Financial Econometrics I

Lecture 5

Fu Ouyang

December 8, 2018

Outline

Heteroscedastic Volatility Models

ARCH Models

GARCH Models

Estimation for GARCH Models

Heteroscedastic Volatility Models

- *Volatility* measures the uncertainty of asset returns, which is usually defined as the conditional standard deviation of an asset return given all the available information up to the present time.
- ARMA models are for modeling the conditional mean of a time series, which facilitates the forecasting of asset returns. However, the white noise assumption says nothing about the conditional volatility.
- To carry out statistical inference for ARMA models, we often assume IID innovations. But this is questionable as time-varying volatility (i.e., conditional *heteroscedasticity*) is often observed in real-world data (e.g., volatility clustering).
- In the following couple of lectures, we are going to introduce popular heteroscedasticity models such as ARCH, GARCH, and stochastic volatility models, which can be used to model time-varying volatility.

Outline

Heteroscedastic Volatility Models

ARCH Models

GARCH Models

Estimation for GARCH Models

ARCH Models

A standard model for asset returns is

$$r_t = \mu_t + X_t$$

where μ_t denotes the conditional mean of r_t, X_t is a diffusion term which may be modeled as

$$X_t = \sigma_t \epsilon_t, \epsilon_t \sim \text{IID}(0, 1)$$

where $\sigma_t > 0$ is determined by the information available before time t.

In practice, often assume $\mu_t = E(r_t) = \mu$, and estimate μ by $\sum_{t=1}^{T} r_t / T$.

σ_t is time varying and often called a *volatility function*. A specification for σ_t is the *autoregressive conditional heteroscedastic* (ARCH) *model*:

$$\sigma_t^2 = a_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2$$

where $a_0 > 0$ and $a_j \geq 0$, $j = 1, ..., p$ are constants. We say $X_t \sim \text{ARCH}(p)$.

ARCH Models

Denote $E_t[\cdot] = E[\cdot|X^t]$ and $V_t[\cdot] = Var[\cdot|X^t]$. It is easy to see that

- $\{X_t\}$ is a martingale difference sequence and unpredictable.

$$E_{t-1}(X_t) = \sigma_t E_{t-1}(\epsilon_t) = 0$$

- σ_t is the conditional standard deviation of X_t given X^{t-1}.

$$V_{t-1}(X_t) = E_{t-1}(X_t^2) = \sigma_t^2$$

- Let $\eta_t \equiv \sigma_t^2(\epsilon_t^2 - 1)$.

$$X_t^2 = \sigma_t^2 + \eta_t = a_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2 + \eta_t$$

$$E_{t-1}(\eta_t) = \sigma_t^2 E(\epsilon_t^2 - 1) = 0$$

i.e., $\{\eta_t\}$ is martingale difference sequence ($\Rightarrow$ WN) and $X_t^2 \sim \text{AR}(p)$.

Example: ARCH(1) Model

ARCH(1) Model

$$X_t = \sigma_t \epsilon_t, \epsilon_t \sim \text{IID}(0, 1)$$

$$\sigma_t^2 = a_0 + a_1 X_{t-1}^2, a_0 > 0, a_1 \geq 0$$

Hence,

$$X_t^2 = a_0 + a_1 X_{t-1}^2 + \eta_t$$

Provided that $\{X_t\}$ is strictly stationary and $E(X_t^4) < \infty$, this AR(1) representation gives that $Corr(X_t^2, X_{t+k}^2) = a_1^{|k|}$.

- $\{X_t\}$ is martingale difference, compatible with the EMH.
- $\{X_t^2\}$ is auto-correlated $\Rightarrow$ predictable volatility.

Example: ARCH(1) Model

ARCH(1) Model

Note that

$$V_t(X_{t+k}) = E_t(X_{t+k}^2) = a_0 + a_1 E(X_{t+k-1}^2) = \frac{a_0(1-a_1^k)}{1-a_1} + a_1^k X_t^2$$

i.e., large $|X_t|$ leads to large volatilities in near future (volatility clustering).

Provided that $E(\epsilon_t^k) < \infty$ and $E(X_t^k) < \infty$ for some $k \geq 1$, by the LIE and the independence of ϵ_t and $\{X_t\}$

$$E(X_t^k) = E(\sigma_t^k \epsilon_t^k) = E[\sigma_t^k E_{t-1}(\epsilon_t^k)] = E(\sigma_t^k) E(\epsilon_t^k)$$

Kurtosis (κ) measures if the data are heavy-tailed (relative to $N(0,1)$), i.e., data sets with high kurtosis tend to have heavy tails, or outliers.

