Volatility Modeling: ARCH, GARCH and MMAR

Mateus Hiro Nagata

LAMFO

Outline

Warning: we are dealing with heteroskedastic things

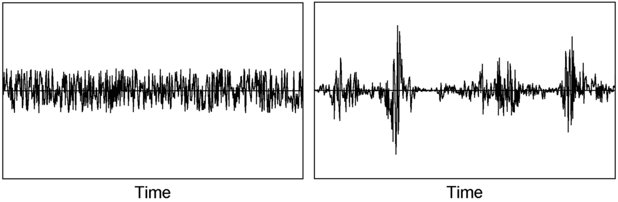

Figura: Font: https://www.value-at-risk.net/stochastic-process-heteroskedasticity/

Volatility: Stylized Facts (Tsay, 2005)

1. There exist volatility clusters
2. Volatility evolves in a continuous mannerthat is, volatility jumps are rare
3. Volatility varies within some fixed range (statistically speaking, this means that volatility is often stationary)
4. Volatility react differently to a big price increase or a big price drop, referred to as the leverage effect

Tsay (2005)

1. Specify a mean equation by testing for serial dependence in the data and, if necessary, building an econometric model (e.g., an ARMA model) for the return series to remove any linear dependence.
2. Use the residuals of the mean equation to test for ARCH effects.
3. Specify a volatility model if ARCH effects are statistically significant and perform a joint estimation of the mean and volatility equations.
4. Check the fitted model carefully and refine it if necessary.

- For most asset return series, the serial correlations are weak
- Thus, building a mean equation amounts to removing the sample mean from the data if the sample mean is significantly different from zero.

Nice to Meet You, ARCH

Autoregressive Conditional Heteroscedasticity (ARCH):

1. Autoregressive: predicts future values based on past values
2. Conditional Heteroscedasticity: nonconstant volatility related to prior periods

- Modeling the volatility of an asset return.
- **DEF.** Volatility means the conditional variance of the underlying asset return

Why Volatility?

- Volatility ⇒ value-at-risk ⇒ mean-variance framework, interval forecast
- Psychological fear
- Price volatility presents opportunities to buy assets cheaply and sell when overpriced;
- Volatility affects pricing of options, being a parameter of the BlackScholes model

ARCH

$$a_t = \sigma_t \epsilon_t, \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2$$

- a_t error term, σ_t time dependent std. dev, ϵ_t stochastic term
- Volatility Clustering: the volatility changes over time and its degree shows a tendency to persist
- $\sigma_t^2 = Var(a_t | a_{t-1}, a_{t-2}, \ldots, a_{t-p})$
- ARCH(p) $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i u_{t-i}^2$
- GARCH(p,q) $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} i u_{t-i}^2 + \sum_{i=1}^{q} \phi_i \sigma_{t-i}^2$

Estimating ARCH

1. Estimate the best fitting autoregressive model AR(q)
   $y_t = a_0 + a_1 y_{t-1} + \cdots + a_q y_{t-q} + \epsilon_t$
2. Obtain the squares of the error
3. Null Hypothesis: absence of ARCH components, we have $\alpha_i = 0$ for all $i = 1, \ldots, q$

Properties of ARCH

1. $E(a_t) = 0$
2. $Var(a_t) = E(a_t^2) = \alpha_0 + \alpha_1 E(a_{t-1}^2) = \alpha_0 + \alpha_1 Var(a_{t-1})$
3. $E(a_t^4) = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$
4. Implications a) $0 \geq \alpha_1^2 < \frac{1}{3}$
5. Implications b) Unconditional kurtosis $\frac{E(a_t^4}{[Var(a_t)]^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} > 3$ (tail heavier than Normal Distribution)

Properties of ARCH

1. $E(a_t) = 0$
2. $Var(a_t) = E(a_t^2) = \alpha_0 + \alpha_1 E(a_{t-1}^2) = \alpha_0 + \alpha_1 Var(a_{t-1})$
3. $E(a_t^4) = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$
4. Implications a) $0 \geq \alpha_1^2 < \frac{1}{3}$
5. Implications b) Unconditional kurtosis $\frac{E(a_t^4}{[Var(a_t)]^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} > 3$ (tail heavier than Normal Distribution)
6. The proofs are exercise to the reader

Weaknesses of ARCH Models

1. Assumes positive and negative shocks affects the same on volatility
2. Restrictive. $\alpha_1^2 \in [0, \frac{1}{3}]$ if has 4th moment. The constraint becomes complicated for higher order ARCH models. In practice, it limits the ability of ARCH models with Gaussian innovations to capture excess kurtosis.
3. (João critique) Does not provide any insight in the causes for the time series' behavior
4. ARCH models are likely to overpredict the volatility because they respond slowly to large isolated shocks to the return series.

Order Determination

- ARCH effect significant $\Rightarrow$ use $PACF(a_t^2)$ to determine ARCH order.
- Model Checking
- ARCH model $\Rightarrow$ standarized residuals are iid r.v.
- Check adequacy if a_t is or not r.v. $\Rightarrow$ Ljung-Box statistics
- H_0: The data are independently distributed

Keywords

1. Volatility
2. ARCH
3. GARCH

Outline

GARCH

- Maximum likelihood estimates of ARCH and GARCH models are efficient and have normal distributions in large samples, such that the usual methods for conducting inference about the unknown parameters can be applied.
- If an autoregressive moving average model (ARMA) model is assumed for the error variance, the model is a generalized autoregressive conditional heteroskedasticity (GARCH) model.

