Quantile Curves without Crossing

Victor Chernozhukov     Iván Fernández-Val     Alfred Galichon

MIT                  Boston University         Ecole Polytechnique

Déjeuner-Séminaire d’Economie
Ecole polytechnique, November 12 2007
Aim of the talk

- Present the methodology and applications of **Quantile Regression**
Aim of the talk

- Present the methodology and applications of **Quantile Regression**
- Identify and correct a **common problem** for several estimation procedures Quantile Regression
Aim of the talk

- Present the methodology and applications of Quantile Regression
- Identify and correct a common problem for several estimation procedures Quantile Regression
- Study the impact of the correction on the estimator of the measure
This talk:

1. Quantile Regression and Two applications
   - The Value-at-Risk
   - Engle curves

2. Rearranging Quantile Curves
   - The crossing problem
   - The rearrangement operation
   - Overview of the literature

3. Properties and illustrations
   - Graphical properties
   - Approximation properties
   - Large-sample properties
The Value-at-Risk

**Aim:** measure & manage risk of portfolio’s contingent loss $Y$. 
Aim: measure & manage risk of portfolio’s contingent loss $Y$.

- $VaR_\alpha(Y) =$ smallest capital amount to cover losses in $\alpha\%$ cases...
The Value-at-Risk

**Aim:** measure & manage risk of portfolio’s contingent loss $Y$.

- $\text{VaR}_\alpha(Y) =$ smallest capital amount to cover losses in $\alpha\%$ cases...

- is robust to tail behaviour (eg. more than variance)
The Value-at-Risk

**Aim:** measure & manage risk of portfolio’s contingent loss $Y$.

- $\text{VaR}_\alpha(Y) =$ smallest capital amount to cover losses in $\alpha\%$ cases...

- is robust to tail behaviour (eg. more than variance)

- has become a market standard for market risk measurement (Basle II 1st pillar)
The Value-at-Risk

**Aim:** measure & manage risk of portfolio’s contingent loss $Y$.

- $VaR_\alpha(Y) =$ smallest capital amount to cover losses in $\alpha\%$ cases...
- is robust to tail behaviour (eg. more than variance)
- has become a market standard for market risk measurement (Basle II 1st pillar)
- is however criticized among both practitioners and academics
Many authors have pointed out the shortfalls of VaR as a management tool.
Many authors have pointed out the shortfalls of VaR as a management tool.

- Desirable property: **subadditivity** “Merging does not make risk appear larger” (Artzner et al., 1999)
Many authors have pointed out the shortfalls of VaR as a management tool.

- Desirable property: **subadditivity** “Merging does not make risk appear larger” (Artzner et al., 1999)
  ⇒ Fails to be satisfied by VaR.
Many authors have pointed out the shortfalls of VaR as a management tool.

- Desirable property: **subadditivity** “Merging does not make risk appear larger” (Artzner et al., 1999)
  \[ \implies \] Fails to be satisfied by VaR.

- Alternative “coherent” measures have been proposed.
  - **Expected Shortfall**: average loss beyond given level
Many authors have pointed out the shortfalls of VaR as a management tool.

- Desirable property: **subadditivity** “Merging does not make risk appear larger” (Artzner et al., 1999)
  \[ \Rightarrow \text{Fails to be satisfied by VaR.} \]

- Alternative “coherent” measures have been proposed.
  - **Expected Shortfall:** average loss beyond given level
  - **distortion measures:** weighted loss average; higher weights toward higher losses.
Recall the definition of a quantile function.
Recall the definition of a quantile function.
Let $F_Y(y) = Pr(Y \leq y)$: distribution function of $Y$, and $Q_Y(u) = F_Y^{-1}(u)$ quantile function.
VaR and Quantile estimation

Recall the definition of a quantile function.
Let \( F_Y(y) = \Pr(Y \leq y) \): distribution function of \( Y \), and \( Q_Y(u) = F_Y^{-1}(u) \) quantile function.