Example: ARCH(1) Model

ARCH(1) Model

By definition and the Cauchy–Schwarz inequality,

$$\kappa_X = \frac{E(X_t^4)}{[E(X_t^2)]^2} = \frac{E(\sigma_t^4)E(\epsilon_t^4)}{[E(\sigma_t^2)]^2[E(\epsilon_t^2)^2} = \kappa_\epsilon \frac{E(\sigma_t^4)}{[E(\sigma_t^2)]^2} \geq \kappa_\epsilon$$

i.e., $\{X_t\}$ has heavier tails than $\{\epsilon_t\}$.

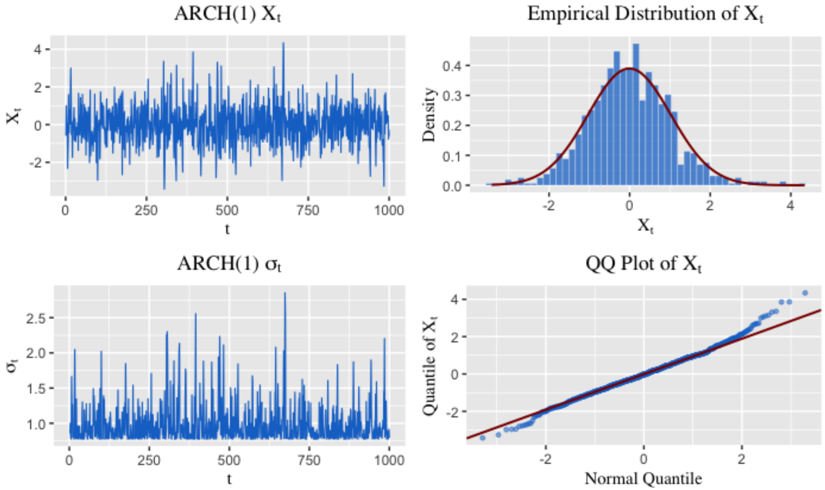
Note that for strongly stationary X_t

$$E(X_t^4) = \frac{a_0^2(1+a_1)E(\epsilon_t^4)}{(1-a_1)(1-a_1^2E(\epsilon_t^4))} < \infty \Rightarrow a_1 < 1/\sqrt{E(\epsilon_t^4)}$$

If $\epsilon_t \sim N(0,1)$, then $a_1 < 1/\sqrt{E(\epsilon_t^4)} = 1/\sqrt{3}$.

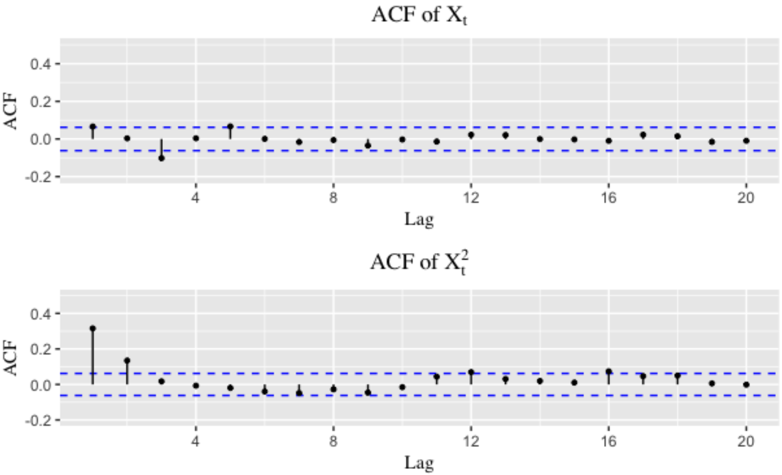
Example: ARCH(1) Model

Figure 1: $X_t = \sigma_t \epsilon_t, \sigma_t^2 = 0.6 + 0.4 X_{t-1}^2, \epsilon_t \sim N(0,1)$



Example: ARCH(1) Model

Figure 2: $X_t = \sigma_t \epsilon_t, \sigma_t^2 = 0.6 + 0.4X_{t-1}^2, \epsilon_t \sim N(0,1)$



Outline

Heteroscedastic Volatility Models

ARCH Models

GARCH Models

Estimation for GARCH Models

GARCH Models

Perhaps the most important extension of ARCH model is the *generalized autoregressive conditional heteroscedastic* (GARCH) *model*:

$$X_t = \sigma_t \epsilon_t, \epsilon_t \sim \text{IID}(0,1)$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i X_{t-i}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2$$

where $a_0 > 0$ and $a_i \geq 0$, $b_j \geq 0$. We say $X_t \sim \text{GARCH}(p,q)$.

