Financial Time Series – Parameter estimation and order selection for GARCH and ARMA-GARCH processes

Andreas Petersson

TMS088/MSA410 – May 2020

Mathematical Sciences, Chalmers University of Technology & University of Gothenburg, Sweden

GARCH models

Definition

A stochastic process $X = (X_t, t \in \mathbb{Z})$ is called a $\mathrm{GARCH}(p, q)$ *process* if it is a stationary solution to the GARCH equations

$$X_t = \sigma_t Z_t,$$

where $Z \sim \mathrm{IID}(0, 1)$,

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j X_{t-j}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2,$$

with $\alpha_0 > 0$, $\alpha_j \geq 0$ for $j = 1, \ldots, p$, $\beta_i \geq 0$ for $i = 1, \ldots, q$.

Usually $Z_t \sim \mathcal{N}(0, 1)$ or $\sqrt{\nu/(\nu-2)} Z_t \sim t_\nu$ for all $t \in \mathbb{Z}$. The factor $\sqrt{\nu/(\nu-2)}$ yields $\mathrm{Var}(Z_t) = 1$, Z_t follows a *generalized* or *non-standardized* t-distribution.

Conditional maximum likelihood estimation

- Motivation: $f_{X,Y}(x, y) = f(x|y) f_Y(y)$

Conditional maximum likelihood estimation

- Motivation: $f_{X,Y}(x,y) = f(x|y)f_Y(y)$
- For observations $(x_j, j = 1, \ldots, n)$ of $\mathbf{X}_n := (X_1, \ldots, X_n)'$ from a $\mathrm{GARCH}(p,q)$-process, we get for the density function:

$$
\begin{aligned}
&f_{\mathbf{X}_n}((x_1, \ldots, x_n)') \\
&\quad = f(x_n|x_{n-1}, x_{n-2}, \ldots, x_1) \cdots f(x_{p+1}|x_p, x_{p-1}, \ldots, x_1) f(x_p, x_{p-1}, \ldots, x_1) \\
&\quad = f(x_p, x_{p-1}, \ldots, x_1) \prod_{t=p+1}^{n} f(x_t|x_{t-1}, x_{t-2}, \ldots, x_1).
\end{aligned}
$$

Conditional maximum likelihood estimation

- Motivation: $f_{X,Y}(x,y) = f(x|y)f_Y(y)$
- For observations $(x_j, j = 1, \ldots, n)$ of $\mathbf{X}_n := (X_1, \ldots, X_n)'$ from a $\mathrm{GARCH}(p,q)$-process, we get for the density function:

$$
\begin{aligned}
&f_{\mathbf{X}_n}((x_1, \ldots, x_n)') \\
&\quad = f(x_n|x_{n-1}, x_{n-2}, \ldots, x_1) \cdots f(x_{p+1}|x_p, x_{p-1}, \ldots, x_1) f(x_p, x_{p-1}, \ldots, x_1) \\
&\quad = f(x_p, x_{p-1}, \ldots, x_1) \prod_{t=p+1}^{n} f(x_t|x_{t-1}, x_{t-2}, \ldots, x_1).
\end{aligned}
$$

- If $(\sigma^2_{p+1-j}, j = 1, 2, \ldots, q)$ is known, then $f(x_t|x_{t-1}, x_{t-2}, \ldots, x_1)$ is known (when $\alpha(1) + \beta(1) < 1$).

Conditional maximum likelihood estimation

- Motivation: $f_{X,Y}(x,y) = f(x|y)f_Y(y)$
- For observations $(x_j, j = 1, \ldots, n)$ of $\mathbf{X}_n := (X_1, \ldots, X_n)'$ from a $\mathrm{GARCH}(p,q)$-process, we get for the density function:

$$
\begin{aligned}
&f_{\mathbf{X}_n}((x_1, \ldots, x_n)') \\
&\quad = f(x_n|x_{n-1}, x_{n-2}, \ldots, x_1) \cdots f(x_{p+1}|x_p, x_{p-1}, \ldots, x_1) f(x_p, x_{p-1}, \ldots, x_1) \\
&\quad = f(x_p, x_{p-1}, \ldots, x_1) \prod_{t=p+1}^{n} f(x_t|x_{t-1}, x_{t-2}, \ldots, x_1).
\end{aligned}
$$

