Financial Econometrics

- Week 2, Lecture 2 -

Financial Analysis of ARCH and GARCH Models

VU Econometrics and Data Science
2024-2025

Paolo Gorgi

Financial analysis of GARCH

Question: What have we done until now?

1. Introduced ARCH and GARCH models;
2. Derived their stochastic properties;
3. Learned how to estimate parameters by Maximum Likelihood.

Question: Why is this useful?

Answer: Financial analysis:

1. Filtering conditional variance;
2. Calculating risk measures;
3. Forecasting risk.

Today's class

1. Filtering volatility
2. Diagnostic analysis and model selection
    - Diagnostic tests
    - Model selection
3. Measuring financial risk and forecasting
    - Value-at-Risk
    - Forecasting
    - News impact curve

Filtering volatility

Filtering the conditional volatility

Question: Can we find an "estimate" (*filter*) of the unknown sequence of conditional variances $\{\sigma_t^2\}_{t=1}^T$?

Answer: Yes! Run the updating equation under $\hat{\theta}_T$.

Note: we have already learned how to estimate θ_0.

Notation: filtered sequence $\{\hat{\sigma}_t^2\}_{t=1}^T$

$$\hat{\sigma}_t^2 = \hat{\omega} + \hat{\beta}_1 \hat{\sigma}_{t-1}^2 + \hat{\alpha}_1 y_{t-1}^2, \qquad \text{for} \quad t = 2, \dots, T.$$

Note 1: $\hat{\sigma}_1^2$ can be set equal to sample variance;

Note 2: Sample variance of first few observations may even be better!

Filtering volatility with R (i)

Estimating the conditional variance with R.

Note: R code available in `analysis_GARCH.R`

First: Take the parameter estimates `omega_hat`, `alpha_hat` and `beta_hat` and the data `x`.

Next: define filter vector `sigma2`

```
n <- length(x)
sigma2 <- rep(0,n)
sigma2[1] <- var(x)
```

Note: You may consider a different initial value for $\hat{\sigma}_1^2$.

Filtering volatility with R (ii)

Finally: We are now ready to filter the conditional variance using a *for loop*

```
for(t in 2:n){
   sigma2[t] = omega_hat + alpha_hat*x[t-1]^2 +
                    beta_hat*sigma2[t-1]
}
```

Note 1: The filter uses the estimated parameters;

Note 2: You can also obtain an estimate of σ_t^2 at time $T+1$ (`n+1`) since `sigma2[n]` and and `x[n]` are available;

Note 3: The value σ_{T+1}^2 is a forecast of the conditional variance!

Filtering volatility with R (iii)

Stacking the code together we have the script:

```
n <- length(x)
sigma2 <- rep(0,n)
sigma2[1] <- var(x)

for(t in 2:n){
   sigma2[t] = omega_hat + alpha_hat*x[t-1]^2 +
                    beta_hat*sigma2[t-1]
}
```

Filtering volatility with R (iv)

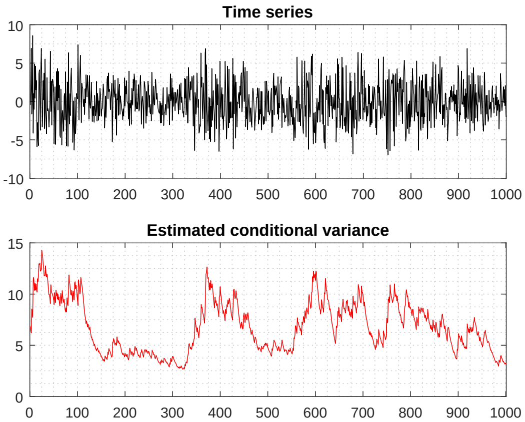

Figure: Time series (first plot) and estimated filtered variance (second plot).

Diagnostic analysis and model selection

Diagnostic tests (i)

Question: Can we trust the filtered variance?
Is it accurate? Valid even?

Answer: Maybe!

Problem: GARCH model may not be appropriate:

1. Observation equation may be too simplistic;
2. Gaussianity assumption may fail;
3. Updating equation may be incorrect;
4. Number of lags may be insufficient.

Solution: Test model specification with *diagnostic tests!*

Diagnostic tests (ii)

ARCH and GARCH models: we can use the residuals $u_t = y_t/\hat{\sigma}_t \approx \epsilon_t$ to test for correct model specification.

Specification tests: fall into two main categories:

1. ***Homoscedasticity tests***:
   - ϵ_t is assumed to have fixed conditional variance (ϵ_t *is iid*);
   - **Hence:** residuals $\{u_t\}_{t=1}^{T}$ should have fixed conditional variance!
2. ***Normality tests***:
   - ϵ_t is assumed to have a normal distribution;
   - **Hence:** residuals $\{u_t\}_{t=1}^{T}$ should have a normal distribution!