- The VaR is is precisely \( \text{VaR}_\alpha(Y) = Q_Y(\alpha) \).
Recall the definition of a quantile function.
Let $F_Y(y) = Pr(Y \leq y)$: distribution function of $Y$, and
$Q_Y(u) = F_Y^{-1}(u)$ quantile function.

- The VaR is precisely $VaR_\alpha(Y) = Q_Y(\alpha)$.
- Distorsion measures can be written $\rho(Y) = \int_0^1 \varphi(u) Q_Y(u) du$, where $\varphi$ is increasing.
Recall the definition of a quantile function. Let $F_Y(y) = Pr(Y \leq y)$: distribution function of $Y$, and $Q_Y(u) = F_Y^{-1}(u)$ quantile function.

- The VaR is precisely $\text{VaR}_\alpha(Y) = Q_Y(\alpha)$.
- Distorsion measures can be written $\rho(Y) = \int_0^1 \varphi(u)Q_Y(u)du$, where $\varphi$ is increasing.

**Example.** Expected Shortfall: $\varphi(u) = 1\{u \geq \alpha\}/(1 - \alpha)$. 
Recall the definition of a quantile function. Let $F_Y(y) = Pr(Y \leq y)$: distribution function of $Y$, and $Q_Y(u) = F_Y^{-1}(u)$ quantile function.

- The VaR is is precisely $VaR_\alpha(Y) = Q_Y(\alpha)$.
- Distortion measures can be written $\rho(Y) = \int_0^1 \varphi(u) Q_Y(u) du$, where $\varphi$ is increasing.

**Example.** Expected Shortfall: $\varphi(u) = 1\{u \geq \alpha\}/(1 - \alpha)$.

Thus estimation of VaR/distortion measure requires estimation of quantile function.
Engel Curves

- $Y$ response variable, $X$ regressor, the $u$-th quantile of $Y$ given $X = x$

$$Q_0(u|x) = \inf\{y : F(y|x) \geq u\}.$$  

- QR estimates a linear approximation to the conditional quantile

$$Q(u|x) = x'\beta(u)$$

- QR fits for different quantiles provide a description of the entire conditional distribution

Example: Buchinsky (1994) uses QR to describe the evolution of the wage distribution in the U.S.

- Here, $y =$ food expenditure ; $x =$ household income.
Engel Curves by Quantile Regression

Chernozhukov, Fernández-Val, Galichon

Rearranging VaR estimators
Engel Curves by Quantile Regression

A. Income = 452 (5% quantile)

B. Income = 884 (Median)

Chernozhukov, Fernández-Val, Galichon

Rearranging VaR estimators
Quantile regression

Given covariate $X$ (information at period $t$), estimate QR model

$$Q_{Y|X} (u|x) = x'\beta (u).$$
Quantile regression

Given covariate $X$ (information at period $t$), estimate QR model

$$Q_{Y|X} (u|x) = x' \beta (u).$$

estimate using

$$\hat{\beta} (u) = \arg \min_{\beta \in \mathbb{R}^d} \sum_{k=1}^{n} u (Y_k - X_k' \beta)^{+} + (1 - u) (Y_k - X_k' \beta)^{-}$$
Quantile regression

Given covariate $X$ (information at period $t$), estimate QR model

$$Q_{Y|X}(u|x) = x' \beta(u).$$

- estimate using

$$\hat{\beta}(u) = \arg \min_{\beta \in \mathbb{R}^d} \sum_{k=1}^{n} u (Y_k - X'_k \beta)^+ + (1 - u) (Y_k - X'_k \beta)^-$$

- Autoregressive case- the covariate $X$ captures past information. Several models: Quantile Autoregression, CaViaR, Dynamic Quantile...
The crossing problem

In the QR procedure, nothing ensures that \( \hat{Q}_{Y \mid X}(u \mid x) = x' \hat{\beta}(u) \) be increasing in \( u \).
The crossing problem

In the QR procedure, nothing ensures that $\hat{Q}_{Y|X}(u|x) = x'\hat{\beta}(u)$ be increasing in $u$.

- In fact it may be non-monotonic if
  - the QR model is misspecified, or
  - the sample size is small
The crossing problem

In the QR procedure, nothing ensures that \( \hat{Q}_{Y|X}(u|x) = x' \hat{\beta}(u) \) be increasing in \( u \).