- If $(\sigma^2_{p+1-j}, j = 1, 2, \ldots, q)$ is known, then $f(x_t|x_{t-1}, x_{t-2}, \ldots, x_1)$ is known (when $\alpha(1) + \beta(1) < 1$).
- A zero-mean *location-scale family* random variable Y fulfills $f_{\sigma Y}(x) = \frac{1}{\sigma} f_Y(\frac{x}{\sigma})$, $\sigma > 0$

Conditional MLE for GARCH models

The conditional MLE $(\hat{\alpha}_0, \ldots, \hat{\alpha}_p, \hat{\beta}_1, \ldots, \hat{\beta}_q, \hat{\theta}_Z)$ are the values that maximize

$$L(\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \theta_Z) = \prod_{t=p+1}^{n} \frac{1}{\sigma_t} f_Z\left(\frac{x_t}{\sigma_t}\right) \tag{1}$$

where f_Z is the density of $Z_t = Z_t(\theta_Z)$.

Conditional MLE for GARCH models

The conditional MLE $(\hat{\alpha}_0, \ldots, \hat{\alpha}_p, \hat{\beta}_1, \ldots, \hat{\beta}_q, \hat{\theta}_Z)$ are the values that maximize

$$L(\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \theta_Z) = \prod_{t=p+1}^{n} \frac{1}{\sigma_t} f_Z\left(\frac{x_t}{\sigma_t}\right) \tag{1}$$

where f_Z is the density of $Z_t = Z_t(\theta_Z)$. L is an implicit function of $\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ via σ_t, which is computed recursively, supposing

$$\sigma_t = \sqrt{\hat{\sigma}^2}, t \leq 0,$$
$$X_t = 0, t \leq 0,$$

where $\hat{\sigma}^2$ is the sample variance of $\{x_1, x_2, \ldots, x_n\}$.

Conditional MLE for GARCH models

The conditional MLE $(\hat{\alpha}_0, \ldots, \hat{\alpha}_p, \hat{\beta}_1, \ldots, \hat{\beta}_q, \hat{\theta}_Z)$ are the values that maximize

$$L(\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \theta_Z) = \prod_{t=p+1}^{n} \frac{1}{\sigma_t} f_Z\left(\frac{x_t}{\sigma_t}\right) \tag{1}$$

where f_Z is the density of $Z_t = Z_t(\theta_Z)$. L is an implicit function of $\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ via σ_t, which is computed recursively, supposing

$$\sigma_t = \sqrt{\hat{\sigma}^2}, t \leq 0,$$
$$X_t = 0, t \leq 0,$$

where $\hat{\sigma}^2$ is the sample variance of $\{x_1, x_2, \ldots, x_n\}$. Equivalently, one can minimize $-\ln L(\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \theta_Z)$.

Conditional MLE for GARCH models

The conditional MLE $(\hat{\alpha}_0, \ldots, \hat{\alpha}_p, \hat{\beta}_1, \ldots, \hat{\beta}_q, \hat{\theta}_Z)$ are the values that maximize

$$L(\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \theta_Z) = \prod_{t=p+1}^{n} \frac{1}{\sigma_t} f_Z\left(\frac{x_t}{\sigma_t}\right) \tag{1}$$

where f_Z is the density of $Z_t = Z_t(\theta_Z)$. L is an implicit function of $\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ via σ_t, which is computed recursively, supposing

$$\sigma_t = \sqrt{\hat{\sigma}^2}, t \leq 0,$$
$$X_t = 0, t \leq 0,$$

where $\hat{\sigma}^2$ is the sample variance of $\{x_1, x_2, \ldots, x_n\}$. Equivalently, one can minimize $-\ln L(\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \theta_Z)$. The *residuals* $(x_t/\hat{\sigma}_t, t = p+1, p+2, \ldots, n)$ should resemble the properties of Z.
Exercise: For $Z \sim \mathrm{IID}\,\mathcal{N}(0,1)$,

$$-\ln L(\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q) = \frac{1}{2} \sum_{t=p+1}^{n} \left(\ln 2\pi + \ln \sigma_t^2 + \frac{X_t^2}{\sigma_t^2}\right).$$

GARCH squares as an ARMA process

As before, $\alpha(1) + \beta(1) < 1 \implies$

$$\mathbb{E}(X_t^2) = \alpha_0 + (\alpha(1) + \beta(1))\,\mathbb{E}(X_t^2)$$

$\implies$

$$\mathbb{E}(X_t^2) = \mathbb{E}(\sigma_t^2) = \frac{\alpha_0}{1 - \alpha(1) - \beta(1)}.$$

