Lecture 3: The GARCH Model

Zheng Tian

Outline

The problem of ARCH models

The principle of parsimony

- Merriam-Webster:
  1. the quality of being careful with money or resources
  2. the quality or state of being stingy
- Econometric modeling
  - Use a concise model specification
  - Object to overparameterization

The problem of ARCH model

- Estimate so many parameters to fully capture higher-order autoregressive relationship in a_t^2.
- Think of how many parameters in an ARCH(m) model?

The GARCH model

Generalized ARCH model

- Bollerslev (1986) proposes an extension of ARCH, known as the Generalized ARCH (GARCH) model.
- A high-order ARCH model may have a more parsimonious GARCH representation.

The mean equation

$$r_t = \mu_t + a_t$$

where

- μ_t is modeled with an appropriate regression model or some ARMA specification.
- a_t is the innovation at time t.

The volatility equation

$$\sigma_t^2 = \mathrm{Var}(\alpha_t^2 | F_{t-1})$$

The GARCH(m, s) model

$$a_t = \sigma_t \epsilon_t, \ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \quad (1)$$

where

- $\epsilon_t \sim i.i.d.(0, 1)$ is a white noise process
- $\alpha_0 > 0$, $\alpha_i \geq 0$ (at least one $\alpha_i > 0$), $\beta_j \geq 0$
- $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$, in which $\alpha_i = 0$ for $i > m$ and $\beta_i = 0$ for $j > s$.

ARCH and GARCH model

When $\beta_j = 0$ for all $j = 1, \ldots, s$

$$GARCH(m, s) \Rightarrow ARCH(m)$$

GARCH v.s. ARCH and AR v.s. ARMA

- An ARCH model can be considered as an AR process of a_t^2.
- A GARCH model can be considered as an ARMA process of a_t^2.
- That is why we can write a higher-order ARCH(m) process with a *parsimonious* GARCH(1, 1) process.
    - What is the AR representation of ARMA?

ARMA representation of GARCH

- Let $\eta_t = \alpha_t^2 - \sigma_t^2$.
- $E(\eta_t) = 0$, $\mathrm{Cov}(\eta_t, \eta_{t-i}) = 0$ for $i \geq 1$, but usually η_t is not i.i.d.
- A GARCH(m, s) model can be written as

$$a_t^2 = \alpha_0 + \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) a_{t-i}^2 + \eta_t - \sum_{j=1}^{s} \beta_j \eta_{t-j}$$

  which can be regarded as an ARMA form for the squared series a_t^2.
- For stationarity of a_t^2, we must require that the characteristic roots of the above ARMA representation lie within the unit circle.

The Properties of GARCH(1, 1)

Consider a GARCH(1, 1)

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

where

$$\alpha_0 > 0, 0 < \alpha_1 \leq 1, 0 \leq \beta_1 \leq 1, \text{ and } \alpha_1 + \beta_1 < 1$$

The mean of a_t

- The unconditional mean: $E(a_t) = E(\sigma_t \epsilon_t) = E(\sigma_t)E(\epsilon_t) = 0$
- The conditional mean: $E_{t-1}(a_t) = \sigma_t E_{t-1}(\epsilon_t) = \sigma_t E(\epsilon_t) = 0$

The variance of ϵ_t

The conditional variance

$$E_{t-1}(a_t^2) = \sigma_t^2 = \alpha_0 + \alpha_1 \alpha_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

The unconditional variance

$$\begin{aligned}
&\alpha_t^2 = \epsilon_t^2(\alpha_0 + \alpha_1 \alpha_{t-1}^2 + \beta_1 \sigma_{t-1}^2) \\
\Rightarrow &E(a_t^2) = E(\epsilon_t^2)\left[\alpha_0 + \alpha_1 E(a_{t-1}^2) + \beta_1 E(\sigma_{t-1}^2)\right] \\
\Rightarrow &E(a_t^2) = \alpha_0 + (\alpha_1 + \beta_1) E(a_{t-1}^2)
\end{aligned}$$

Let $E(a_t^2) = E(a_{t-1}^2)$. We have

$$E(a_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

For the variance must be positive, we require $\alpha_1 + \beta_1 < 1$.

The variance of ϵ_t (cont'd)

From the ARMA representation of a GARCH(m, s) model

$$a_t^2 = \alpha_0 + \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) a_{t-i}^2 + \eta_t - \sum_{j=1}^{s} \beta_j \eta_{t-j}$$

we can also derive the unconditional variance of a stationary a_t^2 series is

$$E(a_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i)}$$

in which we must require $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$.