- In fact it may be non-monotonic if
  - the QR model is misspecified, or
  - the sample size is small
- VaR context: a higher confidence level would require less capital!
The crossing problem

In the QR procedure, nothing ensures that $\hat{Q}_{Y|X}(u|x) = x'\hat{\beta}(u)$ be increasing in $u$.

- In fact it may be non-monotonic if
  - the QR model is misspecified, or
  - the sample size is small
- VaR context: a higher confidence level would require less capital!
- can have adverse managemental effects / lack of trust for the tool...
A proposed solution

Suppose we use the (flawed) estimator \( \hat{Q}_{Y|X}(u|x) \) to simulate \( Y|X = x \).
A proposed solution

Suppose we use the (flawed) estimator $\hat{Q}_{Y|X}(u|x)$ to simulate $Y|X = x$.

- draw $U \sim U[0, 1]$ and take $Y_x := \hat{Q}(U|x)$ (bootstrap).
A proposed solution

Suppose we use the (flawed) estimator $\hat{Q}_{Y|X}(u|x)$ to simulate $Y|X = x$.

- draw $U \sim U[0, 1]$ and take $Y_x := \hat{Q}(U|x)$ (bootstrap).
- take distribution function $\hat{F}(y|x) = Pr(Y_x \leq y)$, ie.

$$\hat{F}(y|x) = \int_0^1 1\{\hat{Q}(u|x) \leq y\} du$$
A proposed solution

Suppose we use the (flawed) estimator \( \hat{Q}_{Y|X}(u|x) \) to simulate \( Y|X = x \).

- draw \( U \sim U[0, 1] \) and take \( Y_x := \hat{Q}(U|x) \) (bootstrap).
- take distribution function \( \hat{F}(y|x) = Pr(Y_x \leq y) \), ie.

\[
\hat{F}(y|x) = \int_0^1 1\{\hat{Q}(u|x) \leq y\}du
\]

- invert to recover rearranged quantile \( \hat{F}^{-1}(u|x) \)
A proposed solution

Suppose we use the (flawed) estimator $\hat{Q}_{Y|X}(u|x)$ to simulate $Y|X = x$.

- draw $U \sim U[0, 1]$ and take $Y_x := \hat{Q}(U|x)$ (bootstrap).
- take distribution function $\hat{F}(y|x) = Pr(Y_x \leq y)$, ie.

$$\hat{F}(y|x) = \int_0^1 1\{\hat{Q}(u|x) \leq y\} du$$

- invert to recover rearranged quantile $\hat{F}^{-1}(u|x)$

If the original estimator $\hat{Q}_X(u|x)$ is monotonic, then $\hat{F}^{-1}(u|x)$ coincides with it.
The rearrangement: illustration
Literature review


Analytical Properties

- As an example, take $Q(u)$ to be a non-monotone function of $u$ - slope changes sign twice in $[0, 1]$.
- Rearranged curve is monotonically increasing and coincides with $Q(u)$ for points where $Q^{-1}(y)$ is uniquely defined.
- The derivative of the rearranged curve is a proper density function, continuous at the regular values of $Q(u)$. 
Analytical Properties: Example

Chernozhukov, Fernández-Val, Galichon

Rearranging VaR estimators
Analytical Properties: Example

\[ \frac{1}{f(F^{-1}(u))} \]

\[ f(y) \]
Estimation Properties

Denote $Q_0(u|x)$: true conditional quantile curve
Denote $Q_0(u|x)$: true conditional quantile curve

- For $p \in [1, \infty]$, rearrangement inequality:

$$\int_0^1 |Q_0(u|x) - \hat{F}^{-1}(u|x)|^p du \leq \int_0^1 |Q_0(u|x) - x'\hat{\beta}(u)|^p du.$$
Estimation Properties

Denote $Q_0(u|x)$: true conditional quantile curve

- For $p \in [1, \infty]$, rearrangement inequality:

$$\int_0^1 |Q_0(u|x) - \hat{F}^{-1}(u|x)|^p du \leq \int_0^1 |Q_0(u|x) - x'\hat{\beta}(u)|^p du.$$

