13. Estimation and Extensions in the ARCH model

MA6622, Ernesto Mordecki, CityU, HK, 2006.

References for this Lecture:

Robert F. Engle. *GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics*, Journal of Economic Perspectives, 15, pp 157-168.[1]

Analysis of Financial Time Series (Chapter 3). Ruey S. Tsay. Wiley (2002) [Available Online]

---

[1]**Recommended** as "Short, readable introduction"

Plan of Lecture 13

(13a) Parameter Estimation in the ARCH model

(13b) Why the ARCH model won the Nobel Prize

(13c) Comments on the ARCH(p) model

(13d) The Generalized ARCH model (GARCH)

(13e) Combining ARMA - ARCH models

(13f) Exponential GARCH model (EGARCH)

13a. Parameter Estimation in the ARCH model

Assume that we want fo model

$$X(1), \ldots, X(n)$$

with an ARCH process with $\{\varepsilon(t)\}$ normally distributed.

We have two parameters ω, α and the estimation is performed through ML. As the variables in the ARCH model are not independent, but are conditionally independent, (i.e. given $X(t-1)$, $X(t)$ is independent) we compute the joint density conditional to the first value $X(0)$, to obtain

$$f[x(n), x(n-1), \ldots, x(1) \mid x(0)]$$
$$= f[x(n) \mid x(n-1)]f[x(n-1) \mid x(n-2)] \ldots f[x(1) \mid x(0)].$$

Here each term is the density of one value conditional on the previous one.

A more explicit notation is

$$f_{X(t)|X(t-1)}(x) = f[x \mid x(t-1)].$$

But, as

$$X(t) = \sigma(t)\varepsilon(t) = \sqrt{\omega + \alpha X(t-1)^2}\,\varepsilon(t),$$

if the value of $X(t-1) = x(t-1)$ is known, $X(t)$ has normal distribution, with zero mean and variance $\sigma(t)^2 = \omega + \alpha X(t-1)^2$. In other terms

$$f_{X(t)|X(t-1)}(x) = \frac{1}{\sigma(t)}\varphi\left(\frac{x}{\sigma(t)}\right).$$

were

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}\exp(-x^2/2).$$

Then (conditional) likelihood is the joint conditional density:

$$\mathcal{L}(\omega, \alpha) = f[x(n), x(n-1), \ldots, x(1) \mid x(0)]$$

$$= \frac{1}{\sigma(n)} \varphi\left(\frac{x(n)}{\sigma(n)}\right) \times \cdots \times \frac{1}{\sigma(1)} \varphi\left(\frac{x(1)}{\sigma(1)}\right)$$

$$= \frac{1}{(2\pi)^{n/2} \sigma(n) \ldots \sigma(1)} \exp\left[-\frac{1}{2} \sum_{t=1}^{n} \frac{x(t)^2}{\sigma(t)^2}\right]$$

Taking the logarithm, we obtain the quantity we should maximize:

$$\ell(\omega, \alpha) = -\frac{1}{2} \sum_{t=1}^{n} \left\{ \log \sigma(t)^2 + \frac{x(t)^2}{\sigma(t)^2} \right\}$$

$$= -\frac{1}{2} \sum_{t=1}^{n} \left\{ \log\left(\omega + \alpha x(t-1)^2\right) + \frac{x(t)^2}{\omega + \alpha x(t-1)^2} \right\}$$

where we discard the constant term $(n/2) \log(2\pi)$.

In this last expression the values $x(t)$ are known, and (ω, α) are the variables.

In order to find the estimators, we equate to zero the derivatives with respect to the variables:

$$\frac{\partial \ell(\omega, \alpha)}{\partial \omega} = -\frac{1}{2} \sum_{t=1}^{n} \left\{ \frac{1}{\omega + \alpha x(t-1)^2} - \frac{x(t)^2}{[\omega + \alpha x(t-1)^2]^2} \right\}$$

$$= -\frac{1}{2} \sum_{t=1}^{n} \left[\frac{1}{\sigma(t)^2} - \frac{x(t)^2}{\sigma(t)^4.} \right] = 0.$$

