Estimation and Testing of Varying Coefficients in Quantile Regression

Xingdong Feng and Liping Zhu *

Abstract

The proofs of Theorems 1, 2 and 3 are given in this supplementary file.

Key Words: Dimension reduction; Hypothesis test; Quantile regression; Singular value decomposition.

* Xingdong Feng is Associate Professor, School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China (Email: feng.xingdong@mail.shufe.edu.cn). Liping Zhu is corresponding author and Professor, School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China (Email: zhu.liping@mail.shufe.edu.cn). They are also affiliated with the Key Laboratory of Mathematical Economics (SUFE), Ministry of Education. This work was supported by National Natural Science Foundation of China (11101254, 11371236 and 11422107), Pujiang Project of Science and Technology Commission of Shanghai Municipality (13PJOC048), Henry Fok Education Foundation Fund of Young College Teachers (141002), Programs for New Century Excellent Talents (NCET-12-0901) and Innovative Research Team in University of China (IRT13077), Ministry of Education of China.
Appendix

Proof of Theorem 1

In the proof, we impose the restriction on the matrix $A$ as given in (6) for estimation, and the argument is similar when restrictions are considered on both matrices $A$ and $B$.

By Assumptions (A3) and (A4), we can consider a bounded space $\Theta_0 \subset \mathbb{R}$ such that all the elements of the matrices $A$, $B$, $A_0$ and $B_0$ belong to this space.

We first establish the consistency of the estimator by minimizing (6). Let $H_n = \sum_{i=1}^{n} z_i z_i^T$,

$$\varpi = H_n \text{vec} \left( \begin{bmatrix} B^T - B_0^T \\ A B^T - A_0 B_0^T \end{bmatrix} \right),$$

and $v_i = H_n^{-1} z_i$. Consider a subspace $\Theta_n$ of $\mathbb{R}^{\zeta_n}$ as

$$\Theta_n = \{ \varpi \in \mathbb{R}^{\zeta_n} : \text{all the elements of matrices } A \text{ and } B \text{ belong to } \Theta_0 \},$$

where $\zeta_n = m(\kappa_n + p)$. By Lemma 3.2, Lemma 3.3 and Proof of Theorem 2.1 of He and Shi (1994), for any $\epsilon$ there exists $L_\epsilon$ such that

$$P \left\{ \inf_{\vartheta^* \in \mathbb{R}^{\zeta_n} : \vartheta^*_n \geq L_\epsilon \kappa_n^{1/2}} \sum_{i=1}^{n} \rho_r(e_i - \kappa_n^{1/2} v_i^T \vartheta^* - x_i^T \Delta_i) > \sum_{i=1}^{n} \rho_r(e_i - x_i^T \Delta_i) \right\} > 1 - \epsilon,$$

as $n \to \infty$. Since $\Theta_n \subset \mathbb{R}^{\zeta_n}$, we have

$$\inf_{\vartheta^* \in \mathbb{R}^{\zeta_n} : \vartheta^*_n \geq L_\epsilon \kappa_n^{1/2}} \sum_{i=1}^{n} \rho_r(e_i - \kappa_n^{1/2} v_i^T \vartheta^* - x_i^T \Delta_i) \geq \inf_{\vartheta^* \in \mathbb{R}^{\zeta_n} : \vartheta^*_n \geq L_\epsilon \kappa_n^{1/2}} \sum_{i=1}^{n} \rho_r(e_i - \kappa_n^{1/2} v_i^T \vartheta^* - x_i^T \Delta_i).$$

