Functional Data Analysis for Generalized Quantile Regression

Mengmeng Guo
Lan Zhou
Wolfgang Karl Härdle
Jianhua Huang

Ladislaus von Bortkiewicz Chair
of Statistics Humboldt-Universität
zu Berlin
Department of Statistics Texas
A&M Univeristy
lvb.wiwi.hu-berlin.de
www.stat.tamu.edu

Generalized Quantile Regression (GQR)

- Quantiles and Expectiles are generalized quantiles, Jones (1994).
- Capture the tail behaviour of conditional distributions.
- Applications in finance, weather, demography, $\cdots$
- Some applications involve MANY GQR curves.

Data

High dimensional and complex data in space and time

- Weather: temperature, rainfall, solar activity
- Electricity: futures and options with different time to maturity
- Medicine: gene expression data

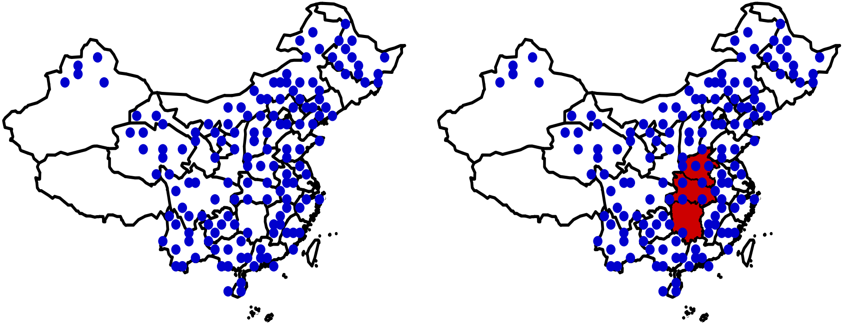


Figure 1: Weather Stations in China

Statistical Challenges

- Traditional: estimate GQR individually
- Directly: estimate GQR jointly

- common structure neglected
- too many parameters, curse of dimensionality

Functional Data Analysis (FDA)

- a tool to capture random curves
- consider dependencies between individuals
- FPCA a tool to reduce dimensionality
- interpretation of factors
- apply "FPCA" and least asymmetric weighted squares (LAWS)


I THOUGHT IT WAS THE CORN THAT SHOULD BE KNEE-HIGH BY JULY.
WISCONSIN STATE JOURNAL 2010 ©

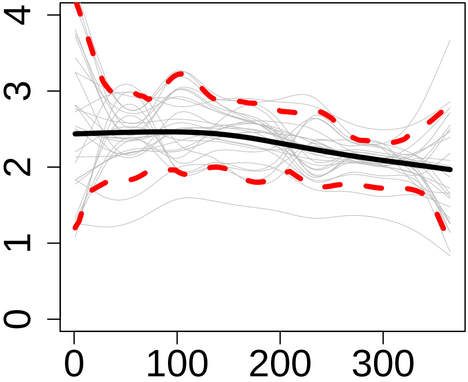
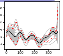


Figure 2: Estimated 95% expectile curves for the volatility of temperature of 30 cities in Germany from 1995-2007.

Go to details



Weather Derivatives

Temperature indices: Cumulative Averages (CAT) over $[\tau_1, \tau_2]$:

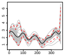
$$CAT(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} T_u du,$$

where $T_u = (T_{u,max} + T_{u,min})/2$.
A CAT temperature future under the non-arbitrage pricing setting:

$$\begin{aligned} F_{CAT(t,\tau_1,\tau_2)} &= \mathsf{E}^{Q_\lambda}\left[\int_{\tau_1}^{\tau_2} T_u du | \mathcal{F}_t\right] \\ &= \int_{\tau_1}^{\tau_2} \Lambda_u du + \mathbf{a}_{t,\tau_1,\tau_2}\mathbf{X}_t + \int_t^{\tau_1} \lambda_u \sigma_u \mathbf{a}_{t,\tau_1,\tau_2}\mathbf{e}_L du \\ &+ \int_{\tau_1}^{\tau_2} \lambda_u \sigma_u \mathbf{e}_1^\top \mathbf{A}^{-1}\left[\exp\{\mathbf{A}(\tau_2 - u)\} - I_L\right]\mathbf{e}_L du \end{aligned} \quad (1)$$

Outline

1. Motivation ✓
2. Generalized Quantile Estimation
3. FDA for GQR
4. Simulation
5. Application
6. Conclusion



Quantile and Expectile

Quantile

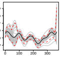
$$F(l) = \int_{-\infty}^{l} dF(y) = \tau$$

$$l = F^{-1}(\tau)$$

Expectile

$$G(l) = \frac{\int_{-\infty}^{l} |y - l| \, dF(y)}{\int_{-\infty}^{\infty} |y - l| \, dF(y)} = \tau$$

$$l = G^{-1}(\tau)$$



Loss Function

Loss function:

$$L(y,\theta) = |y-\theta|^{\alpha} \tag{2}$$

Asymmetric loss function for generalized quantiles:

$$\rho_{\tau}(u) = |\,\mathbf{I}(u \leq 0) - \tau||u|^{\alpha}, \qquad \tau \in (0,1) \tag{3}$$

with $\alpha \in \{1,2\}$ and $u = y - \theta$.