- This property is independent of the sample size (holds in population).
Denote \( Q_0(u|x) \): true conditional quantile curve

- For \( p \in [1, \infty] \), rearrangement inequality:

\[
\int_0^1 |Q_0(u|x) - \hat{F}^{-1}(u|x)|^p du \leq \int_0^1 |Q_0(u|x) - x'\hat{\beta}(u)|^p du.
\]

- This property is **independent** of the sample size (holds in population).
- Rearranged quantile curves have a smaller estimation error than the original curves whenever the latter are not monotone.
Illustration

$ap + bp \leq cp + dp$

Chernozhukov, Fernández-Val, Galichon

Rearranging VaR estimators
Statistical properties (large sample)

- Fix an $x$, and suppose

$$\sqrt{n}x'(\hat{\beta}(u) - \beta(u)) \Rightarrow x'G(u)$$

where $G$ is a Gaussian process. The population curve $u \rightarrow x'\beta(u)$ is assumed to be increasing.
Statistical properties (large sample)

- Fix an $x$, and suppose

$$\sqrt{n}x'(\hat{\beta}(u) - \beta(u)) \Rightarrow x' G(u)$$

where $G$ is a Gaussian process. The population curve $u \rightarrow x' \beta(u)$ is assumed to be increasing.

- Recall $\hat{F}(y|x) = \int_0^1 1\{x' \hat{\beta}(u) \leq y\} du$. One has
Fix an $x$, and suppose

$$\sqrt{n}x'(\hat{\beta}(u) - \beta(u)) \Rightarrow x' G(u)$$

where $G$ is a Gaussian process. The population curve

$u \mapsto x' \beta(u)$ is assumed to be increasing.

Recall $\hat{F}(y|x) = \int_0^1 1\{x'\hat{\beta}(u) \leq y\} du$. One has

$$\sqrt{n}(\hat{F}(y|x) - F(y|x)) \Rightarrow F'(y|x) [x' G(F(y|x))]$$
Statistical properties (large sample)

- Fix an $x$, and suppose

$$\sqrt{n}x'(\hat{\beta}(u) - \beta(u)) \Rightarrow x'G(u)$$

where $G$ is a Gaussian process. The population curve $u \rightarrow x'\beta(u)$ is assumed to be increasing.

- Recall $\hat{F}(y|x) = \int_0^1 1\{x'\hat{\beta}(u) \leq y\} du$. One has

$$\sqrt{n}(\hat{F}(y|x) - F(y|x)) \Rightarrow F'(y|x)[x'G(F(y|x))]$$

- For quantiles, one has

$$\sqrt{n}(\hat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow x'G(u) \text{ in } \ell^\infty((0, 1))$$
Fix an $x$, and suppose
\[ \sqrt{n} x'(\hat{\beta}(u) - \beta(u)) \Rightarrow x' G(u) \]

where $G$ is a Gaussian process. The population curve $u \rightarrow x' \beta(u)$ is assumed to be increasing.

Recall $\hat{F}(y|x) = \int_0^1 1\{x' \hat{\beta}(u) \leq y\} du$. One has
\[ \sqrt{n}(\hat{F}(y|x) - F(y|x)) \Rightarrow F'(y|x)[x' G(F(y|x))] \]

For quantiles, one has
\[ \sqrt{n}(\hat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow x' G(u) \text{ in } \ell^\infty((0, 1)) \]

(same asymptotic limit as for the original curve $u \rightarrow x' \hat{\beta}(u)$.)
The rearranged curve has the same asymptotic error term as the original curve...
Statistical properties, comments

- The rearranged curve has the same asymptotic error term as the original curve!
- Not incompatible with finite sample properties
Statistical properties, comments

- The rearranged curve has the same asymptotic error term as the original curve!
- Not incompatible with finite sample properties
- Convenient for testing purposes, as it does not modify the asymptotic properties of a test, while improving approximation in finite sample.
Empirical Application: Engel Curves

