Quantile Autoregressions: VaR, Causality, Unit Root and Cointegration

Mei-Yuan Chen
Department of Finance
National Chung Hsing University

February 25, 2013
consider that there is a random coefficient model for $y_t$ as

$$y_t = X'_t\beta(U_t) + e_t$$

$$= \beta_1(U_t) + \beta_2(U_t)x_{t2} + \cdots + \beta_k(U_t)x_{tk} + e_t,$$

where $U_t \sim i.i.d.U(0, 1)$ and $\beta_i(U_t), i = 1, \ldots, k$ are assumed to be comonotonic in $U_t$. Then

$$Q_{y_t|x_t}(\tau) = x'_t\beta_{\tau}.$$
Comonotonicity

Two random variables \((X, Y) : \Omega \rightarrow R\) are **comonotonic** if there exists a third random variable \(Z : \Omega \rightarrow R\) and increasing functions \(f\) and \(g\) such that \(X = f(Z)\) and \(Y = g(Z)\).

1. If \(X\) and \(Y\) are comonotonic they have **rank correlation one**.

2. \(F_{X+Y}^{-1}(\tau) = F_X^{-1}(\tau) + F_Y^{-1}(\tau)\).

3. \(X\) and \(Y\) are driven by the same random (uniform) variable.

Besides, for any monotone increasing function \(g\) and standard uniform random variable, \(U\), we have \(Q_{g(U)}(\tau) = g[Q_U(\tau)] = g(\tau)\).
Proof

\[ Q_{yt|x_t}(\tau) = \sum_{i=1}^{k} \beta_i(U_t) x_{ti} + Q_{\beta_1(U_t)x_{t2}}(\tau) + \cdots + Q_{\beta_k(U_t)x_{tk}}(\tau) \]

\[ = \beta_1[Q_{U_t}(\tau)] + \beta_2[Q_{U_t}(\tau)] x_{t2} + \cdots + \beta_k[Q_{U_t}(\tau)] x_{tk} \]

\[ = \beta_1(\tau) + \beta_2(\tau) x_{t2} + \cdots + \beta_k(\tau) x_{tk} \]

\[ = x_t' \beta_{\tau} \]

This result holds for all \( \tau \in (0, 1) \).
Quantile Autoregression

Koenker and Xiao (2006) suggests the quantile autoregression model by the $p$th-order autoregressive process with random coefficients:

$$y_t = \theta_0(U_t) + \theta_1(U_t)y_{t-1} + \cdots + \theta_p(U_t)y_{t-p},$$

where $\{U_t\}$ is a sequence of i.i.d. standard uniform random variables. Providing $\theta_i(U_t), i = 0, \ldots, p$ are monotone increasing in $U_t$, then

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = \theta_0(\tau) + \theta_1(\tau)y_{t-1} + \cdots + \theta_p(\tau)y_{t-p}.$$ 

This is called the QAR($p$) model.
Estimation for Quantile Autoregression

Denote $\mathbf{x}_t = (1, y_{t-1}, \ldots, y_{t-p})'$, the QAR($p$) model can be written as

$$y_t = Q_{y_t}(\tau|\mathcal{F}_{t-1}) + e_t = \mathbf{x}_t'\theta_\tau + e_t.$$

$\theta_\tau$ can be estimated by

$$\hat{\theta}_\tau = \arg \min_{\theta \in \mathbb{R}^{p+1}} \sum_{t=1}^{T} \rho_\tau(y_t - \mathbf{x}_t'\theta),$$

where $\rho_\tau(u) = u[\tau - I(u < 0)]$. Solutions, $\hat{\theta}(\tau)$, are called autoregression quantiles.
Conditional Density Estimation

Then, the conditional density of $y_t$ on $x_t$ can be estimated by the difference quotients,

$$\hat{f}_{y_t|x_t}(\tau) = \frac{\tau_i - \tau_{i-1}}{\hat{Q}_{y_t|x_t}(\tau_i) - \hat{Q}_{y_t|x_t}(\tau_{i-1})},$$

for some appropriately chosen sequence of $\tau$’s, i.e., $\tau \in (\tau_{i-1}, \tau_i)$. 
図: Illustration for the estimator of conditional error density at τ
圖: Illustration for the estimator of conditional error density at $\tau$
Asymptotic Distribution of $\theta_\tau$

Under suitable assumptions, Theorem 2 of Koenker and Xiao (2006) shows that

$$\Sigma^{-1/2} \sqrt{T} (\hat{\theta}_\tau - \theta_\tau) \Rightarrow B_k(\tau),$$

where $B_k(\tau)$ is a $k$-dimensional standard Brownian bridge, $k = p + 1$ and \(\Sigma = \Omega^{-1} \Omega_0 \Omega^{-1}\) with

$$\Omega_1 = \lim_{T \to \infty} T^{-1} \sum_{t=1}^{T} f_{t-1}[F_{t-1}^{-1}(\tau)] \mathbf{x}_t \mathbf{x}_t'$$

$$\Omega_0 = E(\mathbf{x}_t \mathbf{x}_t') = \lim_{T \to \infty} T^{-1} \sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'.$$