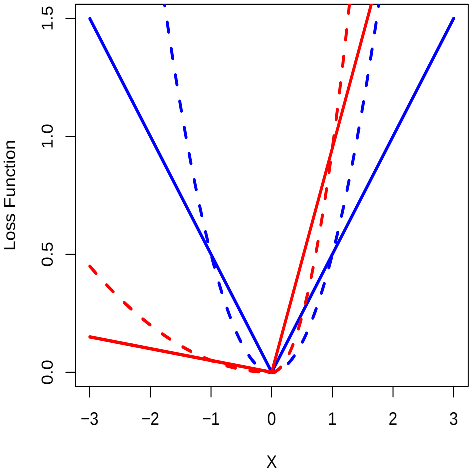


Figure 3: Loss functions for $\tau = 0.9$ (red); $\tau = 0.5$ (blue); $\alpha = 1$ (solid line); $\alpha = 2$ (dashed line).

Weight

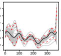
$$w_\alpha(u) = |\mathbf{I}(u \le 0) - \tau||u|^{(\alpha-2)} \tag{4}$$

Minimum contrast approach:

$$\begin{aligned} l_\tau &= \arg\min_\theta \; \mathsf{E}\{\rho_\tau(Y-\theta)\} \\ &= \arg\min_\theta \; \mathsf{E}\, w_\alpha(Y-\theta)|Y-\theta|^2 \end{aligned}$$

Generalized quantile regression curve:

$$\begin{aligned} l_\tau(t) &= \arg\min_\theta \; \mathsf{E}\{\rho_\tau(Y-\theta)|X=t\} \\ &= \arg\min_\theta \; \mathsf{E}\{w_\alpha(Y-\theta)|Y-\theta|^2|X=t\} \end{aligned}$$



Estimation Method

- Kernel Smoothing
  - Quantile: Fan et.al (1994)
  - Expectile: Zhang (1994)
- Penalized Spline Smoothing
  - Quantile: Koenker et.al (1994)
  - Expectile: Schnabel and Eilers (2009)

GQR can be estimated by LAWS.

Single Curve Estimation

Rewrite as regression pb:

$$Y_t = l(t) + \varepsilon_t \tag{5}$$

where $F^{-1}_{\varepsilon|t}(\tau) = 0$ and $G^{-1}_{\varepsilon|t}(\tau) = 0$.
Approximate $l(\cdot)$ by a B-spline basis:

$$l(t) = b(t)^\top \theta_\mu \tag{6}$$

where $b(t) = \{b_1(t), \cdots, b_q(t)\}^\top$ is a vector of cubic B-spline basis and θ_μ is a vector with dimension q.

Estimation

Employ a roughness penalty:

$$
\begin{aligned}
S(\theta_\mu) &= \sum_{t=1}^{T} w_t \{Y_t - b(t)^\top \theta_\mu\}^2 \\
&\quad + \lambda \{\theta_\mu^\top \int \ddot{b}(t) \ddot{b}(t)^\top dt \; \theta_\mu\}
\end{aligned} \tag{7}
$$

where $Y = (Y_1, Y_2, \cdots, Y_T)^\top$, $\ddot{b}(t) = \frac{\partial^2 b(t)}{\partial t^2}$ and $w_t = w_\alpha\{Y_t - l(t)\}$ ($l(t)$ known).

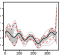
Estimation

The generalized quantile curve:

$$\begin{aligned}\widehat{\theta}_\mu &= \arg\min_{\theta_\mu} S(\theta_\mu)\\ &= \{B^\top W B + \lambda \int \ddot{b}(t)\ddot{b}(t)^\top dt\}^{-1}(B^\top W Y)\end{aligned}$$

$B = \{b(t)\}_{t=1}^{T}$ is the spline basis matrix with dimension $T \times q$, and $W = \mathrm{diag}\{w_t\}$ defined in (4):

$$\widehat{l}(t) = b(t)\widehat{\theta}_\mu \tag{8}$$



Regression Model

$$Y_{ij} = l_i(t_{ij}) + \varepsilon_{ij} \tag{9}$$

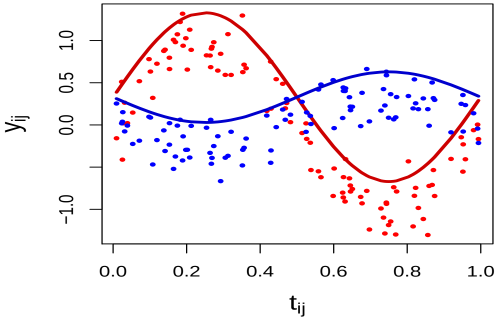


Figure 4: Data design with $\tau = 0.95$. `design`

Mixed effect Model

Observe $i = 1, \cdots, N$ individual curves:

$$l_i(t) = \mu(t) + v_i(t) \tag{10}$$

- $\mu(t)$ common shape
- $v_i(t)$ departure from $\mu(t)$.