It then follows that

$$P \left\{ \inf_{\vartheta^* \in \mathbb{R}^{\zeta_n} : \vartheta^*_n \geq L_\epsilon \kappa_n^{1/2}} \sum_{i=1}^{n} \rho_r(e_i - \kappa_n^{1/2} v_i^T \vartheta^* - x_i^T \Delta_i) > \sum_{i=1}^{n} \rho_r(e_i - x_i^T \Delta_i) \right\} > 1 - \epsilon,$$
Also note that
\[
\sum_{i=1}^{n} \rho_r(e_i - \kappa_n^{1/2}v_i^T\hat{\omega} - x_i^T\Delta_i) = \inf_{\omega \in \Theta_n} \sum_{i=1}^{n} \rho_r(e_i - \kappa_n^{1/2}v_i^T\omega - x_i^T\Delta_i)
\]
holds with probability one, and \(0 \in \Theta_n\). Thus, we have \(P(\|\hat{\omega}\| \leq L\kappa_n^{1/2}) > 1 - \epsilon\). It follows that
\[
\left\| \begin{pmatrix} \hat{B}^T - B_0^T \\ \hat{A}\hat{B}^T - A_0B_0^T \end{pmatrix} \right\|_F = O((\kappa_n/n)^{1/2}).
\]
We immediately have \(\|\hat{B} - B_0\|_F = O((\kappa_n/n)^{1/2})\). Note that \(\hat{A} = \hat{A}\hat{B}^T\{\hat{B}(\hat{B}^T\hat{B})^{-1}\}\), so it follows that \(\|\hat{A} - A_0\|_F = O((\kappa_n/n)^{-1/2})\).

Now we verify Conditions (C0)–(C5) of He and Shao (2000) to establish the Bahadur representation of the estimator. Let
\[
\vartheta = vec \begin{pmatrix} B^T \\ AB^T \end{pmatrix}.
\]
Note that
\[
x_i^T \begin{pmatrix} B^T \\ AB^T \end{pmatrix} \Pi_{\kappa_n}(T_i) = \vartheta^T vec(\Pi_{\kappa_n}(T_i) \otimes x_i),
\]
and
\[
vec \begin{pmatrix} B^T \\ AB^T \end{pmatrix} = \begin{pmatrix} b_1 \\ Ab_1 \\ \vdots \\ b_{\kappa_n+p} \\ Ab_{\kappa_n+p} \end{pmatrix},
\]
where \(b_j\) is the \(j\)-th row of the matrix \(B\). Let \(z_{ij} = \pi_j(T_i) \otimes x_i\), and let \(z_{ij}^{(1)}\) and \(z_{ij}^{(2)}\) be the first \(k\) and last \(m-k\) components of the vector \(z_{ij}\), respectively. We minimize the objective
function

\[ L_n(A, B) = \sum_{i=1}^{n} \rho (Y_i - \sum_{j=1}^{k_n+p} z_{ij}^T \left( \begin{array}{c} b_j \\ A b_j \end{array} \right) ) \].

Thus, the score functions are

\[ \frac{\partial L_n}{\partial b_j} = -\sum_{i=1}^{n} \psi_T \left( Y_i - \sum_{j=1}^{k_n+p} z_{ij}^T \left( \begin{array}{c} b_j \\ A b_j \end{array} \right) \right) (I_k A^T) z_{ij}, \]

and

\[ \frac{\partial L_n}{\partial vec(A)} = -\sum_{i=1}^{n} \psi_T \left( Y_i - \sum_{j=1}^{k_n+p} z_{ij}^T \left( \begin{array}{c} b_j \\ A b_j \end{array} \right) \right) \sum_{j=1}^{n_{n+p}} \left\{ b_j \otimes z_{ij}^{(2)} \right\}, \]

where \( \psi_T(u) = \tau - I(u < 0) \).