Model especification

1. Estimate the best fitting AR(q) model
2. Compute and plot the autocorrelations of t^2 by
3. The asymptotic, that is for large samples, standard deviation of $\rho(i)$is $1/\sqrt{T}$. Individual values that are larger than this indicate GARCH errors. To estimate the total number of lags, use the LjungBox test until the value of these are less than, say, 10% significant. The LjungBox Q-statistic follows χ^2 distribution with n degrees of freedom if the squared residuals ϵ_t are uncorrelated. It is recommended to consider up to T/4 values of n. The null hypothesis states that there are no ARCH or GARCH errors. Rejecting the null thus means that such errors exist in the conditional variance.

Strengths Weaknesses of GARCH

1. Large a_{t-1}^2 or $\sigma_t \Rightarrow$ large $\sigma_t^2 \Rightarrow$ Volatility clustering
2. Tail heavier than Normal Distribution
3. The model provides a simple parametric function that can be used to describe the volatility evolution
4. Responds equally to positive and negative shocks
5. In addition, recent empirical studies of high-frequency financial time series indicate that the tail behavior of GARCH models remains too short even with standardized Student-t innovations.

Prediction with GARCH

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

Types of GARCH

Nonlinear Asymmetric GARCH(1,1)

$$\sigma_t^2 = \omega + \alpha(a_{t-1} - \theta\sigma_{t-1})^2 + \beta\sigma_{t-1}^2$$

GARCH-in-mean

$$y_t = \beta x_t + \lambda\sigma_t + a_t$$

1. FIGARCH: long memory
2. MRS-GARCH: multiple regimes in the data

SVR-GARCH

Peng, Y., Albuquerque, P. H. M., de Sá, J. M. C., Padula, A. J. A., Montenegro, M. R. (2018). The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression. Expert Systems with Applications, 97, 177-192.

How to Choose GARCH(p,q)?

1. Just use GARCH(1,1)
2. Estimate all GARCH(p,q) in some range. Choose with the lowest AIC (forecasting) or lowest BIC (explanation), but... bias-variance tradeoff
3. Estimate and test for autocorrelation (Ljung-Box test) and for no remaining ARCH patterns (Li-Mak test), and other tests. Good for explanation, may not be for forecasting. See "To Explain or to Predict"
4. Train-validation, use a bunch of GARCH(1,1) models and accuracy metric (PENG et al., 2017)

Does (not) anything beat GARCH(1,1)?

1. Does Anything Beat a GARCH(1,1)? (Hansen & Lude, 2001)
2. Does anything NOT beat the GARCH(1,1)?
3. Does Anyone Need a GARCH(1,1)?

Outline

Model

1. Long-tails (Lévy-stable Distribution)
2. Long dependence (Brownian Motion)
3. Trading time property
4. Multiscaling Behavior

Multiscaling Behavior

$$X(ct) \overset{d}{=} c^H X(t)$$

$$X(ct) \overset{d}{=} M(c)X(t)$$

X, M independent random functions. Satisfies $M(ab) \overset{d}{=} M_1(a)M_2(b)$, where M_1, M_2 copies of M, which also satisfies:

$$E(|X(t)|^q) = c(q)t^{\tau(q)+1}$$

Where $c(q)$ and $\tau(q)$ are deterministic functions of q. This scaling rule is the basic property of the multifractality.

The Multifractal Model ofAsset Returns By

$$X(t) \equiv lnP(t) - lnP(0)$$

Brownian motion with the multifractal trading time, we can model the $X(t)$ process Assumptions:

1. $X(t)$ is a compounding process:
2. Trading time (t) is the cdf of the multifractal measure μ defined on interval $[0, T]$
3. $B_H(t)$ and $\theta(t)$ processes are independent.

Partition functions

$$S_q(T, \Delta t) \equiv \sum_{i=0}^{N-1} |X(i\Delta t, \Delta t)|^q$$

Despite the fact that there are temporary correlations, in conjunction with the stable increments property of a multifractal process, addends are distributed identically. When the q-th moments exist, if the X ptq is multifractal, the scaling law satisfies the condition below:

$$logE[S_q(T, \delta t)] = \tau(q)log(\delta t) + c(q)logT$$

The slope of the graph with related q orders is used to test the applicability of MMAR; in other words, the slope of these lines obtained via the OLS method gives an estimation of the scaling function $\tau(q)$.

Properties of the scaling function

1. $\tau(1/H) = 0$
2. $f(\alpha) = \inf_q[\alpha q - \tau(q)]$

The first of these two equations gives the inverse $(1/H)$ of the Hurst exponent H of the $X(t)$ process.

As for the second equation, multifractal spectrum $f(\alpha)$ is the Legendre transformation of the scaling function $\tau(q)$, that is the Legendre transformation enables the obtaining of the multifractal spectrum $f(\alpha)$.

References

1 Mandelbrot, B. B., Fisher, A. J., Calvet, L. E. (1997). A multifractal model of asset returns.
2 Tsay, R. S. (2005). Analysis of Financial Time Series. In Wiley. https://doi.org/10.1177/0951484891004003O5
3 Peng, Y., Albuquerque, P. H. M., de Sá, J. M. C., Padula, A. J. A., Montenegro, M. R. (2018). The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression. Expert Systems with Applications, 97, 177-192.
4 Hansen, P. R. (2010). A winners curse for econometric models: on the joint distribution of in-sample fit and out-of-sample fit and its implications for model selection. Research Paper, 1, 39.
5 Shmueli, G. (2010). To explain or to predict?. Statistical

References

1