Approximate via

$$l_{ij} = l_i(t_{ij}) = b(t_{ij})^\top \theta_\mu + b(t_{ij})^\top \gamma_{ij} \tag{11}$$

where $i = 1, \cdots, N$ and $j = 1, \cdots, T_i$.

- Too many parameters to estimate.
- Very volatile for sparse data, James et.al (2000).

Reduced Model

▸ Mercer's Lemma
▸ Karhunen-Loève Theorem

$$l_i(t) = \mu(t) + \sum_{k=1}^{K} f_k(t)^\top \alpha_{ik} \tag{12}$$

- K the number of factors and f_k k-th factor:

$$f(t) = \{f_1(t), \cdots, f_K(t)\}^\top$$

- $\alpha_i = (\alpha_{i1}, \cdots, \alpha_{iK})^\top$ random scores.

Representation of μ and f:

$$\begin{aligned} \mu(t) &= b(t)^\top \theta_\mu \\ f(t)^\top &= b(t)^\top \Theta_f \end{aligned}$$

where $\theta_\mu \in R^q$ and Θ_f with dimension $q \times K$.

Reduced Model

Rewrite (12)

$$l_{ij} = l_i(t_{ij}) = b(t_{ij})^\top \theta_\mu + b(t_{ij})^\top \Theta_f \alpha_i \tag{13}$$

With $L_i = \{l_i(t_1), \cdots, l_i(T_i)\}^\top$, $B_i = \{b(t_1), \cdots, b(T_i)\}^\top$, the GQR curves:

$$L_i = B_i \theta_\mu + B_i \Theta_f \alpha_i \tag{14}$$

Then the model reads:

$$Y_i = L_i + \varepsilon_i = B_i \theta_\mu + B_i \Theta_f \alpha_i + \varepsilon_i \tag{15}$$

with Y_i is $T_i \times 1$ and α_i is $K \times 1$.

Constraints

Orthogonality requirements of the factors:

$$\int f(t)f(t)^\top dt = \Theta_f^\top \int b(t)^\top b(t)dt \; \Theta_f = \mathrm{I}_K$$

That is to say

$$\begin{aligned} \Theta_f^\top \Theta_f &= \mathrm{I}_K \\ \int b(t)^\top b(t)dt &= \mathrm{I}_q \end{aligned}$$

“Empirical" Loss Function

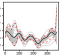
For the GQR regression:

$$S = \sum_{i=1}^{N} \sum_{j=1}^{T_i} w_{ij} \{ Y_{ij} - b(t_j)^\top \theta_\mu - b(t_j)^\top \Theta_f \alpha_i \}^2 \qquad (16)$$

Roughness penalty:

$$\begin{aligned} M_\mu &= \theta_\mu^\top \int \ddot{b}(t) \ddot{b}(t)^\top dt \ \theta_\mu \\ M_f &= \sum_{k=1}^{K} \theta_{kf}^\top \int \ddot{b}(t) \ddot{b}(t)^\top dt \ \theta_{kf} \end{aligned}$$

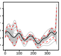
And $w_{ij} = w_\alpha(Y_{ij} - l_{ij})$, where l_{ij} defined in (13).



LAWS

$$
\begin{aligned}
S^* &= S + \lambda_\mu M_\mu + \lambda_f M_f \\
&= \sum_{i=1}^{N} (Y_i - B_i\theta_\mu - B_i\Theta_f\alpha_i)^\top W_i (Y_i - B_i\theta_\mu - B_i\Theta_f\alpha_i) \\
&\quad + \lambda_\mu \{\theta_\mu^\top \int \ddot{b}(t)\ddot{b}(t)^\top dt\ \theta_\mu\} \\
&\quad + \lambda_f \{\sum_{k=1}^{K} \theta_{f,k}^\top \int \ddot{b}(t)\ddot{b}(t)^\top dt\ \theta_{f,k}\} \qquad (17)
\end{aligned}
$$

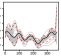
where $\theta_{f,k}$ is the k-th column in Θ_f.



