally obtain

$$\lim_{k \rightarrow \infty} \sum_{m=1}^M \rho_\tau(\mathbf{r}_m^{k+1}) = \sum_{m=1}^M \rho_\tau(\tilde{\mathbf{r}}_m^*) = \sum_{m=1}^M \rho_\tau(\mathbf{r}_m^*),$$

and hence $(\tilde{\beta}^*, \tilde{r}_m^*, \tilde{\beta}_m^*)$ is actually an optimal solution to (10) in the main article.

2 Proof of Theorem 2

2.1 Proof for the IPW-renewable estimator's consistency

We prove this consistency by using the induction technique. In the implementation of IPW-renewable, if a consistent initial estimator is not available, we set the initial value of the algorithm as the solution to $Q_{n_1^*}(\beta) = \mathbf{0}_p$, from which the obtained estimator $\hat{\beta}^{(1)}$ is $\sqrt{n_1^*}$ -consistent for β_0 (He et al. 2023), as this local estimation process does not involve any approximation. Next, we assume that the estimators $\hat{\beta}^{(m)}$ ($m = 1, 2, \dots, M-1, M \geq 1$) are consistent, and prove that $\hat{\beta}^{(M)}$ is also consistent.

From the algorithm details of IPW-renewable given in Algorithm 2 in the main article, we know that $\hat{\beta}^{(M)}$ is obtained by solving

$$Q_{n_M^*}(\beta) + \sum_{m=1}^{M-1} \nabla Q_{n_m^*}(\hat{\beta}^{(m)}) (\beta - \hat{\beta}^{(M-1)}) = \mathbf{0}_p.$$

That is, we have

$$Q_{n_M^*}(\hat{\beta}^{(M)}) + \sum_{m=1}^{M-1} \nabla Q_{n_m^*}(\hat{\beta}^{(m)}) (\hat{\beta}^{(M)} - \hat{\beta}^{(M-1)}) = \mathbf{0}_p. \quad (15)$$

On the other hand, by utilizing the fact that $E[Q(\beta_0)] = \mathbf{0}_p$, the Law of Large Numbers and Assumption 1(c), which implies that $n_1^* \rightarrow \infty$ and further $n^* = \sum_{m=1}^M n_m^* \rightarrow \infty$ under our setup of the initial value for the algorithm, we can obtain

$$\frac{1}{n^*} Q_{n_M^*}(\beta_0) + \frac{1}{n^*} \left[- \sum_{m=1}^{M-1} \nabla Q_{n_m^*}(\hat{\beta}^{(m)}) \right] (\hat{\beta}^{(M-1)} - \beta_0) = o_p(1). \quad (16)$$

Combining (15) and (16), we then have

$$\frac{1}{n^*} \left[Q_{n_M^*}(\beta_0) - Q_{n_M^*}(\hat{\beta}^{(M)}) \right] + \frac{1}{n^*} \left[- \sum_{m=1}^{M-1} \nabla Q_{n_m^*}(\hat{\beta}^{(m)}) \right] (\hat{\beta}^{(M)} - \beta_0) = o_p(1). \quad (17)$$

Besides, by taking the first-order Taylor expansion for $Q_{n_M^*}(\hat{\beta}^{(M)})$ at β_0 , we have

$$Q_{n_M^*}(\hat{\beta}^{(M)}) - Q_{n_M^*}(\beta_0) = [\nabla Q_{n_M^*}(\xi_M) - \nabla Q_{n_M^*}(\beta_0) + \nabla Q_{n_M^*}(\beta_0)] (\hat{\beta}^{(M)} - \beta_0),$$

where

$$\|\nabla Q_{n_M^*}(\xi_M) - \nabla Q_{n_M^*}(\beta_0)\|_2 \leq C_M \|\xi_M - \beta_0\|_2 \leq C_M \left\| \hat{\beta}^{(M)} - \beta_0 \right\|_2,$$

under Assumption 1(b), which ensures the Lipschitz continuity of the Hessian matrix $\nabla Q_{n_M^*}(\beta)$. Here ξ_M lies between $\hat{\beta}^{(M)}$ and β_0 , and C_M is a constant bounded by the local sample size n_M^* . Then we have

$$Q_{n_M^*}(\hat{\beta}^{(M)}) - Q_{n_M^*}(\beta_0) = \nabla Q_{n_M^*}(\beta_0) (\hat{\beta}^{(M)} - \beta_0) + O_p \left(n_M^* \left\| \hat{\beta}^{(M)} - \beta_0 \right\|_2^2 \right). \quad (18)$$

Combining (17) and (18), we obtain

$$\frac{1}{n^*} \left[- \sum_{m=1}^{M-1} \nabla Q_{n_m^*}(\hat{\beta}^{(m)}) - \nabla Q_{n_M^*}(\beta_0) \right] (\hat{\beta}^{(M)} - \beta_0) + O_p \left(\frac{n_M^*}{n^*} \left\| \hat{\beta}^{(M)} - \beta_0 \right\|_2^2 \right) = o_p(1), \quad (19)$$

which then implies $\left\| \hat{\beta}^{(M)} - \beta_0 \right\|_2 = o_p(1)$. This completes the proof for the consistency.