By definition, for any fixed $\tau$, $B_k(\tau)$ is $N(0, \tau(1 - \tau)I_k)$. $F_{t-1}(\cdot)$ is the CDF of $e_t$ and its derivative (PDF) $f_{t-1}(\cdot)$
When \( \{e_t\} \) are i.i.d. random variable with mean 0 and variance \( \sigma^2 < \infty \), then Corollary 2 of Koenker and Xiao (2006) shows that

\[
f[F^{-1}(\tau)] \Omega_0^{-1/2} \sqrt{T}(\hat{\theta}_\tau - \theta_\tau) \Rightarrow B_k(\tau),
\]
Inference on the Quantile Autoregression Process

Given the central limit theorem,

\[ V_T(\tau) = \sqrt{T}[R\Omega_1^{-1}\Omega_0\Omega_1^{-1}R']^{-1/2}[R\hat{\theta}_\tau - r] \Rightarrow B_q(\tau), \]

where \( B_q(\tau) \) is a \( q \)-dimensional standard Brownian bridge, that is \( N(0, \tau(1 - \tau)I_q) \), the Wald test for \( H_0 : R\theta_\tau = r \), where \( R \) is a \( q \times k \) matrix and \( r \) is a \( q \times 1 \) vector, is constructed as

\[ W_T(\tau) = T[R\hat{\theta}_\tau - r]'[\tau(1 - \tau)R\hat{\Omega}_1^{-1}\hat{\Omega}_0\hat{\Omega}_1^{-1}R']^{-1}[R\hat{\theta}_\tau - r] \rightarrow \chi^2(q), \]

where \( \hat{\Omega}_0 \) and \( \hat{\Omega}_1 \) are consistent estimators for \( \Omega_0 \) and \( \Omega_1 \), respectively.
If we are interested in testing $\mathbf{R}\theta(\tau) = \mathbf{r}$ over $\tau \in \mathcal{T}$, then the following Kolmogorov-Smirnov (KS)-type sup-Wald test may be considered:

$$KSW_T = \max_{\tau \in \mathcal{T}} \mathcal{W}_T(\tau) \Rightarrow \sup_{\tau \in \mathcal{T}} Q^2_q(\tau),$$

where $Q_q(\tau) = \| \mathbf{B}_q(\tau) \| / \sqrt{\tau(1 - \tau)}$ is a Bessel process of order $q$, where $\| \cdot \|$ represents the Euclidean norm. For any fixed $\tau$, $Q^2_q \sim \chi^2(q)$ is a centered chi-squared random variable with $q$ degrees of freedom.
Testing for Asymmetric Dynamics

The hypothesis that $\theta_j(\tau), j = 1, \ldots, p$ are constant over $\tau$ [i.e., $\theta_j(\tau) = \mu_j, \forall \tau$] can be represented as $H_0 : R\theta_\tau = r$ with unknown parameter vector $r = (\mu_1, \ldots, \mu_p)$. Koenker and Xiao (2006) consider the process

$$\hat{V}_T(\tau) = \sqrt{T}[R\hat{\Omega}_1^{-1}\hat{\Omega}_0\hat{\Omega}_1^{-1}R']^{-1/2}[R\theta_\tau - \hat{r}],$$

where $\hat{r}$ is a $\sqrt{T}$-consistent estimator for $r$, say OLS estimator, and then under the null

$$\hat{V}_T(\tau) \Rightarrow B_p(\tau) - f(F^{-1}(\tau))[R\Omega_0^{-1}R']^{-1/2}Z,$$

where $Z = \lim \sqrt{T}(\hat{r} - r)$. 
To restore the asymptotically distribution-free nature of inference, Koenker and Xiao (2006) uses a martingale transformation proposed by Khmaladze (1981) over the process $\hat{V}_T(\tau)$. Denote $df(x)/dx = f'(x)$ and define

\[
\dot{g}(\zeta) = [1, (f'/f)(F^{-1}(\zeta))]' \text{ and }
\]

\[
C(s) = \int_s^1 \dot{g}(\zeta)\dot{g}(\zeta)'d\zeta,
\]

then, a martingale transformation on $\hat{V}_T(\tau)$ is constructed as

\[
\tilde{V}_T(\tau) = \hat{V}_T(\tau) - \int_0^\tau \left[ \dot{g}_T(s)'C_T(s) \int_s^1 \dot{g}_T(\zeta)d\hat{V}_T(\zeta) \right] ds
\]

where $\dot{g}_T(s)$ and $C_T(s)$ are uniformly consistent estimators of $\dot{g}(\zeta)$ and $C(s)$ over $\tau \in T$. 
The following KS-type test based on the transformed process:

\[
KH_T = \max_{\tau \in T} \| \tilde{V}_T(\tau) \|.
\]