Solutions

Minimizing S^*:

$$
\begin{aligned}
\widehat{\theta}_\mu &= \left\{ \sum_{i=1}^{N} B_i^\top W_i B_i + \lambda_\mu \int \ddot{b}(t)\ddot{b}(t)^\top dt \right\}^{-1} \\
&\quad \left\{ \sum_{i=1}^{N} B_i^\top W_i (Y_i - B_i \widehat{\Theta}_f \widehat{\alpha}_i) \right\} \\
\widehat{\theta}_{f,j} &= \left\{ \sum_{i=1}^{N} \widehat{\alpha}_{ij}^2 B_i^\top W_i B_i + \lambda_f \int \ddot{b}(t)\ddot{b}(t)^\top dt \right\}^{-1} \\
&\quad \left\{ \sum_{i=1}^{N} \widehat{\alpha}_{ij} B_i^\top W_i (Y_i - B_i \widehat{\theta}_\mu - B_i Q_{ij}) \right\} \qquad (18)
\end{aligned}
$$



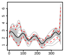
$$\widehat{\alpha}_i = \left\{ \widehat{\Theta}_f^\top B_i^\top W_i B_i \widehat{\Theta}_f \right\}^{-1} \left\{ \widehat{\Theta}_f^\top B_i^\top W_i (Y_i - B_i \widehat{\theta}_\mu) \right\} \quad (19)$$

Where

$$Q_{ij} = \sum_{k \neq j} \hat{\theta}_{f,k} \hat{\alpha}_{ik}$$

and $i = 1, \cdots, N$, $j = 1, \cdots, K$.

- initial values ▸ Details
- updated procedure ▸ Details



Auxiliary Parameters

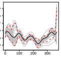
- Number of knots is not crucial, James et.al (2000)
- Use 5-fold cross validation (CV) to choose the number of factors and the penalty parameters

$$CV(K, \lambda_\mu, \lambda_f) = \frac{1}{5} \sum_{i=N-(m-1)\times 5}^{N-m\times 5} \sum_{j=1}^{T_i} \widehat{w}_{ij} |Y_{ij} - \widehat{l}_{ij}|^2 \qquad (20)$$

where $m = 1, 2, \cdots, [N/5]$, $\widehat{w}_{ij} = w_\alpha(Y_{ij} - \widehat{l}_{ij})$ and

$$\widehat{l}_{ij} = b(t_{ij})^\top \widehat{\theta}_\mu + b(t_{ij})^\top \widehat{\Theta}_f \widehat{\alpha}_i$$

.



Simulation

$$Y_{ij} = \mu(t_j) + f_1(t_j)\alpha_{1i} + f_2(t_j)\alpha_{2i} + e_{ij} \tag{21}$$

with $i = 1, \cdots, N$, $j = 1, \cdots, T_i$ and t_j is equal distanced on $[0, 1]$.

The common shape curve and factor functions:

$$\begin{aligned} \mu(t) &= 1 + t + \exp\{-(t-0.6)^2/0.05\} \\ f_1(t) &= \sin(2\pi t)/\sqrt{0.5} \\ f_2(t) &= \cos(2\pi t)/\sqrt{0.5} \end{aligned}$$

where $\alpha_{1i} \sim \mathrm{N}(0, 36)$, $\alpha_{2i} \sim \mathrm{N}(0, 9)$.

Scenarios

- $e_{ij} \sim N(0, 0.5)$
- $e_{ij} \sim N(0, \mu(t) \times 0.5)$
- $e_{ij} \sim t(5)$

- small sample: $N = 20$, $T = T_i = 100$
- large sample: $N = 40$, $T = T_i = 150$

Theoretical τ quantile and expectile for individual i:

$$l_{ij} = \mu(t_j) + f_1(t_j)\alpha_{1i} + f_2(t_j)\alpha_{2i} + \varepsilon_{ij}$$

where ε_{ij} represents the corresponding theoretical τ-th quantile and expectile of the distribution of e_{ij} ($\varepsilon_{ij} = e_{ij} + \sqrt{0.5} \cdot \Phi^{-1}(\tau)$).

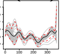
Estimators

- The individual curve:

$$\begin{aligned} l_i &= \mu + \sum_{k=1}^{K} f_k \alpha_{ik} \\ \widehat{l}_{i,fp} &= B_i \widehat{\theta}_\mu + B_i \widehat{\Theta}_f \widehat{\alpha}_i \\ \widehat{l}_{i,in} &: \text{Single curve, see (8)} \end{aligned}$$

- The mean curve:

$$\begin{aligned} m &= \mu(t) + e_\tau \\ m_{fp} &= \frac{1}{N} \sum_{i=1}^{N} B_i \widehat{\theta}_\mu \\ m_{in} &= \frac{1}{N} \sum_{i=1}^{N} \widehat{l}_{i,in} \end{aligned} \tag{22}$$