Similarly

$$\frac{\partial \ell(\omega, \alpha)}{\partial \alpha} = -\frac{1}{2} \sum_{t=1}^{n} x(t-1)^2 \left[\frac{1}{\sigma(t)^2} - \frac{x(t)^2}{\sigma(t)^4} \right] = 0.$$

The difference with our first Example is here we can not find explicit formulas for the estimated parameters $\bar{\omega}$ and $\bar{\alpha}$

In order to overpass this difficulty, Engle (1982) proposed a recursive method, known as the BHHH method[2]

To introduce this method we use vectorial notation:

$$\theta = (\omega, \alpha)', \qquad \text{and} \qquad \frac{\partial \ell}{\partial \theta} = \left(\frac{\partial \ell}{\partial \omega}, \frac{\partial \ell}{\partial \alpha}\right)'$$

For each t we denote

$$\ell_t(\theta) = -\frac{1}{2}\left[\log \sigma(t)^2 + \frac{x(t)^2}{\sigma(t)^2}\right]$$

$$= -\frac{1}{2}\left[\log\left(\omega + \alpha x(t-1)^2\right) + \frac{x(t)^2}{\omega + \alpha x(t-1)^2}\right].$$

In this way

$$\ell = \sum_{t=1}^{n} \ell_t. \qquad \frac{\partial \ell}{\partial \theta} = \sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}$$

[2]Berndt, E., B. Hall, R. Hall, and J. Hausman, (1974), "Estimation and Inference in Nonlinear Structural Models", Annals of Social Measurement, Vol. 3, 653-665.

The BHHH iteration method consist in the following:

Choose an initial value $\theta_0 = (\omega_0, \alpha_0)$.

At step k, compute the vector

$$v_k = \sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}(\theta_k).$$

Compute for this value the matrix

$$B = \sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}\left(\frac{\partial \ell_t}{\partial \theta}\right)'$$

Invert the matrix B,

Compute the new approximation by

$$\theta_{k+1} = \theta_k + B^{-1}v_k.$$

Stop the algorithm when the values of the estimates stabilizes.

Remarks

- The method is a modification of the Newton Raphson method

- It does not use the second derivatives

- It is very sensitive to the initial values θ_0. (remember that $\omega + \alpha < 1$)

- One can check that the initial parameters fit the sample variance and kurtosis:

$$\mathbf{var}\, X(t) = \frac{\omega}{1-\alpha} \sim \sum_{t=1}^{n} \left(X(t) - \bar{X} \right)^2 = \bar{\sigma}^2,$$

and

$$\kappa_{X(t)} = \frac{6\alpha^2}{1 - 3\alpha^2} \sim \frac{1}{\bar{\sigma}^4} \sum_{t=1}^{n} \left(X(t) - \bar{X} \right)^4 - 3.$$

13b. Why the ARCH model won the Nobel Prize

The ARCH model (and a large list of subsequent related models) produced a revolution in financial time series modelling, specially because it allows to take into account several (if not all) of the observed empirical facts, that ARMA modelling was not able to capture.

(1) Return series show little serial correlation. ARCH time series are uncorrelated.

(2) Series of absolute or squared returns show profound serial correlation. As we have seen:

$$\rho(X(t)^2, X(t-1)^2) = \alpha.$$

that is, the ARCH effect ($\alpha \neq 0$) is equivalent to this fact. We are not in the random walk hypothesis.

(3) Conditional expected returns are close to zero. And we have

checked that the conditional means vanishes:

$$\mathbf{E}(X(t) \mid \mathcal{F}(t-1)) = 0,$$

we are in the martingale hypotesis.

(4) Volatility appears to vary over time.

This is the fundamental property of ARCH modelling.

(5) Return series are leptokurtic or heavy-tailed. We have cheked that the kurtosis is always positive, and infinite if $3\alpha^2 \geq 1$ (even for gaussian noises $\{\varepsilon(t)\}$).

(6) Extreme returns apperar in clusters: volatility clustering phenomena. The volatility modelling in ARCH implies that big returns at t produces high volatility at $t+1$.

13c. Comments on the ARCH(p) model

The ARCH(p) model assumes that $\{X(t)\}$ is a stationary process satisfying:

$$X(t) = \sigma(t)\varepsilon(t),$$
$$\sigma(t)^2 = \omega + \alpha_1 X(t-1)^2 + \cdots + \alpha_p X(t-p)^2,$$

where $\{\varepsilon(t)\}$ is a strict white noise with variance one, and similar conditions on the parameters.