Theorem 4 in Koenker and Xiao (2006) shows that

\[
KH_T = \max_{\tau \in T} \| \tilde{V}_T(\tau) \| \Rightarrow \sup_{\tau \in T} \| W_p(\tau) \|,
\]

where \( W_p(\tau) \) is a \( p \)-dimensional standard Brownian motion. In R package, quantreg, this test can be conducted by using the command

\[ T1 < - KhmaladzeTest(y \sim X, taus = -1, nullH="location") \]

1. \[ T2 < - KhmaladzeTest(y \sim X, taus = 10:290/300, nullH="location", se="ker") \].
Threshold Quantile Autoregressive Models

The following model is considered in Galvao, et al. (2008)

\[ y_t = \theta_0 + \theta_1 y_{t-1} + \delta |y_{t-1}| u_t, \]

and the model can be written as a two-regime model:

\[
\begin{align*}
    y_t &= \begin{cases} 
        \theta_0 + \theta_1 y_{t-1} - \delta y_{t-1} u_t, & y_{t-1} \leq 0, \\
        \theta_0 + \theta_1 y_{t-1} + \delta y_{t-1} u_t, & y_{t-1} > 0.
    \end{cases}
\end{align*}
\]

Although the model is nonlinear, its first conditional moments are identical across regimes,

\[
\begin{align*}
    E[y_t|y_{t-1} \leq 0] &= E[y_t|y_{t-1} > 0] = \theta_0 + \theta_1 y_{t-1} \\
    \operatorname{var}[y_t|y_{t-1} \leq 0] &= \operatorname{var}[y_t|y_{t-1} > 0] = \delta^2 y_{t-1}^2 \sigma_u^2.
\end{align*}
\]
Given the two-regime model, the conditional $\tau$-quantile process, assuming $\delta > 0$ is

\[
\begin{align*}
Q_{y_t}(\tau|y_{t-1} \leq 0) & = \theta_0 + [\theta_1 - \delta F_u^{-1}(\tau)]y_{t-1} \\
Q_{y_t}(\tau|y_{t-1} > 0) & = \theta_0 + [\theta_1 + \delta F_u^{-1}(\tau)]y_{t-1}.
\end{align*}
\]
Consider a nonlinear process \( \{y_t\} \) with two possible regimes defined by the known function \( q_t = q(x_t) \) with
\[
x_t = (1, y_{t-1}, y_{t-2}, \ldots, y_{t-p})',
\]
and threshold parameter \( \gamma(\tau) \), \(-\infty < \gamma(\tau) < \infty\) is allowed for different values across quantiles. Then the QAR\((p)\) with a threshold model is represented as
\[
y_t = \begin{cases} 
\theta_{10} + \theta_{11}(\tau)y_{t-1} + \cdots + \theta_{1p}(\tau)y_{t-p} + u_t(\tau) & q \leq \gamma(\tau), \\
\theta_{20} + \theta_{21}(\tau)y_{t-1} + \cdots + \theta_{2p}(\tau)y_{t-p} + u_t(\tau) & q > \gamma(\tau), 
\end{cases}
\]
\[
= z_t(\gamma(\tau))'\theta(\tau) + e_t(\tau),
\]
with \( \theta(\tau) \) is a \( 2 \times (1 + p) \) parameter vector and
\[
z_t(\gamma(\tau)) = x_t \otimes (1\{q \leq \gamma(\tau)\}, 1\{q > \gamma(\tau)\}),
\]
with \( 1\{\cdot\} \) an indicator function and \( \otimes \) the Kronecker product.
Estimation and Inference

For models with a known threshold parameter $\gamma(\tau) = \gamma_0(\tau)$, the quantile regression estimators are obtained by

$$
\hat{\theta}_{\gamma_0(\tau)}(\tau) = \arg\min_{\theta} \sum_{t=1}^{T} \rho_{\tau}(y_t - z_t(\gamma_0(\tau))'\theta).
$$
However, the threshold parameter $\gamma(\tau)$ is unknown, two-stage estimation procedures are suggested by Galvao, et al. (2008) as follows. For fixed $\tau$, consider a grid of $\gamma(\tau)$ values and for each value estimate the quantile regression estimator and save $\hat{\theta}_{\gamma(\tau)}(\tau)$. Then minimize

$$
\hat{\gamma}(\tau) = \arg\min_{\gamma} \sum_{t=1}^{T} \rho_{\tau}(y_t - z_t(\gamma)'\hat{\theta}_{\gamma(\tau)}(\tau)).
$$
Nonlinearity Tests

For a linear hypothesis at a given $\tau$ and a known $\gamma_0(\tau)$, $H_0 : \mathbf{R}\theta_{\gamma_0(\tau)}(\tau) = \mathbf{r}$, the Wald test statistic is

$$\mathcal{W}_T(\tau, \gamma_0(\tau)) = \frac{T(\hat{\mathbf{R}}\theta_{\gamma_0(\tau)}(\tau) - \mathbf{r})'[\mathbf{R}\hat{\Sigma}(\tau, \gamma_0(\tau))\mathbf{R}']^{-1}(\hat{\mathbf{R}}\theta_{\gamma_0(\tau)}(\tau) - \mathbf{r})}{\tau(1 - \tau)},$$

with $\hat{\Sigma}(\tau, \gamma_0(\tau)) = \hat{\Omega}_1(\tau, \gamma_0(\tau))^{-1}\hat{\Omega}_0(\tau, \gamma_0(\tau))\hat{\Omega}_1(\tau, \gamma_0(\tau))^{-1}$. 