Let

\[ \phi(\theta, h_i) = \psi_T \left( Y_i - \theta^T z_i \right) \left( \begin{array}{c} I_{n+k+p} \otimes \left( \begin{array}{c} I_k \\ A^T \end{array} \right) \\ B^T \otimes \left( \begin{array}{c} 0 \\ I_{m-k} \end{array} \right) \end{array} \right) z_i, \]

where \( h_i = (x_i^T, Y_i, T_i)^T \), and \( z_i = \Pi(T_i) \otimes x_i \). The score can be represented as

\[ \frac{\partial L_n}{\partial \theta} = \sum_{i=1}^{n} \phi(\theta, h_i). \]

With the similar arguments as the proof of Theorem 3.2 of He and Shao (1996), we have

\[ E_{\theta_2} \sup_{\theta_2: \|\theta_2 - \theta_1\| \leq d} \| \eta_i(\theta_1, \theta_2) \|^2 \leq K \| z_i \|^2 d, \]

for some constant \( K \), where

\[ \eta_i(\theta_1, \theta_2) = \phi(\theta_1, h_i) - \phi(\theta_2, h_i) - E \{ \phi(\theta_1, h_i) - \phi(\theta_2, h_i) \}. \]

It follows from Assumptions (A2) that \( \sum_{i=1}^{n} E \| \phi(\theta_0, h_i) \|^2 = O(n \kappa_n) \). Hence, Conditions (C0)–(C2) of He and Shao (2000) are satisfied under Assumptions (A2), (A5) and (A6).
Consider $\alpha \in S_{k(\kappa_n + p + m - k)}$. Under Assumption (A1), we have

$$\sum_{i=1}^{n} |\alpha^T \eta_i(\theta_1, \theta_2)|^2$$

$$\leq K_1 \sum_{i=1}^{n} \left\{ (\tilde{\alpha}_1^T z_i)^2 |\psi_\tau (Y_i - \vartheta^T z_i)|^2 + (\tilde{\alpha}_2^T z_i)^2 \|\theta_1 - \theta_2\|^2 \right\}$$

$$\leq K_2 \sum_{i=1}^{n} (\tilde{\alpha}_1^T z_i)^2 \left[ I\{|Y_i - \vartheta^T z_i| \leq |(\vartheta_1 - \vartheta_2)^T z_i|\} + \{(\vartheta_1 - \vartheta_2)^T z_i\}^2 \right]$$

$$+ K_3 \sum_{i=1}^{n} (\tilde{\alpha}_2^T z_i)^2 \|\theta_1 - \theta_2\|^2,$$

where $\tilde{\alpha}_1, \tilde{\alpha}_2 \in S_{k(\kappa_n + p + m - k)}$, and $K_l (l = 1, 2, 3)$ are some constants. Therefore, by Lemma 2.2 of He and Shao (1996), Conditions (C4) and (C5) of He and Shao (1996) are satisfied under Assumptions (A5) and (A6) with

$$\sup_{\|\theta_2 - \theta_1\| \leq K(\kappa_n/n)^{1/2}} \sum_{i=1}^{n} E_{\theta_1} |\alpha^T \eta_i(\theta_1, \theta_2)|^2 = O((n\kappa_n)^{1/2}),$$

and

$$\sup_{\|\theta_2 - \theta_1\| \leq K(\kappa_n/n)^{1/2}} \sum_{i=1}^{n} |\alpha^T \eta_i(\theta_1, \theta_2)|^2 = O_p((n\kappa_n)^{1/2}),$$

where $K > 0$.

Note that

$$E \left\{ \psi_\tau (Y_i - \vartheta^T z_i) |x_i, T_i \right\}$$

$$= f_i(0) \left\{ (\vartheta - \vartheta_0)^T z_i - x_i^T \Delta_i \right\} + O(||z_i^T (\vartheta - \vartheta_0) - x_i^T \Delta_i||^{3/2}).$$