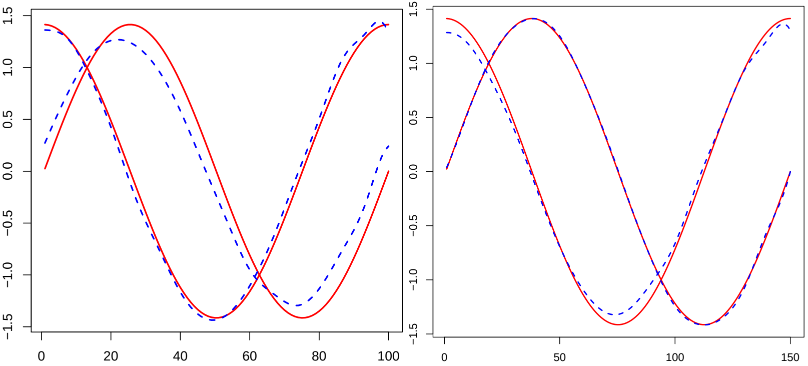


Figure 5: The estimated factors (dashed blue) compared with the true ones (solid red) for the 95% expectile with the error term normally distributed. The left part is for $N = 20$, $T = 100$. The right one is for $N = 40$, $T = 150$.

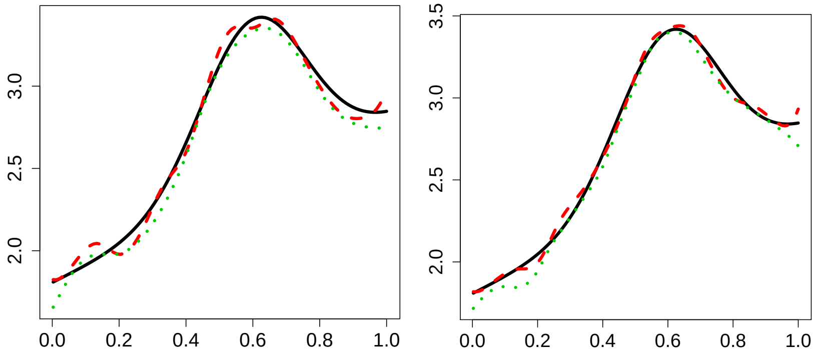


Figure 6: The estimated common shape compared with the true mean for the 95% expectile with the error term normally distributed. The left part is for $N = 20$, $T = 100$. The right one is for $N = 40$, $T = 150$.

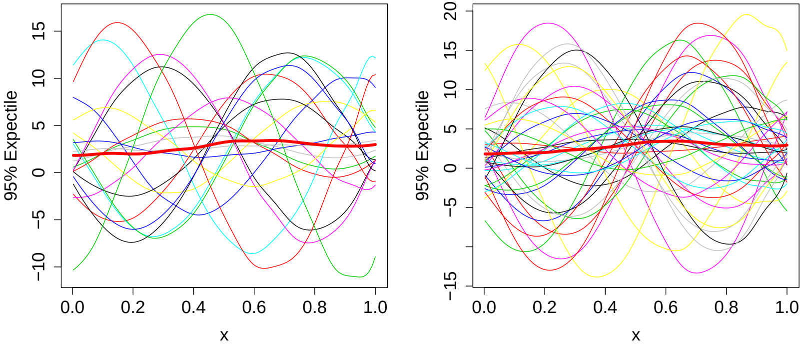


Figure 7: The estimated 95% expectile curves. The thick red line is the common mean curve with the error term normally distributed. The left part is for $N = 20$, $T = 100$. The right one is for $N = 40$, $T = 150$.

	Individual		*Mean*	
Sample Size	*FDA*	*Single*	*FDA*	*Single*
$N = 20, T = 100$	0.0469	0.0816	0.0072	0.0093
$N = 40, T = 150$	0.0208	0.0709	0.0028	0.0063
$N = 20, T = 100$	0.1571	0.2957	0.0272	0.0377
$N = 40, T = 150$	0.1002	0.2197	0.0118	0.0172
$N = 20, T = 100$	0.2859	0.5194	0.0454	0.0556
$N = 40, T = 150$	0.1531	0.4087	0.0181	0.0242

Table 1: The mean squared errors (MSE) of the FDA and the single curve estimation for expectile curves with error term is normally distributed with mean 0 and variance 0.5 (Top), with variance $\mu(t) \times 0.5$ (Middle) and $t(5)$ distribution (Bottom).