M.-Y. Chen
When $\gamma(\tau_0)$ is unknown

For a fixed $\tau_0$ and $\gamma(\tau_0) \in G$ is unknown, the null hypothesis for the linearity is $H_0 : R\theta_{\gamma(\tau_0)}(\tau_0) = r$ for all $\gamma(\tau_0) \in G$, the test statistics are

$$T_1(\tau_0) = \max_{\gamma(\tau_0) \in G} \mathcal{W}_{T}^{(1)}(\tau_0, \gamma(\tau_0)) \Rightarrow \sup_{\gamma(\tau_0) \in G} Q_{q}^{2}(\gamma(\tau_0))$$

$$T_2(\tau_0) = \text{ave}_{\gamma(\tau_0) \in G} \mathcal{W}_{T}^{(1)}(\tau_0, \gamma(\tau_0)) \Rightarrow \text{ave}_{\gamma(\tau_0) \in G} Q_{q}^{2}(\gamma(\tau_0))$$

with

$$\mathcal{W}_{T}^{(1)}(\tau_0, \gamma(\tau_0)) = T(\hat{R}\theta_{\gamma(\tau_0)}(\tau_0) - r)'[\hat{R}\Sigma(\tau_0, \gamma(\tau_0))\hat{R}']^{-1}(\hat{R}\theta_{\gamma(\tau_0)}(\tau_0) - r),$$

where

$$\hat{\Sigma}(\tau_0, \gamma(\tau_0)) = \tau_0(1 - \tau_0)\hat{\Omega}_1(\tau_0, \gamma(\tau_0))^{-1}\hat{\Omega}_0(\tau_0, \gamma(\tau_0))\hat{\Omega}_1(\tau_0, \gamma(\tau_0))^{-1}.$$
Linearity Tests for Entire $\tau$ with Unknown $\gamma(\tau)$

For testing for linearity for the entire $\tau$-quantile process the Kolmogorov-Smirnov type test statistics are

$$T_1 = \max_{\tau \in T} \max_{\gamma(\tau) \in G} \mathcal{W}_T^{(2)}(\tau, \gamma(\tau)) \Rightarrow \sup_{\tau \in T} \sup_{\gamma(\tau) \in G} Q^2_q(\gamma(\tau))$$

$$T_2 = \max_{\tau \in T} \text{ave}_{\gamma(\tau) \in G} \mathcal{W}_T^{(2)}(\tau, \gamma(\tau)) \Rightarrow \sup_{\tau \in T} \text{ave}_{\gamma(\tau) \in G} Q^2_q(\gamma(\tau))$$

with

$$\mathcal{W}_T^{(2)}(\tau, \gamma(\tau)) = T(\mathbf{R} \hat{\theta}_{\gamma(\tau)}(\tau) - \mathbf{r})' [\mathbf{R} \hat{\Sigma}(\tau, \gamma(\tau)) \mathbf{R}']^{-1} (\mathbf{R} \hat{\theta}_{\gamma(\tau)}(\tau) - \mathbf{r}),$$

where

$$\hat{\Sigma}(\tau, \gamma(\tau)) = \tau(1 - \tau) \hat{\Omega}_1(\tau, \gamma(\tau))^{-1} \hat{\Omega}_0(\tau, \gamma(\tau)) \hat{\Omega}_1(\tau, \gamma(\tau))^{-1}.$$
Quantile Autoregression with a Unit Root

The following autoregression model is considered as

$$y_t = \alpha y_{t-1} + u_t, t = 1, \ldots, T,$$

focusing on the case that $\alpha = 1$. Consider the QAR(1) model as

$$Q_{y_t}(\tau|y_{t-1}) = Q_u(\tau) + \alpha y_{t-1} = \mathbf{x}_t' \alpha(\tau),$$

where $\mathbf{x}_t = (1, y_{t-1})'$ and $\alpha(\tau) = (\alpha_0(\tau), \alpha_1(\tau))'$ with $\alpha_0(\tau) = Q_u(tau)$ and $\alpha_1(\tau) = \alpha$. Denote $\hat{\alpha}_\tau$ as the quantile regression estimator.
Denoting $\psi_{\tau}(u) = \tau - I(u < 0)$, by definition of $u_{t\tau}$, we have $E[\psi_{\tau}(u_{t\tau})|_{t-1}] = 0$. Assuming the partial sums of the vector process $[u_t, \psi_{\tau}(u_{t\tau})]'$ follow a bivariate invariance principle:

$$T^{-1/2} \sum_{t=1}^{[T_r]} [u_t, \psi_{\tau}(u_{t\tau})]' \Rightarrow [B_u(r), B_{\psi_{\tau}}(r)]' = B(0, \Sigma(\tau)),$$