Let $\eta_t \equiv \sigma_t^2(\epsilon_t^2 - 1)$, $a_{p+j} = b_{q+j} = 0, \forall j \geq 1$ and $p \vee q \equiv \max\{p,q\}$. We have

$$\begin{aligned} X_t^2 = \sigma_t^2 + \eta_t &= a_0 + \sum_{i=1}^{p} a_i X_{t-i}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2 + \eta_t \\ &= a_0 + \sum_{i=1}^{p \vee q} (a_i + b_i) X_{t-i}^2 + \eta_t - \sum_{j=1}^{q} b_j \eta_{t-j} \end{aligned}$$

Stationarity of GARCH Models

It is easy to verify that $\{\eta_t\}$ is a sequence of martingale differences, and so $\{X_t^2\}$ is an ARMA($p \vee q, q$) process.

If $\{X_t^2\}$ is stationary, provided that $E(X_t^2) < \infty$,

$$E(X_t^2) = E(a_0 + \sum_{i=1}^{p} a_i X_{t-i}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2 + \eta_t)$$
$$= a_0 + \sum_{i=1}^{p} a_i E(X_t^2) + \sum_{j=1}^{q} b_j E(X_t^2)$$

Solving this gives the *long-run variance*:

$$Var(X_t) = E(X_t^2) = \frac{a_0}{1 - \sum_{i=1}^{p} a_i - \sum_{j=1}^{q} b_j} \Rightarrow \sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j < 1, a_0 > 0$$

Stationarity of GARCH Models

In fact, $\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j < 1$ is also sufficient for $\{X_t^2\}$ being stationary.

To see this, note that the characteristic polynomial

$$a(z) = 1 - \sum_{i=1}^{p \vee q} (a_i + b_i) z^i$$

has all roots outside the unit circle when $\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j < 1$ since for all $|z| \leq 1$

$$|a(z)| \geq 1 - \sum_{i=1}^{p \vee q} (a_i + b_i) |z^i| \geq 1 - \sum_{i=1}^{p \vee q} (a_i + b_i) > 0$$

Then, this gives us the following general results:

Stationarity of GARCH Models

Stationarity of GARCH Models

The necessary and sufficient condition for a GARCH(p, q) model defining a unique strictly stationary process $\{X_t, t = 0, \pm 1, ...\}$ with $E(X_t^2) < \infty$ is

$$\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j < 1$$

Furthermore, $E(X_t) = 0$ and

$$Var(X_t) = E(X_t^2) = \frac{a_0}{1 - \sum_{i=1}^{p} a_i - \sum_{j=1}^{q} b_j}$$

Stationarity of GARCH Models

- The necessary and sufficient condition for a ARCH(p) model being unique strictly stationary is

$$\sum_{i=1}^{p} a_i < 1$$

- Why weak stationarity $\Rightarrow$ strict stationarity? Note that

$$\sigma_t^2 = b(B)^{-1}(a_0 + \sum_{i=1}^{p} a_i X_{t-i}^2)$$

where $b(z) = 1 - \sum_{j=1}^{q} b_j z^j$.

$$\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j < 1 \Rightarrow \sum_{j=1}^{q} b_j < 1 \Rightarrow \sigma_t^2 \sim \mathrm{ARCH}(\infty)$$

Example: GARCH(1,1) Model

GARCH(1,1) Model

Consider GARCH(1,1) model

$$X_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = a_0 + a_1 X_{t-1}^2 + b_1 \sigma_{t-1}^2, a_0 > 0, a_1 \geq 0, b_1 \geq 0$$

Note that for stationary GARCH model, $a_1 + b_1 < 1$, and so

$$\begin{aligned}
\sigma_t^2 &= (1 - b_1 B)^{-1}(a_0 + a_1 X_{t-1}^2) \\
&= (1 + b_1 B + b_1^2 B^2 + \cdots)(a_0 + a_1 X_{t-1}^2) \\
&= \frac{a_0}{1 - b_1} + a_1(X_{t-1}^2 + b_1 X_{t-2}^2 + b_1^2 X_{t-3}^2 + \cdots)
\end{aligned}$$

Then GARCH(1,1) has an ARCH(∞) representation.

Example: GARCH(1,1) Model

GARCH(1,1) Model

Recall that GARCH(1,1) model also has an ARMA(1,1) representation

$$X_t^2 = a_0 + (a_1 + b_1)X_{t-1}^2 + \eta_t - b_1\eta_{t-1}$$

which gives us a convenient way to compute $\gamma(k)$.