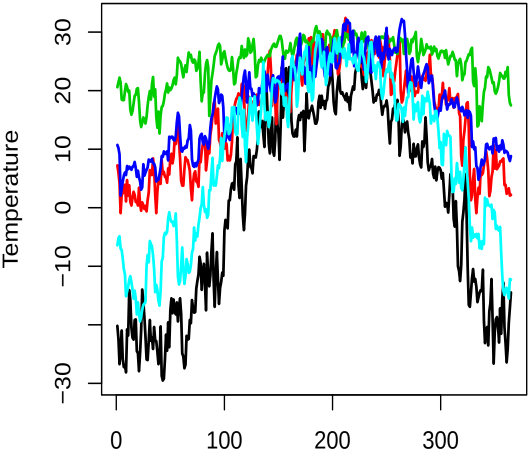


Figure 8: time series plot of 5 selected weather stations (south, north, east, west and middle)from 150 weather stations in China

Data

- Daily temperature data in 2010 in 150 weather stations in China
- B cubic spline
- The number of knots $q = 16$

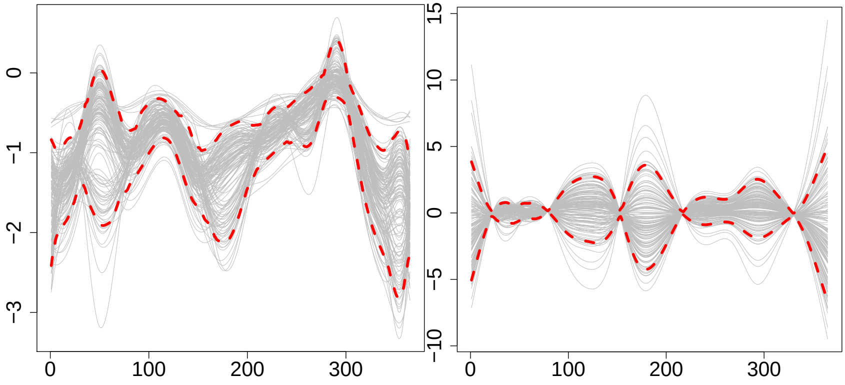


Figure 9: 25% (left) and 50% (right) estimated expectile curves of the temperature variations for 150 weather stations in China in 2010.

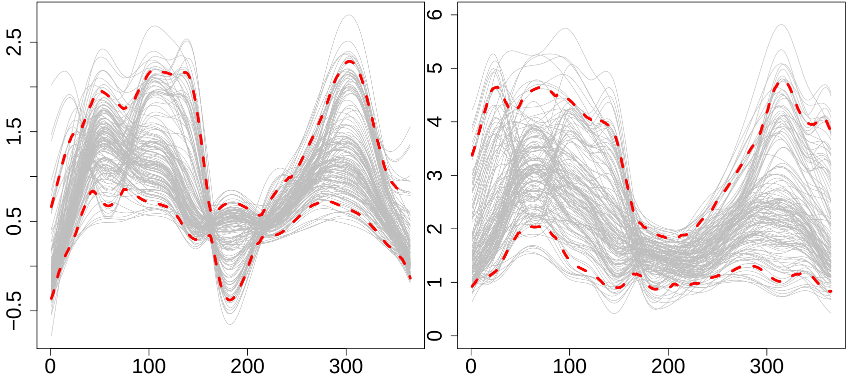


Figure 10: 75% (left) and 95% (right) estimated expectile curves of the temperature variations for 150 weather stations in China in 2010.

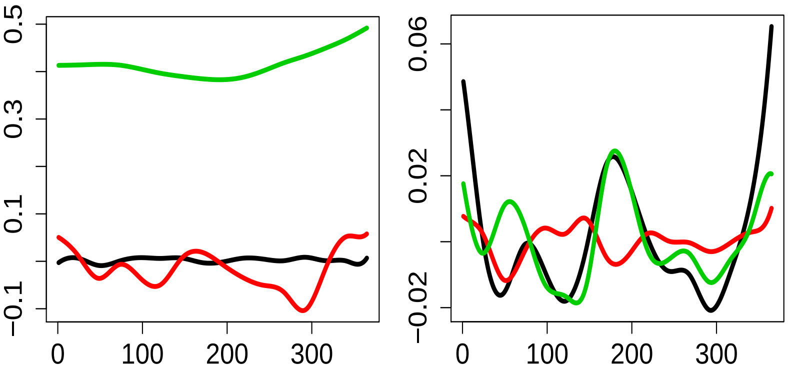


Figure 11: The estimated three factors for 25% (left) and 50% (right) expectile curves of the temperature variation. The black one is the first eigenfunction, the red one is the second and the green one represents the third factor.