where $\Sigma(\tau) = E\{[u_t, \psi_{\tau}(u_{t\tau})][u_t, \psi_{\tau}(u_{t\tau})]'\}$. Consequently, we have

$$T^{-1} \sum_{t=1}^{T} y_{t-1} \psi_{\tau}(u_{t\tau}) \Rightarrow \int_0^1 B_u dB_{\psi_{\tau}}.$$
Since the limiting variate $B_{\psi_\tau}(r)$ can be viewed as random function of $\tau$, is a Brownian bridge over $\tau \in (0, 1)$. Thus, the two parameter process $B_{\psi_\tau}(r)$ is partially Brownian motion and partially Brownian bridge in the sense that for fixed $r$, $B_{\psi_\tau}(r)$ is a rescaled Brownian bridge, while for each $\tau$, $T^{-1/2} \sum_{t=1}^{[Tr]} \psi_\tau(u_{t\tau})$ converges weakly to a Brownian motion with variance $\tau(1 - \tau)$. Thus, for each fixed pair $(\tau, r)$, $B_{\psi_\tau}(r) \sim N(0, \tau(1 - \tau)r)$. 

M.-Y. Chen
Limiting Distribution of Unit Root Test

If \( y_t = y_{t-1} + u_t \), under suitable stable law for \( u_t \),

\[
T(\hat{\alpha}_1(\tau) - 1) \Rightarrow \frac{1}{f[F_u^{-1}(\tau)]} \frac{\int_0^1 [B_u(r) - \int_0^1 B_u(r)dr]dB_{\psi \tau}(r)}{\int_0^1 [B_u(r) - \int_0^1 B_u(r)dr]^2dr}.
\]
Consider the augmented Dicky-Fuller (ADF) regression model:

\[ y_t = \alpha_1 y_{t-1} + \sum_{j=1}^{q} \alpha_{j+1} \triangle y_{t-j} + u_t. \]

The \( \tau \)-th conditional quantile of \( y_t \), conditional on \( \mathcal{F} = \sigma(u_s, s \leq t) \), is given by

\[ Q_{y_{t-1}}(\tau|\mathcal{F}_{t-1}) = Q_u(\tau) + \alpha_1 y_{t-1} + \sum_{j=1}^{q} \alpha_{j+1} \triangle y_{t-j}. \]
Let $\alpha_0(\tau) = Q_u(\tau)$, $\alpha_j(\tau) = \alpha_j$, $j = 1, \ldots, q + 1$, and define $\alpha_\tau = (\alpha_0(\tau), \alpha_1, \ldots, \alpha_{q+1})'$, $x_t = (1, y_{t-1}, \Delta y_{t-1}, \ldots, \Delta y_{t-q})'$, then

$$Q_{y_{t-1}}(\tau|\mathcal{F}_{t-1}) = x_t' \alpha_\tau.$$
Denote $\hat{\alpha}_\tau = (\hat{\alpha}_0(\tau), \hat{\alpha}_1, \ldots, \hat{\alpha}_{q+1})$ and 
$D_n = \text{diag}(\sqrt{T}, T, \sqrt{T}, \ldots, \sqrt{T})$ and assume the following FCLT holds under the unit root assumption

$$T^{-1/2} \sum_{t=1}^{[Tr]} (w_t, \psi_\tau(u_{t\tau}))' \Rightarrow (B_w(r), B_\psi^\tau(r))' = B(0, \Sigma_{\tau}(r)),$$

where

$$\Sigma(\tau) = \begin{bmatrix} \sigma_w^2 & \sigma_{w\psi}(\tau) \\ \sigma_{w\psi}(\tau) & \sigma_{\psi}^2(\tau) \end{bmatrix}$$

is the long run covariance matrix of the bivariate Brownian motion and can be written as $\Sigma_0(\tau) + \Sigma_1(\tau) + \Sigma(\tau)'$, where

$$\Sigma_0 = E[(w_t, \psi_\tau(u_{t\tau}))'(w_t, \psi_\tau(u_{t\tau}))],$$

$$\Sigma_1 = \sum_{s=2}^{\infty} E[(w_1, \psi_\tau(u_{1\tau}))'(w_s, \psi_\tau(u_{s\tau}))].$$
Since the limiting distribution of $\sqrt{T}(\hat{\alpha}_1(\tau))$ is invariant to the estimation of $\hat{\alpha}_j(\tau), j = 2, \ldots, q + 1$ and the lag length $q$, we have

$$\sqrt{T}(\hat{\alpha}_1(\tau)) \Rightarrow \frac{1}{f[F_u^{-1}(\tau)]} \frac{\int_0^1 [B_w(r) - \int_0^1 B_w(r)dr]dB_{\psi_\tau}(r)}{\int_0^1 \left[ B_w(r) - \int_0^1 B_w(r)dr \right]^2 dr}.$$
Augmented Dicky-Fuller $t$-ratio Test