Homoscedasticity tests (i)

Simple homoscedasticity test:

- Plot ACF of squared residuals $\{u_t^2\}_{t=1}^{T}$;
- Verify squared residuals are uncorrelated by looking at the ACF.

In R:

1. Obtain the residual vector `u`

```
u <- x/sqrt(sigma2)
```

2. Plot the ACF of squared residuals

```
acf(u^2, main="")
```

3. Verify that `ACF` vector is within the 95% confidence intervals.

Homoscedasticity tests (ii)

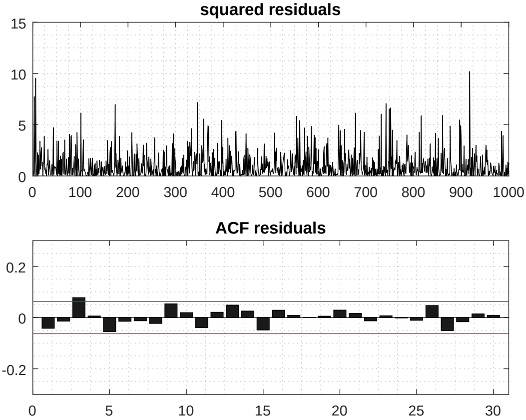

Figure: Squared residuals (first plot) and autocorrelation function of squared residuals (second plot).

Normality test

Normality tests: *Jarque-Bera* (JB) test

JB test statistic: $$JB = \frac{T+1}{6}\left(\hat{\mu}_3^2 + \frac{1}{4}(\hat{\mu}_4 - 3)^2\right).$$

- $\hat{\mu}_3$ is sample skewness and $\hat{\mu}_4$ is sample kurtosis
- H_0 : residuals are normal , H_1 : residuals not normal
- $JB \sim \chi^2_2$ under H_0 (Chi-square distribution with 2 df).

In R:

```
jarque.bera.test(u)
```

- The output gives you the p-value of the test along with other information.

Model selection

Important: In practice we need to choose the ARCH or GARCH model that best describes the data. In other words, we need to decide the order p and q of a GARCH(p,q).

Question: Can we just look at diagnostic tests to choose between a GARCH(1,1) and an ARCH(3) for instance? **Answer:** No!

- Several nested models may be well specified;
- Several competing models may seem well specified!

Solution: We can use **model selection criteria**!

Note: The log-likelihood offers natural comparison term;

However: Model selection by comparing log-likelihoods leads to overfitting (intro Econometrics).

Overfitting with log-likelihood

Overfitting: Larger nested models always increase log-Likelihood.

Example: sample log-likelihood of GARCH(2,2) is always larger than that of GARCH(1,1)... *even if the GARCH(1,1) is the correct model!!*

Problem: *sample* log-likelihood increases *not* because the model is better, but simply because it is able to *overfit* the data!

Solution: penalize the number of parameters in the model.

Note: This is the main idea behind the majority of the so-called *information criteria*.

Information Criteria

***Akaike's information criterion* (AIC):**

$$\mathrm{AIC} = 2k - 2\log L(y_1, ..., y_T; \hat{\theta}_T) \;,$$

***Bayesian Information Criterion* (BIC)**

$$\mathrm{BIC} = \log(T)k - 2\log L(y_1, ..., y_T; \hat{\theta}_T) \;.$$

Notes:

1. Both the AIC and BIC are based on a negative log-likelihood;
2. Lower value of the criterion indicates better model;
3. AIC and BIC form *reasonable* basis for model selection.

Measuring financial risk and forecasting

Value-at-Risk (i)

Financial analysis: focus on **risk measures**.

***Value-at-Risk* (VaR):** is a popular risk measure!

Idea of VaR: the daily α-VaR is the minimum amount the investor stands to loose with probability α in one day.

Example: If a portfolio has a daily 10%-VaR of 1 million euros. Then, there is a 10% probability that the value of the portfolio will fall by more than 1 million euros in one day.

Important: the VaR is typically stated in *percentage loss*:

- If a portfolio has a daily 5%-VaR of 17%, THEN there is a 5% probability that the value of the portfolio will fall by more than 17% of its value in one day.