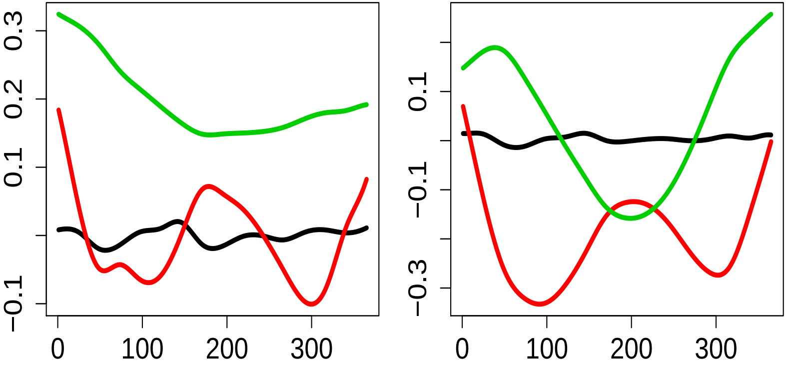


Figure 12: The estimated three factors for 75% (left) and 95% (right) expectile curves of the temperature variation. The black one is the first factor f_1, the red one is the second f_2 and the green one represents the third factor f_3.

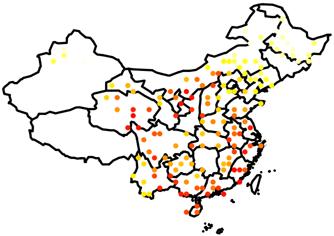


Figure 13: The estimated first random scores α_1 for 25%, 50%, 75% and 95% expectile curves of the temperature variation.

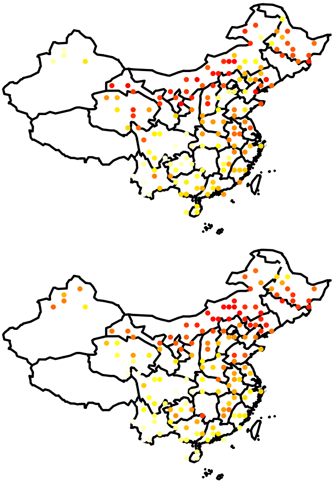


Figure 14: The estimated second random scores α_2 for 25%, 50%, 75% and 95% expectile curves of the temperature variation.

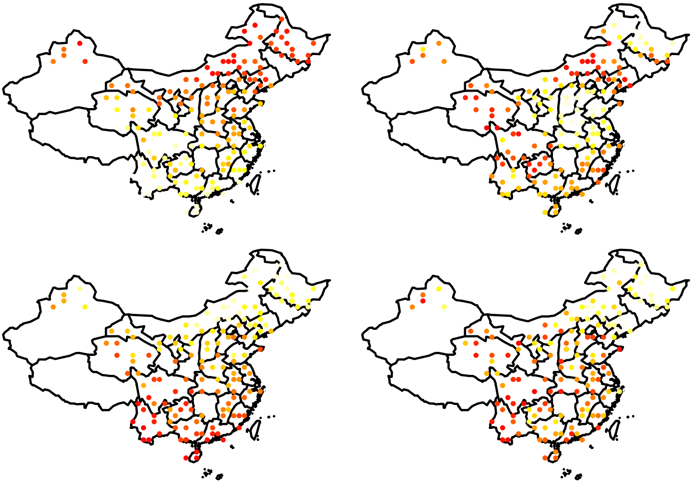


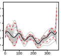
Figure 15: The estimated third random scores α_3 for 25%, 50%, 75% and 95% expectile curves of the temperature variation.

	Min	*Max*	*Median*	*Mean*	*SD*
$\tau = 0.25$	-68.48	168.30	-14.09	0.00	46.27
$\tau = 0.5$	-129.50	199.50	-18.02	0.00	52.00
$\tau = 0.75$	-22.64	61.20	-8.86	0.00	19.94
$\tau = 0.95$	-60.93	142.60	-12.64	0.00	44.56

Table 2: Statistical Summary of α_1

Conclusion

- Dimension Reduction technique applied to a nonlinear object.
- Provides a novel way to estimate several generalized quantile curves simultaneously.
- Outperforms the single curve estimation, especially when the data is very volatile.
- Pricing weather derivatives more precisely can be possible.



Reference

J. Fan and T. C. Hu and Y. K. Troung
Robust nonparametric function estimation
Scandinavian Journal of Statistics, 21:433-446, 1994.

M. Guo and W. Härdle
Simulateous Confidence Bands for Expectile Functions
Advances in Statistical Analysis, 2011,
DOI:10.1007/s10182-011-0182-1.

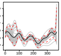
G. James and T. Hastie and C. Sugar
Principal Component Models for Sparse Functinal Data
Biometrika, 87:587-602, 2000.

M. Jones
Expectiles and M-quantiles are Quantiles
Statistics & Probability Letters, 20:149-153, 1994.

R. Koenker and P. Ng and S. Portnoy
Quantile Smoothing Splines
Biometrika, 81(4):673-680, 1994.

B. Zhang
Nonparametric Expectile Regression
Nonparametric Statistics, 3:255-275, 1994

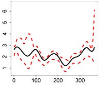
L. Zhou and J. Huang and R. Carroll
Joint Modelling of Paired Sparse Functional Data Using Principal Components
Biometrika, 95(3):601-619, 2008.