2.2 Proof for the IPW-renewable estimator's asymptotic normality

Firstly, by taking the first-order Taylor expansion for $Q_{n_1^*}(\beta_0)$ at $\hat{\beta}^{(1)}$, we have

$$Q_{n_1^*}(\beta_0) = -\nabla Q_{n_1^*}(\hat{\beta}^{(1)}) \left(\hat{\beta}^{(1)} - \beta_0 \right) + O_p \left(n_1^* \left\| \hat{\beta}^{(1)} - \beta_0 \right\|_2^2 \right),$$

where we utilized the fact that $Q_{n_1^*}(\hat{\beta}^{(1)}) = \mathbf{0}_p$. Besides, by combining (15) and (18), we have

$$\begin{aligned} Q_{n_m^*}(\theta_0) &= \left[-\sum_{j=1}^{m-1} \nabla Q_{n_j^*}(\hat{\beta}^{(j)}) - \nabla Q_{n_m^*}(\beta_0) \right] \left(\hat{\beta}^{(m)} - \beta_0 \right) \\ &\quad + \left[\sum_{j=1}^{m-1} \nabla Q_{n_j^*}(\hat{\beta}^{(j)}) \right] \left(\hat{\beta}^{(m-1)} - \beta_0 \right) + O_p \left(n_m^* \left\| \hat{\beta}^{(m)} - \beta_0 \right\|_2 \right), \end{aligned}$$

holding for $m \geq 2$. Cumulating $Q_{n_m^*}(\theta_0)$'s by m , we then obtain

$$\begin{aligned} \sum_{m=1}^M Q_{n_m^*}(\theta_0) &= \left[-\sum_{m=1}^{M-1} \nabla Q_{n_m^*}(\hat{\beta}^{(M)}) - \nabla Q_{n_M^*}(\beta_0) \right] \left(\hat{\beta}^{(M)} - \beta_0 \right) + O_p \left(\sum_{m=1}^M n_m^* \left\| \hat{\beta}^{(m)} - \beta_0 \right\|_2^2 \right) \\ &= \left[-\sum_{m=1}^M \nabla Q_{n_m^*}(\beta_0) \right] \left(\hat{\beta}^{(M)} - \beta_0 \right) + O_p \left(\sum_{m=1}^M n_m^* \left\| \hat{\beta}^{(m)} - \beta_0 \right\|_2^2 \right), \end{aligned} \tag{20}$$

and further

$$\frac{1}{n^*} \sum_{m=1}^M Q_{n_m^*}(\beta_0) = \frac{1}{n^*} \left[-\sum_{m=1}^M \nabla Q_{n_m^*}(\beta_0) \right] \left(\hat{\beta}^{(M)} - \beta_0 \right) + O_p \left(\frac{1}{n^*} \sum_{m=1}^M n_m^* \left\| \hat{\beta}^{(m)} - \beta_0 \right\|_2^2 \right). \tag{21}$$

Note that all the estimators $\hat{\boldsymbol{\beta}}^{(m)}$'s ($m \geq 1$) are consistent, and thus (21) implies that the term $\frac{1}{n^*} \left[-\sum_{m=1}^M \nabla Q_{n_m^*}(\boldsymbol{\beta}_0) \right] \left(\hat{\boldsymbol{\beta}}^{(M)} - \boldsymbol{\beta}_0 \right)$ is asymptotically equivalent to $\frac{1}{n^*} \sum_{m=1}^M Q_{n_m^*}(\boldsymbol{\beta}_0)$. On the other hand, by the Central Limit Theorem we have

$$\frac{1}{\sqrt{n^*}} \sum_{m=1}^M Q_{n_m^*}(\boldsymbol{\beta}_0) \rightarrow_d \mathcal{N}(\mathbf{0}_p, E(Q(\boldsymbol{\beta}_0) Q^T(\boldsymbol{\beta}_0))),$$

and by the Law of Large Numbers we have

$$\frac{1}{n^*} \left[\sum_{m=1}^M \nabla Q_{n_m^*}(\boldsymbol{\beta}_0) \right] \rightarrow \nabla Q(\boldsymbol{\beta}_0). \quad (22)$$

And the asymptotic normality follows.

References

He X, Pan X, Tan KM, et al (2023) Smoothed quantile regression with large-scale inference. *Journal of Econometrics* 232(2):367–388. <https://doi.org/10.1016/j.jeconom.2021.07.010>