- This one was the model introduced by Engle (1982),

- In this context, ARCH model should be denoted ARCH(1).

- Properties and estimation are similar.

- New problem: model selection (i.e. choose p)

- In order to choose p we use criteria similar to the the Akaike criteria for ARMA(p,q) models.

13d. The Generalized ARCH model (GARCH)

The most important generalization of the ARCH model is the GARCH model, that assumes: $\{X(t)\}$ is a stationary process satisfying:

$$X(t) = \sigma(t)\varepsilon(t), \tag{1}$$

$$\sigma(t)^2 = \omega + \alpha X(t-1)^2 + \beta\sigma(t-1)^2. \tag{2}$$

introduced by T. Bollerslev[3] in (1986)[4]

The parameters should satisfy $\alpha + \beta < 1$. The value $\beta = 0$ corresponds to the ARCH model.

Tipical values encountered in GARCH modelling are $\alpha = 0.1$, $\beta = 0.85$.

We have a GARCH(p,q) model when (1) holds, and the volatility

---

[3]Tim Bollerslev was at that time a student of R. Engle.

[4]Tim Bollerslev. "Generalized Autoregressive Conditional Heteroskedasticity", Journal of Econometrics, 31:307-327, 1986.

is modelled by

$$\sigma(t)^2 = \omega + \alpha_1 X(t-1)^2 + \cdots + \alpha_p X(t-p)^2$$
$$+ \beta_1 \sigma(t-1)^2 + \cdots + \beta_q \sigma(t-q)^2.$$

The GARCH model has become a benchmark in modelling financial data, specially daily log returns.

The main advantage of GARCH over ARCH is that, ARCH(p) modelling requires relatively high values of p for good fitting, while GARCH(1,1) is usually enough for fitting financial data.

13e. Combining ARMA - ARCH models

All the description up to now refers to $\{X(t)\}$ being a weak white noise.

It is the natrual to model financial time series with and ARMA time series, with residuals of the ARCH type:

This lead us to consider time series, in the case of AR(1)-ARCH case, of the form

$$R(t) = \phi R(t-1) + \mathbf{X}(t),$$

where now the residuals $\{X(t)\}$ is an ARCH time series.

In this model, called an AR(1)-ARCH(1) model, we have 3 parameters: ϕ, ω, and α.

Similar model constructions can be carried for higher orders in both parts of the model, obtaining an ARMA(p,q)-GARCH(r,s) model.

13f. Exponential GARCH model (EGARCH)

An empirical phenomena not reported yet, that the GARCH modelling does not cover is

(7) Yesterday's rertuns $X(t-1)$ and today's volatility $\sigma(t)$ are negatively correlated. This phenomena, known as the leverage effect explains that after drops in prices an increase in volatility is expected.

In order to take into account this effect Nelson (1990)[5] propses the Exponential GARCH model, or EGARCH.

---

[5]D. B. Nelson, "Conditional Heteroskedasticity in Asset Returns: A New Approach", Econometrics, 59 (1990)

EGARCH model assumes that $\{X(t)\}$ is stationary, with

$$X(t) = \sigma(t)\varepsilon(t),$$

$$\log \sigma(t)^2 = \omega + \lambda \left| \frac{\varepsilon(t-1)}{\sigma(t-1)} \right| + \gamma \frac{\varepsilon(t-1)}{\sigma(t-1)} + \beta \log \sigma(t-1)^2.$$

An ARMA-EGARCH model has been proposed to model the HSI[6]
The values provided for the EGARCH part are

Period	ω	β	λ	γ
Before (3/93-6/97)	-0.39	0.87	0.18	-0.02
Crisis (7/97-8/98)	-0.92	0.81	0.30	-0.26
After (9/98-12/00)	-0.18	0.88	0.30	-0.12

[6]K.C.Chan, L.T.W. Cheng, P.P. Lung, "Testing the net buying pressure hypothesis during the Asian Financial Crisis: Evidence from Hang Seng Index Options, The Journal of Financial Reserarch (2006)