Exercise: If $\mathbb{E}(\sigma_t^4) < +\infty$, $X^2 = (X_t^2, t \in \mathbb{Z}) \sim \mathrm{ARMA}(\max\{p, q\}, q)$ with

$$\phi(z) = 1 - \alpha(z) - \beta(z),$$
$$\theta(z) = 1 - \beta(z)$$

and mean $\alpha_0(1 - \alpha(1) - \beta(1))^{-1}$. Alternatively:

$$X_t^2 - \sum_{i=1}^{\max\{p,q\}} (\alpha_i + \beta_i) X_{t-i}^2 = \alpha_0 + \eta_t - \sum_{j=1}^{q} \beta_j \eta_{t-j},$$

where $\alpha_i := 0$ for $i > p$, $\beta_i := 0$ for $i > q$ and $\eta_t := X_t^2 - \sigma_t^2$ is white noise.

ARMA alternative

Method (Two-pass estimation of GARCH)

Assume that a zero-mean set of observations $(x_j, j = 1, \ldots, n)$ is given. Use the ordinary maximum likelihood method to estimate the parameters of the ARMA representation for $(x_j^2, j = 1, \ldots, n)$, denoted by $\hat{\phi}_i$ and $\hat{\theta}_i$. Obtain the parameter estimates of the GARCH coefficients by setting

$$\hat{\beta}_i := \hat{\theta}_i \quad \text{and} \quad \hat{\alpha}_i := \hat{\phi}_i - \hat{\theta}_i.$$

Order selection

For conditional maximum likelihood:

Method (AICC criterion)

Choose $p, q, \alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ and θ_Z to minimize

$$-2\frac{n}{n-p}\ln L(\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \theta_Z) + 2n\frac{m+1}{n-m-2},$$

where n is the sample size and m is the number of non-zero estimated parameters.

ARMA–GARCH processes

Definition

A time series X is said to be an $\mathrm{ARMA-GARCH}$ *process* if it is an $\mathrm{ARMA}(p,q)$ process driven by GARCH, noise, i.e., if it is stationary and

$$X_t - \sum_{j=1}^{p} \phi_j X_{t-j} = \sigma_t Z_t + \sum_{j=1}^{q} \theta_j \sigma_{t-j} Z_{t-j}, \tag{2}$$

where $Z \sim \mathrm{IID}(0,1)$, the polynomials $(1 - \sum_{j=1}^{p} \phi_j z^j)$ and $(1 + \sum_{j=1}^{q} \theta_j z^j)$ have no common zeros and

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{\tilde{p}} \alpha_j \sigma_{t-j}^2 Z_{t-j}^2 + \sum_{i=1}^{\tilde{q}} \beta_i \sigma_{t-i}^2,$$

with $\alpha_0 > 0$, $\alpha_j \geq 0$ for $j = 1, \ldots, \tilde{p}$, $\beta_i \geq 0$ for $i = 1, \ldots, \tilde{q}$. Furthermore, it is called an $\mathrm{ARMA\text{-}GARCH}$ *process with mean* μ if $X - \mu$ is an $\mathrm{ARMA\text{-}GARCH}$ process.

A Gaussian noise example

If $Z \sim \mathrm{IID}\,\mathcal{N}(0,1)$ and if $\phi(z) \neq 0$ and $\theta(z) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| \leq 1$ and if $\alpha(1) + \beta(1) < 1$, then *conditioned on* $(X_{t-j}, j \in \mathbb{N})$:

$$X_t \sim \mathcal{N}\left(\sum_{j=1}^{p} \phi_j X_{t-j} + \sum_{j=1}^{q} \theta_j \sigma_{t-j} Z_{t-j}, \alpha_0 + \sum_{j=1}^{\tilde{p}} \alpha_j \sigma_{t-j}^2 Z_{t-j}^2 + \sum_{i=1}^{\tilde{q}} \beta_i \sigma_{t-i}^2\right)$$

for all $t \in \mathbb{Z}$. X is a *conditional mean/variance model*.

Model building for an ARMA–GARCH process

- Remove trend and seasonality
- If the conditional mean is non-zero, fit an ARMA process to the data
- Determine if the residuals look like a GARCH process
- Order selection for the residuals
- Final model estimation by conditional maximum likelihood estimation
- Check final goodness of fit