The conventional augmented Dicky-Fuller $t$-ratio test statistic is

$$t_T(\tau) = \frac{\hat{f}[F_u^{-1}(\tau)]}{\sqrt{\tau(1-\tau)}} (y'_{-1} P x_t y_{-1})^{1/2} (\hat{\alpha}_1(\tau) - 1),$$

where $\hat{f}[F_u^{-1}(\tau)]$ is a consistent estimator of $f[F_u^{-1}(\tau)]$, $y_{-1}$ is the
Quantile ARCH\((q)\) Models

The ARCH\((q)\) quantile model (or linear ARCH\((q)\) models) of Koenker and Zhao (1996) is described as

\[
\begin{align*}
    r_t & = \mu_t + \epsilon_t \\
    \epsilon_t | F_{t-1} & = \sqrt{h_t} z_t \\
    & = (\gamma_0 + \gamma_1 |\epsilon_{t-1}| + \cdots + \gamma_q |\epsilon_{t-q}|) z_t.
\end{align*}
\]

In which the fundamental innovation \(z_t\) is from an unknown distribution \(F_{z_t}\). This linear ARCH\((q)\) model is less sensitive to extreme returns. Then, the VaR value is calculated as

\[
V_t(\alpha) \equiv Q_{r_t}(\alpha|F_{t-1}) = \mu_t + Q_\epsilon(\alpha|F_{t-1}).
\]
\[
Q_{rt}(\alpha|\mathcal{F}_{t-1})
\]
\[
= \mu_t + Q_{\epsilon}(\alpha|\mathcal{F}_{t-1})
\]
\[
= \mu_t + \sqrt{h_t} Q_z(\alpha)
\]
\[
= \mu_t + (\gamma_0 + \gamma_1 |\epsilon_{t-1}| + \cdots + \gamma_q |\epsilon_{t-q}|) Q_z(\alpha)
\]
\[
= \mu_t + (\gamma_0 Q_z(\alpha) + \gamma_1 Q_z(\alpha) |\epsilon_{t-1}| + \cdots + \gamma_q Q_z(\alpha) |\epsilon_{t-q}|)
\]
\[
= \mu_t + (1, |\epsilon_{t-1}|, \ldots, |\epsilon_{t-q}|)(\gamma_0 Q_z(\alpha), \gamma_1 Q_z(\alpha), \ldots, \gamma_q Q_z(\alpha))'
\]
Denote $X_t = (1, |\epsilon_{t-1}|, \ldots, |\epsilon_{t-q}|)'$ and then

$$Q_\epsilon(\alpha|\mathcal{F}_{t-1}) = X'_t \lambda(\alpha),$$

with

$$\lambda(\alpha) = (\gamma_0 Q_z(\alpha), \gamma_1 Q_z(\alpha), \ldots, \gamma_q Q_z(\alpha))'$$

$$\equiv: (\lambda_0(\alpha), \lambda_1(\alpha), \ldots, \lambda_q(\alpha))'$$

and $Q_z(\alpha) = F_z^{-1}(\alpha)$. 
The parameter vector $\lambda(\alpha)$ can be estimated by a quantile regression at quantile $\alpha$ for regressing $\hat{e}_t = r_t - \hat{\mu}_t$ on $1, |\hat{e}_{t-1}|, \ldots, |\hat{e}_{t-q}|$. That is,

$$
\hat{\lambda}(\alpha) = \arg\min \left[ \sum_{t \in \{ t : \hat{e}_t \geq X'_t \lambda \}} \alpha|\hat{e}_t - X'_t \lambda| + \sum_{t \in \{ t : \hat{e}_t < X'_t \lambda \}} (1 - \alpha)|\hat{e}_t - X'_t \lambda| \right].
$$

Then the estimated VaR is

$$
\hat{V}_t(\alpha) = \hat{\mu}_t + X'_t \hat{\lambda}(\alpha).
$$
In practice, $\epsilon_t$ is replaced with their OLS estimators

$$\hat{\epsilon}_t = r_t - \hat{\mu}_t.$$
Quantile GARCH\((p, q)\) Models

The quantile GARCH\((p, q)\) quantile model (or linear GARCH\((p, q)\) models) of Xiao and Koenker (2009) is described as

\[
\begin{align*}
    r_t &= \mu_t + \epsilon_t \\
    \epsilon_t | \mathcal{F}_{t-1} &= \sqrt{h_t} z_t \\
    &= (\beta_0 + \beta_1 \sqrt{h_{t-1}} + \cdots + \beta_p \sqrt{h_{t-p}} + \gamma_1 |\epsilon_{t-1}| + \cdots + \gamma_q |\epsilon_{t-q}|) z_t.
\end{align*}
\]