Value-at-Risk (ii)

Mathematically: given a % return $y_t = (p_t - p_{t-1})/p_{t-1}$, the 5%-VaR is defined as the value c that satisfies

$$P(y_t \leq c) = 0.05.$$

Note 1: in other words, the α-VaR is the **quantile** of level α of y_t

Note 2: in practice we can use log-returns since log-returns are a good approximation of % returns, i.e.

$$y_t = \log(p_t) - \log(p_{t-1}) \approx (p_t - p_{t-1})/p_{t-1}.$$

Conditional VaR

Important: GARCH models allow us to obtain a **conditional VaR** at each time point $t+1$.

GARCH: we know that $y_{t+1}|Y^t \sim N(0, \sigma^2_{t+1})$.

As a result: the **conditional VaR** at time $t+1$ (α-VaR$_{t+1}$) is the quantile of $y_{t+1}|Y^t$.

- The α-VaR$_{t+1}$ is the value q that satisfies $P(y_{t+1} \leq q|Y^t) = \alpha$,
- Therefore, the α-VaR$_{t+1}$ is:

$$\alpha\text{-VaR}_{t+1} = z_\alpha \sigma_{t+1},$$

where z_α is the quantile of level α of the standard normal.

Conditional VaR in R (i)

Conditional VaR with R: `analysis_GARCH.R` calculates the α-VaR for $\alpha = 0.1$, 0.05, and 0.01

First: Obtain the conditional variance `sigma2`

Finally: Use R's quantile function for the Normal distribution `qnorm()` to obtain the desired quantiles

```
VaR10 <- qnorm(0.1,0,sqrt(sigma2))
VaR05 <- qnorm(0.05,0,sqrt(sigma2))
VaR01 <- qnorm(0.01,0,sqrt(sigma2))
```

Conditional VaR with R (ii)

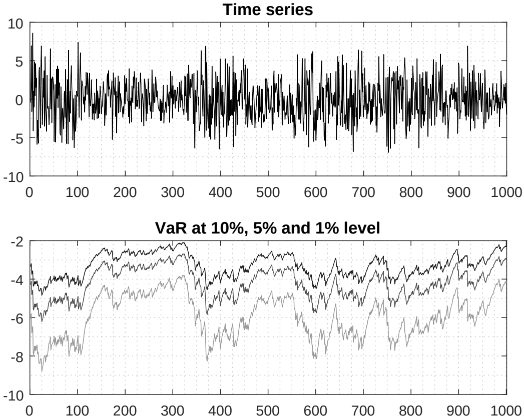

Figure: Conditional α-VaR for $\alpha = 0.1$, $\alpha = 0.05$, and $\alpha = 0.01$, estimated from a GARCH(1,1) model.

Forecasting conditional volatility (i)

Important: we are often interested in forecasting the risk of financial assets. In the following, we shall see how to **forecast volatility**.

Problem: assume we are at time T and we want to forecast the variance of y_{T+h} for $h = \{1, 2, \dots\}$.

Solution: we can use the conditional variance of y_{T+h} given Y^T as forecast. The variance forecast h steps ahead is

$$\sigma_T^2(h) = \mathbb{V}\mathrm{ar}(y_{T+h}|Y^T).$$

Forecasting conditional volatility (ii)

The forecast $\sigma_T^2(h)$ is given by:

$$
\begin{aligned}
\sigma_T^2(h) = \mathbb{V}\mathrm{ar}(y_{T+h}|Y^T) &= \mathbb{E}(y_{T+h}^2|Y^T) && (\mathbb{E}(y_{T+h}|Y^T) = 0)\\
&= \mathbb{E}(\sigma_{T+h}^2\epsilon_{T+h}^2|Y^T) && (\text{by definition})\\
&= \mathbb{E}(\sigma_{T+h}^2|Y^T) \times \mathbb{E}(\epsilon_{T+h}^2|Y^T) && (\epsilon_{T+h} \perp \sigma_{T+h}^2)\\
&= \mathbb{E}(\sigma_{T+h}^2|Y^T) \times \mathbb{E}(\epsilon_{T+h}^2) && (\epsilon_{T+h} \perp Y^T)\\
&= \mathbb{E}(\sigma_{T+h}^2|Y^T) && (\epsilon_{T+h} \sim N(0,1))
\end{aligned}
$$

Conclusion: we must obtain $\mathbb{E}(\sigma_{T+h}^2|Y^T)$

- **(for $h = 1$)** $\sigma_T^2(1) = \mathbb{E}(\sigma_{T+1}^2|Y^T) = \sigma_{T+1}^2$ (σ_{T+1}^2 is constant given Y^T)
- **(for $h > 1$)** $\sigma_T^2(h) = \mathbb{E}(\sigma_{T+h}^2|Y^T)$ depends on the model we use!

