# High-quantile regression for tail-dependent time series by Ting zhang

Zhenyi Lei

October 29, 2024

# Introduction

Quantile regression is a powerful method for quantile analysis with given regressors. It can capture the unequal variation of one variable for different ranges of another variable and are very useful in ecology, Climatology and Economics.

High-quantile regression : quantile level can grow with the sample size to capture the tail phenomena.

Existing framework: strong mixing condition (cons: need additional condition which many extreme value data cannot meet )

Improvement: propose a new framework can handle with dependent data and find the CLT for the estimator.

# Tail adversarial stability

consider a row-wise stationary triangular array of random variables $U_{1,n}, \cdots, U_{n,n}$ with marginal distribution function $F_n(u) = Pr(U_{1,n} \leq u)$, so $F_n^{-1}(1-\alpha)$ be the empirical $1-\alpha$ quantile of $F_n(\cdot)$. Assume :

$$U_{i,n} = G_n(\mathcal{F}_i), \quad \mathcal{F}_i = (\ldots, \epsilon_{i-1}, \epsilon_i) \tag{1}$$

Then denote $\mathcal{F}_k^* = (\mathcal{F}_{-1}, \epsilon_0^*, \epsilon_1, \ldots, \epsilon_k)$, where $\epsilon_0^*$ be identically distributed as $\epsilon_0$ and independent with other $\epsilon_i$.We have $U_{i,n}^* = G_n(\mathcal{F}_i^*)$

With given notation, the paper defines:

$$\theta_{n,\alpha}(k) = \sup_{a\in(0,\alpha],N\geq n} \Pr\left\{ U^*_{k,N} \leq F_N^{-1}(1-a) \,\middle|\, U_{k,N} > F_N^{-1}(1-a) \right\}, \quad \alpha \in (0,1) \tag{2}$$

This is the adversarial tail dependence measure.

$$\Theta_{n,\alpha,q}(m) = \sum_{k=m}^{\infty} \{\theta_{n,\alpha}(k)\}^{1/q}, \quad m > 0 \tag{3}$$

which measures the cumulative tail adversarial effect of the current innovation $\epsilon_0$ on future observations from time m.

Tail adversarial q-stability measure:

$$\lim_{\alpha\downarrow 0} \lim_{n\to\infty} \Theta_{n,\alpha,q}(0) < \infty \tag{4}$$

# TAS$_q$ *example*

For moving-maximum model where $U_i = \sum_{l=0}^{\infty} a_l \, \epsilon_{i-l}$ for $i = 1, \ldots, n$

$$\lim_{\alpha \downarrow 0} \Theta_{\alpha,q}(0) \leq 2(\sum_{l=0}^{\infty} a_l^{\gamma})^{-1}(\sum_{k=0}^{\infty} a_k^{\gamma/q}) \tag{5}$$

under the condition $(\sum_{l=0}^{\infty} a_l^{\gamma})^{-1} > 0$and $\sum_{k=0}^{\infty} a_k^{\gamma/q} < \infty$

Given triangular array of random variables $Y_{1,n}, \cdots, Y_{n,n}$ associated with a set of regressors $x_{1,n}, \cdots, x_{n,n} \in \mathbb{R}^p$ have the high-quantile regression:

$$Y_{i,n} = x_{i,n}^T \beta_n + U_{i,n}, \quad (i = 1, \ldots, n) \tag{6}$$

So the esitmator is :

$$\hat{\beta}_n = \arg\min_{\eta \in \mathbb{R}^p} \sum_{i=1}^{n} \phi_{1-\alpha_n} \left( Y_{i,n} - x_{i,n}^T \eta \right) \tag{7}$$

where$\phi_{1-\alpha_n}(y) = (1 - \alpha_n) y^+ + \alpha_n (-y)^+$ and $y^+ = \max(y, 0)$

# Standralize

For convenience, define:$\Sigma_n = \frac{1}{n}\sum_{i=1}^{n} x_{i,n}x_{i,n}^T \in \mathbb{R}^{p\times p}$ where $\Sigma_n$ is nonsingular for all large $n$, and let $z_{i,n} = \Sigma_n^{-1/2}x_{i,n}$. Then,$\frac{1}{n}\sum_{i=1}^{n} z_{i,n}z_{i,n}^T = I_{p\times p}$ and $\varphi_n = \Sigma_n^{1/2}\beta_n$. By rewriting the function: the problem is to solve

$$\hat{\varphi}_n = \arg\min_{\eta\in\mathbb{R}^p} \sum_{i=1}^{n} \phi_{1-\alpha_n}\left(Y_{i,n} - z_{i,n}^T\eta\right) \tag{8}$$

# Order of $\hat{\varphi}_n$

*Condition* 1. The triangular array $(U_{i,n}) \in \text{TAS}_q$ for some $q \geqslant 2$.