Similarly, the VaR value is calculated as

\[
V_t(\alpha) \equiv Q_{r_t}(\alpha | \mathcal{F}_{t-1}) = \mu_t + Q_{\epsilon}(\alpha | \mathcal{F}_{t-1}).
\]
\[ Q_{r_t}(\alpha|\mathcal{F}_{t-1}) \]
\[ = \mu_t + Q_\epsilon(\alpha|\mathcal{F}_{t-1}) \]
\[ = \mu_t + \sqrt{h_t} Q_z(\alpha) \]
\[ = \mu_t + \left( \beta_0 + \beta_1 \sqrt{h_{t-1}} + \cdots + \beta_p \sqrt{h_{t-p}} \right. \]
\[ + \gamma_1 |\epsilon_{t-1}| + \cdots + \gamma_q |\epsilon_{t-q}| \big) Q_z(\alpha) \]
\[ = \mu_t + \left( \beta_0 Q_z(\alpha) + \beta_1 \sqrt{h_{t-1}} Q_z(\alpha) + \cdots + \beta_p \sqrt{h_{t-p}} Q_z(\alpha) \right. \]
\[ + \gamma_1 Q_z(\alpha)|\epsilon_{t-1}| + \cdots + \gamma_q Q_z(\alpha)|\epsilon_{t-q}| \big) \]
\[ = \mu_t + (1, \sqrt{h_{t-1}}, \cdots, \sqrt{h_{t-p}}, |\epsilon_{t-1}|, \cdots, |\epsilon_{t-q}|) \]
\[ (\beta_0 Q_z(\alpha), \beta_1 Q_z(\alpha), \cdots, \beta_p Q_z(\alpha), \gamma_1 Q_z(\alpha), \cdots, \gamma_q Q_z(\alpha))^t. \]
Also denote

\[
X_t = (1, \sqrt{h_{t-1}}, \ldots, \sqrt{h_{t-p}}, |\epsilon_{t-1}|, \ldots, |\epsilon_{t-q}|)'
\]

\[
\lambda(\alpha) = (\beta_0 Q_z(\alpha), \beta_1 Q_z(\alpha), \ldots, \beta_p Q_z(\alpha), \gamma_1 Q_z(\alpha), \ldots, \gamma_q Q_z(\alpha))'
\]

\[
\equiv: (\lambda_0(\alpha), \lambda_1(\alpha), \ldots, \lambda_p(\alpha), \lambda_{p+1}(\alpha), \ldots, \lambda_{p+q}(\alpha))'
\]

and then

\[
Q_{\epsilon_t}(\alpha|\mathcal{F}_{t-1}) = X_t' \lambda(\alpha),
\]

with \(Q_z(\alpha) = F_z^{-1}(\alpha)\).
Since $\sqrt{h_{t-j}} Q_z(\alpha) = Q_{\epsilon_{t-j}}(\alpha|F_{t-j-1})$, the conditional quantiles $Q_{\epsilon_t}(\alpha|F_{t-1})$ follows CAViaR($p, q,$) representation of Engle and Manganelli (2004) as

$$Q_{\epsilon_t}(\alpha|F_{t-1}) = X_t' \lambda(\alpha)$$

$$= \left( \beta_0 Q_z(\alpha) + \beta_1 \sqrt{h_{t-1}} Q_z(\alpha) + \cdots + \beta_p \sqrt{h_{t-p}} Q_z(\alpha) \\
+ \gamma_1 Q_z(\alpha) |\epsilon_{t-1}| + \cdots + \gamma_q Q_z(\alpha) |\epsilon_{t-q}| \right)$$

$$= \left( \beta_0 Q_z(\alpha) + \beta_1 Q_{\epsilon_{t-1}}(\alpha|F_{t-2}) + \cdots + \beta_p Q_{\epsilon_{t-p}}(\alpha|F_{t-p-1}) \\
+ \gamma_1 Q_z(\alpha) |\epsilon_{t-1}| + \cdots + \gamma_q Q_z(\alpha) |\epsilon_{t-q}| \right)$$

$$= \beta_0^*(\alpha) + \sum_{i=1}^{p} \beta_i^* Q_{\epsilon_{t-i}}(\alpha|F_{t-i-1}) + \sum_{j=1}^{q} \gamma_j^* |\epsilon_{t-j}|.$$
Quantile Regression Estimation of GARCH Models

Since $X_t$ contains $\sqrt{h_{t-i}}, i = 1, \ldots, p$ which in turn depend on unknown parameters $\theta = (\beta_0, \beta_1, \ldots, \beta_p, \gamma_1, \ldots, \gamma_q)'$, thus $X_t$ depends on $\theta$ and is written as $X_t(\theta)$. This dependence makes the estimation of conditional quantiles of $\epsilon_t$ troublesome. To overcome this problem, Xiao and Koenker (2009) suggests a two-step estimator for the conditional quantile of $\epsilon_t$. This two-step estimator can be rationale as follows.
As an GARCH\((p, q)\) model, under some regularity assumptions, can be represented as an ARCH\((\infty)\) model:

\[
\sqrt{h_t} = a_0 + \sum_{j=1}^{\infty} a_j |\epsilon_{t-j}|,
\]