Forecasting conditional volatility: ARCH(1)

ARCH(1) model:

$$\begin{aligned}\sigma_T^2(h) = \mathbb{E}(\sigma_{T+h}^2|Y^T) &= \mathbb{E}\Big(\omega + \alpha_1 y_{T+h-1}^2|Y^T\Big)\\ &= \omega + \alpha_1\mathbb{E}(y_{T+h-1}^2|Y^T)\\ &= \omega + \alpha_1\mathbb{E}(\sigma_{T+h-1}^2|Y^T)\\ &= \omega + \alpha_1\sigma_T^2(h-1)\end{aligned}$$

Therefore: The forecast is given by the following recursion

$$\sigma_T^2(h) = \omega + \alpha_1\sigma_T^2(h-1),$$

where the recursion is initialized at $\sigma_T^2(1) = \sigma_{T+1}^2$.

Forecasting conditional volatility: GARCH(1,1)

GARCH(1,1) model:

$$\sigma_T^2(h) = \omega + (\alpha_1 + \beta_1)\sigma_T^2(h-1),$$

where the recursion is initialized at $\sigma_T^2(1) = \sigma_{T+1}^2$.

Note:

1. $\sigma_T^2(h)$ converges to the unconditional variance as $h \to \infty$
   1. GARCH(1,1): $\lim_{h\to\infty} \sigma_T^2(h) = \omega/(1-\beta_1-\alpha_1)$;
   2. In practice: $\sigma_T^2(h) \approx \omega/(1-\beta_1-\alpha_1)$ for large h.
2. The true $\sigma_T^2(h)$ cannot be obtained because θ_0 is unknown
   - In practice: we use the estimate $\hat{\sigma}_T^2(h)$ evaluated at $\hat{\theta}_T$.

Forecasting the conditional density (i)

Recall: ARCH and GARCH describe conditional density of y_t given Y^{t-1}

$$y_t|Y^{t-1} \sim N(0, \sigma_t^2) \ .$$

Question: what is the density of y_{T+1} conditional on the sample $y_1, ..., y_T$?

Answer:

$$y_{T+1}|Y^T \sim N(0, \sigma_{T+1}^2) \ .$$

Note: forecasted density for time $T+1$ is easy to obtain because σ_{T+1}^2 is given!

Forecasting the conditional density (ii)

Note: the conditional density is intractable for $h > 1$

$$y_{T+2}|Y^T = \sigma_{T+2}\epsilon_{T+2}|Y^T.$$

- $\sigma_{T+2}\epsilon_{T+2}|Y^T = (\sigma_{T+2}|Y^T) \times \epsilon_{T+2}$;
- ϵ_{T+2} is Gaussian;
- $\sigma_{T+2}|Y^T = \sqrt{\omega + \alpha_1 y_{T+1}^2 + \beta\sigma_{T+1}^2}|Y^T$;
- $y_{T+1}^2|Y^T$ has a generalized χ^2 distribution;
- Hence: $\sigma_{T+2}\epsilon_{T+2}|Y^T$ is the product between the square root of a generalized χ^2 and a normal... :(

Note: Conditional densities for $h > 1$ can be obtained by simulations (*here we focus on one-step-ahead forecasts*).

News Impact Curve

***News impact curve* (NIC):**

- Interesting piece of information obtained immediately upon estimating the parameters of an ARCH or GARCH model;
- Commonly used to describe volatility dynamics.

Definition: the NIC is the updating function that maps values of y_t to values of σ^2_{t+1}.

In essence: we fix σ^2_t to some value and look at the GARCH update as a function of y_t only.

News Impact Curve: GARCH(1,1) plot

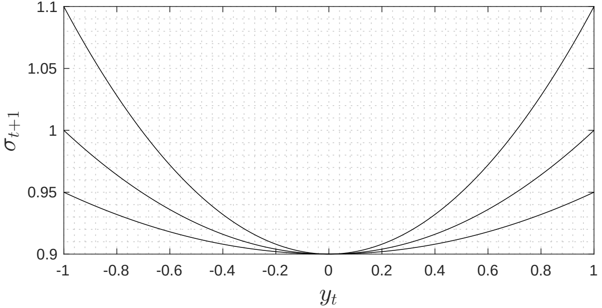

Figure: GARCH(1,1) NIC for $\omega = 0.1$, $\beta = 0.8$, and $\alpha = 0.05, 0.1, 0.2$.

Note: small absolute value of y_t lead to a decrease in conditional volatility

Note: large values of $|y_t|$ lead to explosive increase in the conditional volatility

News Impact Curve: Extensions

Advanced Econometrics (Master's course):

1. Robust NIC;
2. Leverage effects;
3. Breaks;
4. General nonlinear filter.