The autocorrelation and kurtosis

The autocorrelation function

$$E(a_t a_{t-i}) = E(\sigma_t \epsilon_t \sigma_{t-i} \epsilon_{t-i}) = 0$$

The kurtosis

Assume that $\epsilon_t \sim N(0, 1)$. Given that $1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2 > 0$, the kurtosis of a_t is

$$\frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3$$

That is, the tail distribution of a GARCH(1, 1) process is heavier than that of a normal distribution.

Volatility persistence

The roles of α_1 and β_1 in volatility persistence are different

- The larger is α_1, the larger is the response of σ_t^2 to new information.

$$\text{A shock of } \epsilon_t \to a_t \to \sigma_{t+1}^2$$

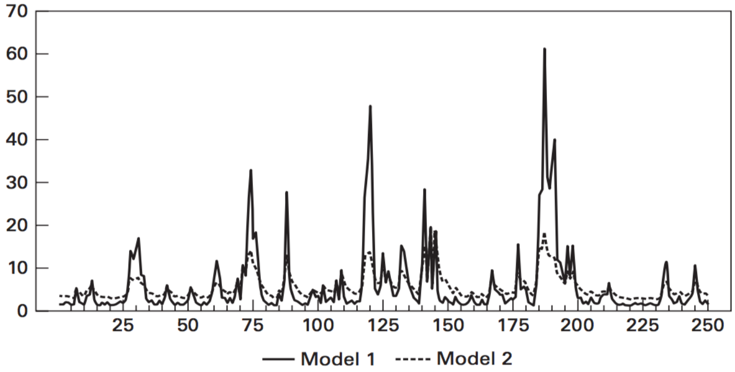
- The larger is β_1, the more persistence is the conditional variance.

$$\text{A shock of } \epsilon_t \to a_t \to \sigma_{t+1}^2 \to \sigma_{t+2}^2$$

Volatility persistence

Consider two GARCH(1, 1) models

$$\sigma_t^2 = 1 + 0.6a_{t-1}^2 + 0.2\sigma_{t-1}^2$$
$$\sigma_t^2 = 1 + 0.2a_{t-1}^2 + 0.6\sigma_{t-1}^2$$



Maximum likelihood estimation

The conditional log-likelihood function is similar to that of ARCH model

$$\ell(\boldsymbol{\alpha}, \boldsymbol{\beta} | a_1, \ldots, a_T) = \sum_{t=1}^{T} \left[-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{1}{2} \frac{a_t^2}{\sigma_t^2} \right] \tag{2}$$

The difference is that now σ_t^2 is a GARCH model

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

Check model adequacy

Compute the standardized residuals

$$\tilde{a}_t = \frac{\hat{a}_t}{\hat{\sigma}_t}$$

Check the mean equation

Use the Ljung-Box statistic for $\{\tilde{a}_t\}$.

Check the volatility equation

Use the Ljung-Box statistic for $\{\tilde{a}_t\}$.

Model diagnosis

Goodness of fit

- SSR. Since $\epsilon_t = a_t^2/\sigma_t^2$, we can compute SSR as

$$SSR = \sum_{t=1}^{T} \frac{\hat{a}_t^2}{\hat{\sigma}_t^2}$$

- The log-likelihood function.

$$2\ell = \sum_{t=1}^{T} \left[\ln(\hat{\sigma}_t^2) + \frac{\hat{a}_t^2}{\hat{\sigma}_t^2} \right] - T \ln(2\pi)$$

Information criteria

- $AIC = -2\ell + 2n$
- $BIC = -2\ell + n \ln(T)$

Forecasting

1-step-ahead forecast

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2$$

2-step-ahead forecast

$$\begin{aligned} \sigma_{h+2}^2 &= \alpha_0 + \alpha_1 a_{h+1}^2 + \beta_1 \sigma_{h+1}^2 \\ &= \alpha_0 + (\alpha_1 + \beta_1)\sigma_{h+1}^2 + \alpha_1 \sigma_{h+1}^2 (\epsilon_{h+1}^2 - 1) \end{aligned} \quad (3)$$

Given that $E(\epsilon_{h+1}^2 - 1 | F_h) = 0$, the 2-step-ahead forecast is

$$\sigma_h^2(2) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(1)$$

Forecasting (cont'd)

The ℓ-step-ahead forecast

$$\sigma_h^2(\ell) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(\ell - 1), \text{ for } \ell > 1$$

As $\ell \to \infty$

$$\sigma^2(\ell) = \frac{\alpha_0 \left[1 - (\alpha_1 + \beta_1)^{\ell-1}\right]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{\ell-1}\sigma_h^2(1)$$

Therefore,

$$\sigma^2(\ell) \to \frac{\alpha_0}{1 - \alpha_1 - \beta_1}, \text{ as } \ell \to \infty$$

provided that $\alpha_1 + \beta_1 < 1$.