\(a_0 = 1\) is set for identification. Then

\[
\epsilon_t = \left( a_0 + \sum_{j=1}^{\infty} a_j |\epsilon_{t-j}| \right) z_t,
\]

and

\[
Q_{\epsilon_t}(\alpha|\mathcal{F}_{t-1}) = b_0(\alpha) + \sum_{j=1}^{\infty} b_j(\alpha) |\epsilon_{t-j}|,
\]

where \(b_j(\alpha) = a_j Q_{z_t}(\alpha), j = 0, 1, 2, \ldots\)
Suppose $a_j$ decreases geometrically, so letting $m = m(T)$ denote a truncation parameter so that the truncated quantile autoregression of Koenker and Xiao (2006) is considered:

$$Q_{\epsilon_t}(\alpha|\mathcal{F}_{t-1}) \approx b_0(\alpha) + \sum_{j=1}^{m} b_j(\alpha)|\epsilon_{t-j}|.$$

By choosing $m$ suitably small relative to the sample size $T$, but large enough to avoid serious bias, a sieve approximation for the GARCH model is obtained and can be estimated as

$$\tilde{Q}_{\epsilon_t}(\alpha|\mathcal{F}_{t-1}) = \tilde{b}_0(\alpha) + \tilde{b}_1(\alpha)|\epsilon_{t-1}| + \cdots + \tilde{b}_m(\alpha)|\epsilon_{t-m}|,$$

where $\tilde{b}_j(\alpha)$ are the quantile autoregression estimates.
Suppose the $m$-th order quantile autoregression is estimated as

$$\tilde{b} = \arg \min_b \sum_{t=m+1}^{T} \rho_\alpha \left( \epsilon_t - b_0 - \sum_{j=1}^{m} b_j |\epsilon_{t-j}| \right)$$

(1)

at quantiles $(\alpha_1, \ldots, \alpha_K)$, and then obtain estimates

$$\tilde{b}(\alpha_k), k = 1, \ldots, K.$$
Denote

\[ \xi = (a_1, \ldots, a_m, Q_z(\alpha_1), \ldots, Q_z(\alpha_K))' =: (a', Q')' \]

\[ \tilde{B} = (\tilde{b}(\alpha_1)', \ldots, \tilde{b}(\alpha_K)')' \]

and then

\[ \phi(\xi) = Q \otimes a = (Q_z(\alpha_1), a_1 Q_z(\alpha_1), \ldots, a_m Q_z(\alpha_1), Q_z(\alpha_2), a_1 Q_z(\alpha_2), \ldots, a_m Q_z(\alpha_2), \ldots, Q_z(\alpha_K), a_1 Q_z(\alpha_K), \ldots, a_m Q_z(\alpha_K))' \]

Consider the following estimator

\[ \tilde{\xi} = \arg \min_{\xi} (\tilde{B} - \phi(\xi))' A_T (\tilde{B} - \phi(\xi)), \quad (2) \]

where \( A_T \) is a \([K(m+1) \times K(m+1)]\) positive definite matrix.
Summary of Two-step Estimator: Step 1

1. Estimate the $m$-th order quantile regression ($\beta\alpha_k$) at quantiles $(\alpha_1, \ldots, \alpha_K)$ and obtain $\tilde{b}(\alpha_k), k = 1, \ldots, K$;

2. By setting $\tilde{a}_0 = 1$ and solving the minimum distance estimation problem ($\xi$), we obtain $\tilde{\xi}$ in which estimates $(\tilde{a}_0, \tilde{a}_1, \ldots, \tilde{a}_m)$ are included.

3. Thus, $\sqrt{h_t}$ can be estimated by

$$ \sqrt{\tilde{h}_t} = \tilde{a}_0 + \tilde{a}_1|\epsilon_{t-1}| + \cdots + \tilde{a}_m|\epsilon_{t-m}|. $$
Summary of Two-step Estimator: Step 2

1. Quantile regression of $\epsilon_t$ on

$$\tilde{X}_t = \left(1, \sqrt{\tilde{h}_{t-1}}, \ldots, \sqrt{\tilde{h}_{t-p}}, |\epsilon_{t-1}|, \ldots, |\epsilon_{t-q}|\right)'$$

as

$$\hat{\lambda}(\alpha) = \arg \min_{\lambda} \sum \rho_{\alpha}(\epsilon_t - \tilde{X}_t'\lambda(\alpha)).$$

2. The $\alpha$-th conditional quantile of $\epsilon_t$ is estimated by

$$\hat{Q}_{\epsilon_t}(\alpha|\mathcal{F}_{t-1}) = \tilde{X}_t'\hat{\lambda}(\alpha).$$

3. Iteration can be applied to the above procedure for further improvement.
Finally, the VaR value is calculated as

\[
\hat{V}_t(\alpha) = \hat{Q}_{r_t}(\alpha|\mathcal{F}_{t-1}) = \hat{\mu}_t + \hat{Q}_\epsilon(\alpha|\mathcal{F}_{t-1}).
\]

As before, the unknown error \( \epsilon_t \) is approximated by the OLS residuals, \( \hat{\epsilon}_t = r_t - \hat{\mu}_t \). In the simulation study of Xiao and Koenker (2009), \( m = 3T^{1/4}, A_T = I_T \) and \( \alpha_k = 5k \% \), \( k = 1, \ldots, 19 \) are considered.