The derivative of the vector

$$f_i(0)(\vartheta - \vartheta_0)^T z_i$$

$$\left( \begin{array}{c} I_{k_n+p} \otimes \left( \begin{array}{c} I_k \ A^T \end{array} \right) \\ B^T \otimes \left( \begin{array}{c} 0 \ I_{m-k} \end{array} \right) \end{array} \right) z_i,$$
with respect to $\vartheta$, can be written as

$$D_{n\kappa_n}(\theta) = \begin{pmatrix} D_{11}(\theta) & D_{12}(\theta) \\ D_{12}^T(\theta) & D_{22}(\theta) \end{pmatrix},$$

where

$$D_{11}(\theta) = \sum_{i=1}^{n} f_i(0) \begin{pmatrix} (I_k \ A^T) z_{i1} \\ \vdots \\ (I_k \ A^T) z_{i,\kappa_n+p} \end{pmatrix},$$

$$D_{12}(\theta) = -\sum_{i=1}^{n} f_i(0) \left( \begin{pmatrix} (I_k \ A^T) z_{i1} \\ \vdots \\ (I_k \ A^T) z_{i,\kappa_n+p} \end{pmatrix} \right) k_{n+p} \sum_{j=1}^{k_n+p} \left\{ b_j^T \otimes z_{ij}^{(2)T} \right\},$$

$$-\sum_{i=1}^{n} f_i(0) \sum_{j=1}^{k_n+p} z_{ij}^T \left( b_j - b_{0j} \right) \left( \begin{pmatrix} I_k \otimes z_{i1}^{(2)T} \\ \vdots \\ I_k \otimes z_{i,\kappa_n+p}^{(2)T} \end{pmatrix} \right),$$

and

$$D_{22}(\theta) = \sum_{i=1}^{n} f_i(0) \left[ \sum_{j=1}^{k_{n+p}} \left\{ b_j^T \otimes z_{ij}^{(2)T} \right\} \right] \left[ \sum_{j=1}^{k_{n+p}} \left\{ b_j^T \otimes z_{ij}^{(2)T} \right\} \right].$$

Thus,

$$D_{n\kappa_n}(\theta_0) = \begin{pmatrix} I_{\kappa_n+p} \otimes \left( I_k \ A_0^T \right) \\ B_0^T \otimes \left( 0 \ I_{m-k} \right) \end{pmatrix} \left\{ \sum_{i=1}^{n} f_i(0) z_i z_i^T \right\} \left( I_{\kappa_n+p} \otimes \left( I_k \ A_0^T \right) \right)^T.$$

Under Assumptions (A5) and (A6), we have

$$\sup_{||\theta-\theta_0|| \leq K(\kappa_n/n)^{1/2}} \left| \alpha^T \sum_{i=1}^{n} E(\phi(\theta, h_i) - \phi(\theta_0, h_i)) - \alpha^T D_{n}(\theta_0)(\theta - \theta_0) \right| = o(n^{1/2}).$$
uniformly for \( \alpha \in \mathbb{S}_{k(\kappa_0 + p + m - k)} \). By Lemma 4 of Horowitz and Mammen (2004), \( \| n^{-1} D_{n\kappa} (\widehat{\theta}_0) - D_{0n} \|^2 = O_p(\kappa_n^2 / n) \), where \( D_{0n} \) is defined in (10). The Bahadur representation (9) of the estimator \( \widehat{\theta} \) then follows from Lemma 3.2 of He and Shao (2000). \( \diamond \)

**Proof of Theorem 2**

In the proof, we consider the general case where no restrictions are imposed on the matrices \( A_0 \) and \( B_0 \), and the proof will be similar if we consider the forms (8).

Denote the columns of the matrices \( \widehat{A}, A_0, \widehat{B} \) and \( B_0 \) as \( \{ \widehat{p}_1, \cdots, \widehat{p}_k \}, \{ p_{01}, \cdots, p_{0k} \}, \{ \widehat{q}_1, \cdots, \widehat{q}_k \} \) and \( \{ q_{01}, \cdots, q_{0k} \} \), respectively. There exist some \( \overline{v}_{\kappa_n}(k) \times m \) matrices \( P_j \) and \( \overline{v}_{\kappa_n}(k) \times (\kappa_n + p) \) matrices \( Q_j \) composed of zeros and ones, \( j = 1, \cdots, k \), such that \( \widehat{p}_j = P_j^T \theta_0, p_{0j} = P_j^T \theta_0, \widehat{q}_j = Q_j^T \theta_0, q_{0j} = Q_j^T \theta_0, P_j^T 1_{\overline{v}_{\kappa_n}(k)} = 1_m \), and \( Q_j^T 1_{\overline{v}_{\kappa_n}(k)} = 1_{\kappa_n+p} \), where \( \overline{v}_{\kappa_n}(k) = (k+1)(\kappa_n + p + m) \), \( \widehat{\theta} = (vec(\widehat{B}^T), vec(\widehat{A}))^T \) and \( \theta_0 = (vec(B_0^T), vec(A_0))^T \). Actually, the matrices \( P_j \) and \( Q_j \) are used to map the vectors back to matrix forms, which is the inverse operation of vector operator \( vec(\cdot) \).