Assume $E(X_t^4) < \infty$ and define

$$\sigma^2 = E(\eta_t^2) = E[(\epsilon_t^2 - 1)^2]E(\sigma_t^4)$$

Then for $\{X_t^2\}$

$$\gamma(0) = \frac{1 + b_1^2 - 2(a_1 + b_1)b_1}{1 - (a_1 + b_1)^2}\sigma^2$$

Example: GARCH(1,1) Model

GARCH(1,1) Model

$$\gamma(1) = \frac{-b_1 + (a_1 + b_1)[1 + b_1^2 - (a_1 + b_1)b_1]}{1 - (a_1 + b_1)^2}\sigma^2$$

and

$$\gamma(k) = (a_1 + b_1)^{k-1}\gamma(1), k \geq 1$$

Note that if $a_1 = 0$, b_1 is not identifiable. This result can be extended to general GARCH(p, q) models where $a_1, ..., a_p$ cannot be all zeros, otherwise $b_1, ..., b_q$ are unidentifiable (and the model is trivial, right?).

Existence of Fourth Moments

Recall that to derive $\gamma(k)$ of $\{X_t^2\}$, we assume $E(X_t^4) < \infty$. However, this is not always true. To see this, first note that

$$E(X_t^4) = E(\sigma_t^4)E(\epsilon_t^4) = \frac{E(\epsilon_t^4)}{[E(\epsilon_t^2)]^2}E(\sigma_t^4) = \kappa_\epsilon E(\sigma_t^4)$$

For $\epsilon_t \overset{iid}{\sim} N(0,1)$, $\kappa_\epsilon < \infty$, then $E(X_t^4) < \infty$ if and only if $E(\sigma_t^4) < \infty$.

Consider GARCH(1,1) model, where we have (with some algebra)

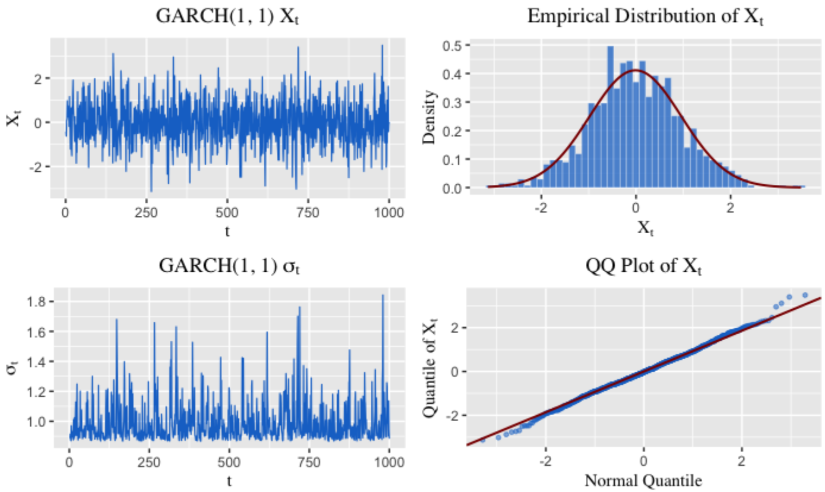
$$E(\sigma_t^4) = a_0^2 + (a_1^2\kappa_\epsilon + 2a_1b_1 + b_1^2)E(\sigma_t^4) + 2a_0(a_1 + b_1)E(\sigma_t^2)$$

which implies that $E(\sigma_t^4) < \infty$ if and only if $a_1^2\kappa_\epsilon + 2a_1b_1 + b_1^2 < 1$.

Note that this is stronger than the stationarity condition $a_1 + b_1 < 1$.

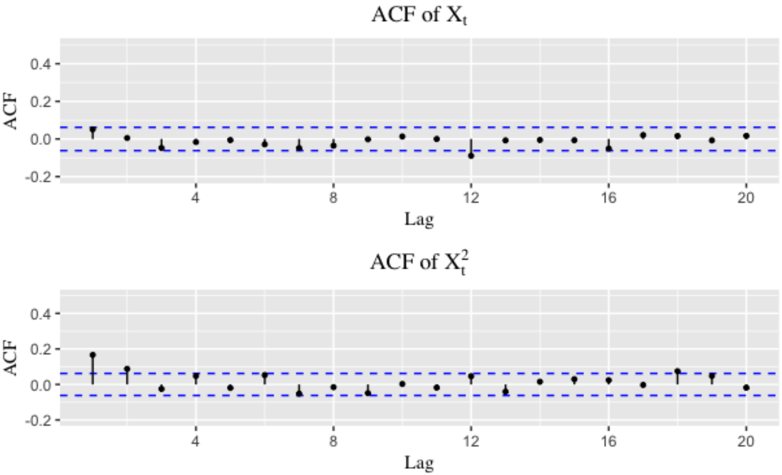
Example: GARCH(1,1) Model

Figure 3: $X_t = \sigma_t \epsilon_t, \sigma_t^2 = 0.6 + 0.2X_{t-1}^2 + 0.2\sigma_{t-1}^2, \epsilon_t \sim N(0,1)$