Functional Data Analysis for Generalized Quantile Regression

Mengmeng Guo
Lan Zhou
Wolfgang Karl Härdle
Jianhua Huang

Ladislaus von Bortkiewicz Chair of Statistics Humboldt-Universität zu Berlin
Department of Statistics Texas A&M Univerisity
lvb.wiwi.hu-berlin.de
www.stat.tamu.edu



Volatility of Temperature

Return

- The temperature T_{it} on day t for city i:

$$T_{it} = X_{it} + \Lambda_{it}$$

- The seasonal effect Λ_{it}:

$$\Lambda_{it} = a_i + b_i t + \sum_{m=1}^{M} c_{im} \cos\{\frac{2\pi(t - d_{im})}{365}\}$$

- X_{it} follows an $AR(p_i)$ process:

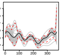
$$X_{it} = \sum_{j=1}^{p_i} \beta_{ij} X_{i,t-j} + \varepsilon_{it} \tag{23}$$

$$\widehat{\varepsilon}_{it} = X_{it} - \sum_{j=1}^{p_i} \hat{\beta}_{ij} X_{i,t-j}$$

Initial Values

▸ Return

1. Estimate N single curves $\widehat{l}_i$ individually.
2. Linear regression for $\widehat{\theta}_{\mu 0}$: $\widehat{l}_i = B_i \theta_\mu + \varepsilon_i$
3. Calculate $\widetilde{l}_{i0} = \widehat{l}_i - B_i \widehat{\theta}_{\mu 0}$, and $\widehat{\Gamma}_0 = (\widehat{\Gamma}_{10}, \cdots, \widehat{\Gamma}_{N0})$.
$$\widetilde{l}_{i0} = B_i \Gamma_i + \varepsilon_i$$
4. Apply SVD to decompose $\widehat{\Gamma}_{i0}$:
$$\widehat{\Gamma}_{i0} = UDV^\top = \Theta_{f0} \alpha_{i0}$$
5. Choose the first K factors from U as $\widehat{\Theta}_{f0}$, and regress $\widehat{\Gamma}_{i0}$ on $\widehat{\Theta}_{f0}$ to get $\widehat{\alpha}_{i0}$:
$$\widehat{\Gamma}_{i0} = \widehat{\Theta}_{f0}(\alpha_{i1}, \cdots, \alpha_{iK}) + \varepsilon_i \tag{24}$$



Update Procedure

▸ Return

1. Plug $\widehat{\Theta}_{f0}$ and $\widehat{\alpha}_{i0}$ into (18) to update θ_μ, and get $\widehat{\theta}_{\mu 1}$.
2. Plugging $\hat{\theta}_{\mu 1}$ and $\widehat{\alpha}_{i0}$ into the second equation of (18) gives $\widehat{\Theta}_{f1}$.
3. Given $\widehat{\theta}_{\mu 1}$ and $\widehat{\Theta}_{f1}$, estimate $\widehat{\alpha}_i$.
4. Recalculate the weight matrix:

$$w_{ij}^{'} = w_\alpha(Y_{ij} - \widehat{l_{ij}})$$

   where $\widehat{l_{ij}}$ is the j-th element in $\widehat{l_i} = B_i \widehat{\theta}_{\mu 1} + B_i \widehat{\Theta}_{f1} \widehat{\alpha}_i$
5. Repeat step (1) to (4) until the solutions converge.

Mercer's Lemma

The covariance operator K

$$K(s,t) = \mathrm{Cov}\{I(s), I(t)\}, \mathrm{E}\{I(t)\} = \mu(t), s, t \in \mathcal{T} \tag{25}$$

There exists an orthonormal sequence (ψ_j) and non-increasing and non-negative sequence (κ_j),

$$\begin{aligned} (K\psi_j)(s) &= \kappa_j \psi_j(s) \\ K(s,t) &= \sum_{j=1}^{\infty} \kappa_j \psi_j(s) \psi_j(t) \\ \sum_{j=1}^{\infty} \kappa_j &= \int_{\mathrm{I}} K(t,t) dt < \infty \end{aligned} \tag{26}$$

Return

Karhunen-Loève Theorem

Under assumptions of Mercer's lemma

$$l(t) = \mu(t) + \sum_{j=1}^{\infty} \sqrt{\kappa_j}\xi_j\psi_j(t) \tag{27}$$

where $\xi_j \stackrel{\text{def}}{=} \frac{1}{\sqrt{\kappa_j}} \int \; l(t)\psi_j(s)ds$, and $\mathsf{E}(\xi_j) = 0$

$$\mathsf{E}(\xi_j\xi_k) = \delta_{j,k} \qquad j,k \in \mathbb{N}$$

and $\delta_{j,k}$ is the Kronecker delta.

▸ Return