Consider matrix difference

\[
\| \widehat{A} \widehat{B}^T - A_0 B_0^T \|_F
= \left\| \sum_{j=1}^{k+1} (\widehat{p}_j \widehat{q}_j^T - p_{0j} q_{0j}^T) \right\|_F \\
\leq \left\| \sum_{j=1}^{k+1} \left\{ (\widehat{p}_j - p_{0j}) q_{0j}^T + p_{0j} (\widehat{q}_j - q_{0j})^T \right\} \right\|_F + \left\| \sum_{j=1}^{k+1} (\widehat{p}_j - p_{0j}) (\widehat{q}_j - q_{0j})^T \right\|_F \\
= \left\| \sum_{j=1}^{k+1} \left\{ (\widehat{p}_j - p_{0j}) q_{0j}^T + p_{0j} (\widehat{q}_j - q_{0j})^T \right\} \right\|_F + O_p(\kappa_n / n) \\
= \left\| \sum_{j=1}^{k+1} P_j^T \left\{ (\widehat{\theta} - \theta_0) \theta_0^T + \theta_0 (\widehat{\theta} - \theta_0)^T \right\} Q_j \right\|_F + O_p(\kappa_n / n). \]
With the Bahadur representation (9) of the estimator, we have

\[
\tilde{PH} = (\hat{P} - \tilde{P}_0) \tilde{H} = -\sum_{i=1}^{n} \psi_\tau(e_i) C^T \left( \sum_{j=1}^{k} P^T_j (z_j \theta_0^T + \theta_0 z_i^T) Q_j \right) \tilde{H} + r_1 + r_2 + r_3,
\]

where \( \|r_1\|_F = O_p(\kappa_n^{-r+1/2}) \), \( \|r_2\|_F = o_p(\kappa_n^{1/2}) \), and \( \|r_3\|_F = O_p(\kappa_n/n) \).

Thus, under Assumption (A5)–(A6), we have

\[
T_n^{(1)} = \left( \sum_{i=1}^{n} \tilde{x}_i^{(1)} x_i^{(1)T} \right)^{-1/2} \sum_{i=1}^{n} \psi_\tau(e_i) \tilde{x}_i^{(1)T} / \{\tau(1 - \tau)\}^{1/2} + o_p(1).
\]

The result follows directly from Theorem 4.1 of Portnoy (1985).

Proof of Theorem 3

The proof is similar to that of Theorem 2.

References

He, X. and Shao, Q. (1996), “A general Bahadur representation of M-estimators and its application to linear regression with nonstochastic designs,” The Annals of Statistics, 24, 2608–2630.

— (2000), “On parameters of increasing dimensions,” Journal of Multivariate Analysis, 73, 120–135.

He, X. and Shi, P. (1994), “Convergence rate of B-spline estimator of nonparametric conditional quantile functions,” Nonparametric Statistics, 3, 299–308.
Horowitz, J. L. and Mammen, E. (2004), “Nonparametric estimation of an additive model with a link function,” *The Annals of Statistics*, 32, 2412–2443.

Portnoy, S. (1985), “Asymptotic behavior of M-estimators of $p$ regression parameters when is $p^2/n$ is large; II. Normal approximation,” *The Annals of Statistics*, 13, 1403–1417.