Example: GARCH(1,1) Model

Figure 4: $X_t = \sigma_t \epsilon_t, \sigma_t^2 = 0.6 + 0.2X_{t-1}^2 + 0.2\sigma_{t-1}^2, \epsilon_t \sim N(0,1)$



Outline

Heteroscedastic Volatility Models

ARCH Models

GARCH Models

Estimation for GARCH Models

Estimating GARCH Model using its ARMA Representation?

To fit a GARCH(p, q) model, one may consider the ARMA representation for $\{X_t^2\}$:

$$X_t^2 = a_0 + \sum_{i=1}^{p \vee q} (a_i + b_i) X_{t-i}^2 + \eta_t - \sum_{j=1}^{q} b_j \eta_{t-j}$$

However, ARMA coefficients are determined by $\gamma(k)$. When $E(X_t^4) = \infty$, $\gamma(k)$ is not well-defined.

In practice, estimating GARCH models using ARMA representations is *not* recommended.

Estimating GARCH Model using its ARMA Representation?

Example: ARCH(1)

Consider ARCH(1) model with $(a_0, a_1) = (0.1, 0.9)$, $\epsilon_t \sim N(0, 1)$.

$$X_t^2 = a_0 + a_1 X_{t-1}^2 + \eta_t$$

is stationary. But $E(X_t^4) = \infty$ as

$$a_1^2 \kappa_\epsilon + 2a_1 b_1 + b_1^2 = 0.9^2 \times 3 + 0 + 0 = 2.43 > 1$$

Conditional MLE for GARCH Models

Suppose that $X^T \equiv \{X_1, ..., X_T\}$ are observations from GARCH(p, q) model:

$$X_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i X_{t-i}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2$$

- Assume $\epsilon_t \sim$ IID F_ϵ (e.g. $N(0, 1)$, t_{ν_0}, generalized Gaussian).
- Assume $\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j < 1$.
- We want to estimate $\theta \equiv (a_0, a_1, ..., a_p, b_1, ..., b_q)$ for fixed (p, q).

The most often used estimators for θ are conditional maximum likelihood estimator (MLE).

Conditional MLE for GARCH Models

Let f_ϵ be the PDF of ϵ_t. The conditional density function for $X_t|X^{t-1}$ is

$$f(X_t|X^{t-1};\theta) = f_\epsilon(X_t/\sigma_t)/\sigma_t$$

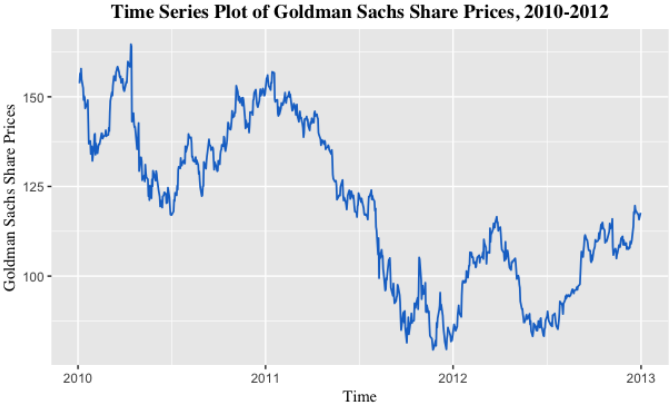
Let $\nu = p \vee q + 1$.

$$\begin{aligned} f(X_v, ..., X_T|X^{\nu-1};\theta) &= f(X_T|X^{T-1};\theta)\cdots f(X_\nu|X^{\nu-1};\theta) \\ &= \prod_{t=\nu}^{T} f(X_t|X^{t-1};\theta) = \prod_{t=\nu}^{T}[f_\epsilon(X_t/\sigma_t)/\sigma_t] \end{aligned}$$

The conditional MLE is defined as

$$\begin{aligned} \widehat{\theta} &= \arg\max_\theta \log f(X_v, ..., X_T|X^{\nu-1};\theta) \\ &= \arg\max_\theta \sum_{t=\nu}^{T}[-\log\sigma_t + \log f_\epsilon(X_t/\sigma_t)] \end{aligned}$$

Application: Daily Returns of Goldman Sachs Share Prices, 2010-2012



See [R markdown]