- We use the original Engel (1857) data to estimate the relationship between food expenditure and annual household income.
- Data set is based on 235 budget surveys of 19th century working-class Belgium households.
- Plot of quantile regression process (as a function of $u$) shows quantile-crossing for 5% percentile of income. No crossing problem for the sample median of income.
- Rearrangement procedure produces monotonically increasing curves - coincides with QR for the median of income.
Empirical Application: Engel Curves

A. Income = 394 (1% quantile)

B. Income = 452 (5% quantile)

C. Income = 884 (Median)

D. Income = 2533 (99% quantile)

Chernozhukov, Fernández-Val, Galichon

Rearranging VaR estimators
Uniform Inference: Engel Curves

- **Quantile-uniform inference** can be performed using the rearranged quantile curves.
- Next figure plots simultaneous 90% confidence intervals for the conditional quantile process of food expenditure for two different values of income, the sample median and the 1 percent sample percentile.
- Bands for QR are obtained by **bootstrap** using 500 repetitions and a grid of quantiles \{0.10, 0.11, ..., 0.90\}.
- Bands for rearranged curves are constructed assuming that estimand of QR is monotonically correct.
- Rearranged bands lie within QR bands - points towards lack of monotonicity due to small sample size.
Uniform Inference: Engel Curves

A. Income = 394 (1% quantile)

B. Income = 884 (Median)
Smoothing: Engel Curves

- Uniform bands can be constructed for smoothed quantile regression and rearranged curves.
- Previous bootstrap procedure is valid for smoothed rearranged curves even under population non monotonicity.
- Smoothed estimates with box kernel and bandwidth = 0.05.
- Almost perfect overlap of the bands - indication of population monotonicity.
- Smoothing reduces widths of the bands, but it is not enough to monotonize quantile regression curves.
Smoothing: Engel Curves

A. Income = 394 (1% quantile)

B. Income = 884 (Median)
We consider two versions of the location-scale shift model:

\[ y_i = x_i' \alpha + (x_i' \gamma) \epsilon_i, \quad Q_0(u|x_i) = x_i' (\alpha + \gamma F_{\epsilon}^{-1}(u)) \]

(1) **Linear**: \( x_i = (1, z_i) \)
(2) **Piecewise Linear**: \( x_i = (1, z_i, 1\{z_i > Med[z]\} \times z_i) \)

- Parameters calibrated to Engel application
- 1,000 Monte Carlo samples of \( n = 235 \) from a normal with same mean and variance as the residuals \( \epsilon_i = (y_i - x_i' \alpha)/(x_i' \gamma) \).
- Regressors fixed to the values of income in Engel data set.
- We estimate a linear model: \( Q(u|x_i) = x_i' \beta(u) \) for \( x_i = (1, z_i) \)
## Approximation Properties: Monte Carlo

<table>
<thead>
<tr>
<th></th>
<th>(1) Correct Specification</th>
<th>(2) Incorrect Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original</td>
<td>Rearranged</td>
</tr>
<tr>
<td>$L^1$</td>
<td>6.79</td>
<td>6.61</td>
</tr>
<tr>
<td>$L^2$</td>
<td>7.99</td>
<td>7.69</td>
</tr>
<tr>
<td>$L^3$</td>
<td>8.93</td>
<td>8.51</td>
</tr>
<tr>
<td>$L^4$</td>
<td>9.70</td>
<td>9.17</td>
</tr>
<tr>
<td>$L^\infty$</td>
<td>17.14</td>
<td>15.32</td>
</tr>
</tbody>
</table>

Each entry of the table gives a Monte Carlo average of

$$L^p(\tilde{Q}) := \left( \int |Q_0(u|x_0) - \tilde{Q}(u|x_0)|^p du \right)^{1/p}$$

for $\tilde{Q}(u|x_0) = x_0'\hat{\beta}(u)$, $\tilde{Q}(u|x_0) = \hat{F}^{-1}(u|x_0)$, and $x_0 = 452$.
Further research directions and extensions:

- Probability curves
- Demand curves
- Growth curves
- Yield curves
Thank you!
alfred.galichon@polytechnique.edu