*Condition* 2. There exists an $\alpha \in (0,1)$ such that $F_n(\cdot)$ is continuously differentiable with uniformly bounded and strictly positive derivative $f_n(\cdot)$ in its upper tail $\{F_n^{-1}(1-\alpha), F_n^{-1}(1)\}$ with $\liminf_{n\to\infty} |F_n^{-1}(1) - F_n^{-1}(1-\alpha)| > 0$ for all large $n$.

*Condition* 3. The rescaled design satisfies $\max_{1\leqslant i\leqslant n} |z_{i,n}| = o\{(n\alpha_n)^{1/2}\}$.

THEOREM 1. *Assume that Conditions 1–3 hold, $\alpha_n \to 0$, and $n\alpha_n \to \infty$. If*

$$\tau_n = (n\alpha_n)^{1/2} \frac{f_n(0)}{1 - F_n(0)} \to \infty,$$

$\max_{1\leqslant i\leqslant n} |z_{i,n}| = o(\tau_n)$, *and*

$$\max_{1\leqslant i\leqslant n} \sup_{|\eta|\leqslant c} \left| \frac{f_n(\tau_n^{-1} z_{i,n}^{\mathrm{T}} \eta) - f_n(0)}{f_n(0)} \right| \to 0$$

*for any $c < \infty$, then $\hat{\varphi}_n - \varphi_n \to 0$ in probability and*

$$\hat{\varphi}_n - \varphi_n = O_{\mathrm{p}}(\tau_n^{-1}).$$

THEOREM 2. *Assume that the conditions of Theorem 1 hold. If the limits*

$$\rho_k = \lim_{n\to\infty} \operatorname{cor}(\mathbb{1}_{\{U_{0,n}>0\}}, \mathbb{1}_{\{U_{k,n}>0\}}), \quad \Upsilon_k = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n-|k|} z_{i,n} z_{i+|k|,n}^{\mathrm{T}}$$

*exist for each $k \in \mathbb{Z}$, then the matrix*

$$\Gamma = \sum_{k\in\mathbb{Z}} \rho_k \Upsilon_k$$

*is positive semidefinite with bounded eigenvalues. If in addition the eigenvalues of $\Gamma$ are bounded away from zero, then we have the central limit theorem*

$$\tau_n(\hat{\varphi}_n - \varphi_n) \to N(0, \Gamma) \tag{11}$$

*in distribution.*

# Simulation

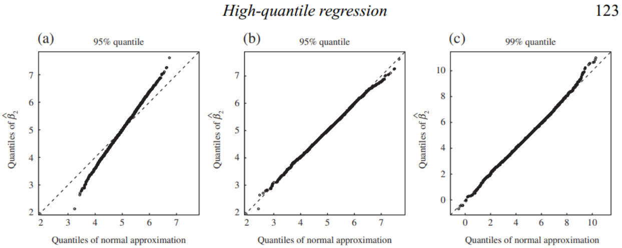


Fig. 1. Q-Q plots for the high-quantile regression estimator $\hat{\beta}_2$: plotted are quantiles of $\hat{\beta}_2$ against quantiles of the normal approximation (a) ignoring the tail dependence, (b) based on the central limit theorem for the 95% quantile, and (c) based on the central limit theorem for the 99% quantile; the dashed lines in all plots have unit slope and zero intercept.

# real data

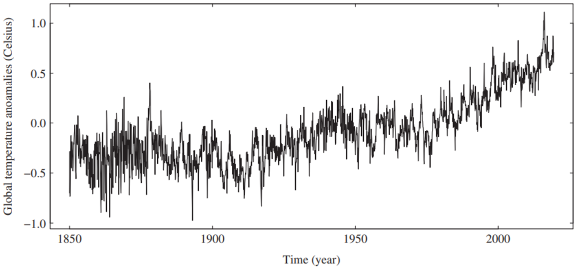


Fig. 2. Time series of monthly global temperature anomalies in degrees Celsius from January 1850 to June 2019.

# real data

Table 1. *High-quantile regression estimators for the cubic coefficient and their associated p-values for testing a zero null hypothesis against a two-sided alternative*

| | 95% quantile | | 99% quantile | |
|---|---|---|---|---|
| Tail dependence | Adjusted | Ignored | Adjusted | Ignored |
| Cubic coefficient estimate | 1.211 | 1.211 | 3.421 | 3.421 |
| *p*-value | 0.270 | 0.008 | 0.002 | 0.000 |