Applied Economic Forecasting

9. G(ARCH) Models

1. Introduction
2. GARCH Models
3. GARCH Models in R
4. A Quick Application
5. Additional Notes

Section 1

Introduction

Market Returns

Recall, the price returns of a series can be computed as

$$R_t = \frac{(P_t - P_{t-1})}{P_{t-1}} \approx \log\left(\frac{P_t}{P_{t-1}}\right)$$

Where P_t is the current price of the stock (for example) and P_{t-1} is yesterday's price.

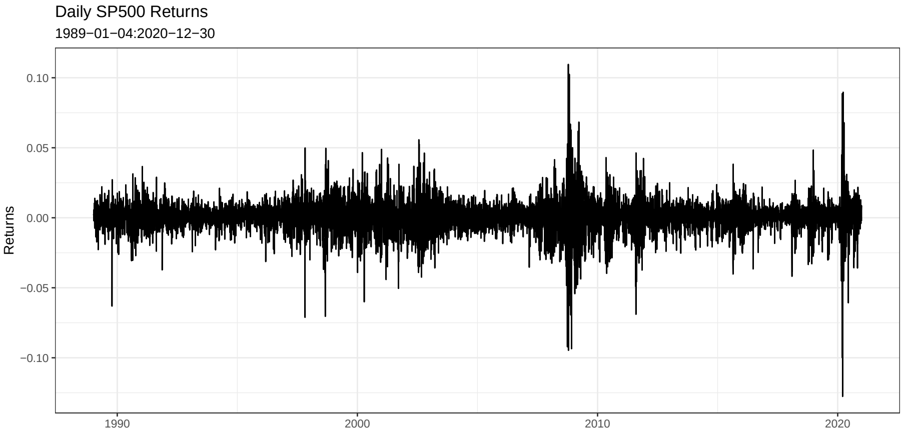
The S&P 500

```
getSymbols(Symbols = "^GSPC", src = "yahoo", from = "1989-01-04",
  to = "2020-12-30")

sp500 <- GSPC$GSPC.Close %>% log %>% diff()
 sp500 %>% autoplot() +
  labs(title = "Daily SP500 Returns",
    subtitle = "1989-01-04:2020-12-30",
    y = "Returns", x = "")
```

Market Returns

The S&P 500



Annualized Volatility

The S&P 500

Key Points

- SP500 Returns are mean zero.
- The variability of the price return changes over the sample period.
  - We can use the standard deviation (σ) to measure this variability (volatility) over time.

```
sd(sp500, na.rm = TRUE) -> sdsp500
```

- Over the full sample, the standard deviation of returns was approximately 1.14% per day.
- We can annualize the daily volatility by multiplying σ by the square root of the number of trading days (252).

That is:

```
sdsp500 * sqrt(252) -> annual.sdsp500
```

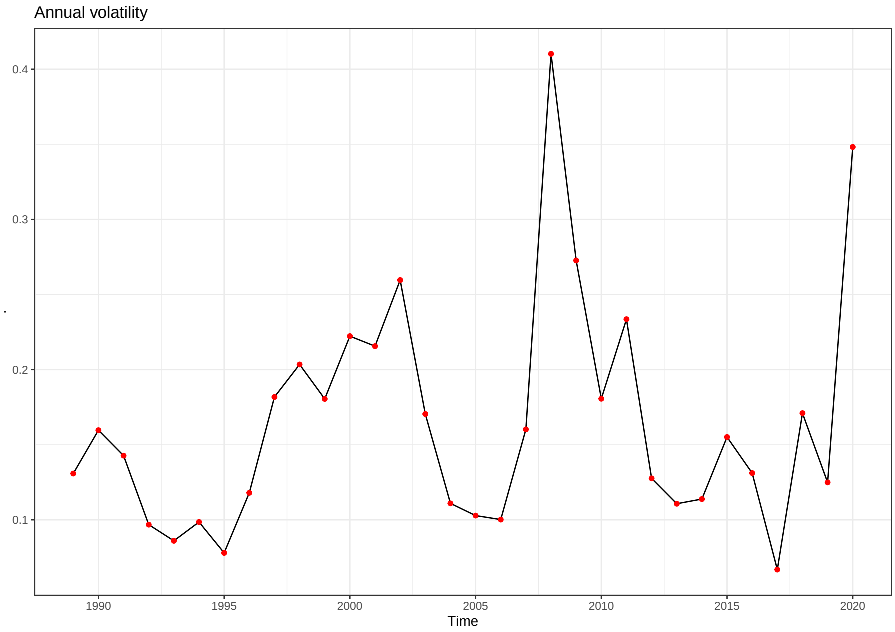
The SP500 has an annualized volatility of 18.11%.

Volatility across time

The S&P 500

Instead of the static approach above, and realizing that the volatility varies over time, we could compute the standard deviation across years instead.

```
years <- 1989:2020
sd.t <- c()
for(i in seq_along(years)){
  sd.t[i] <- sd(sp500[paste0("", years[i],"")],
                na.rm = TRUE)*sqrt(252)
}
ts(sd.t, start= 1989) %>% autoplot() +
  labs(title = "Annual volatility") +
  geom_point(color = "red")
```


Annual volatility
0.4
0.3
0.2
0.1
1990
1995
2000
2005
2010
2015
2020
Time

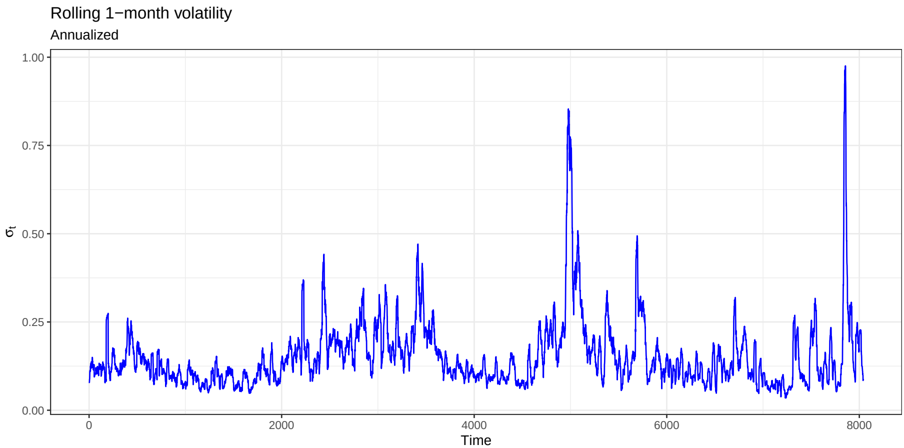
Rolling Volatility

The S&P 500

Conversely, we can compute the volatility using a rolling window approach. In the example, we will compute the 1-month rolling volatility. We will use a standard 21-days window (the average number of trading days within a month). We could just as easily compute the 2-month (42-days window) or 3-month (63-days window) rolling volatility.

In each step, we drop the oldest observation and add the most recent observation then reestimate the volatility.

```
sd.roll <- c()
for(i in 1:(length(sp500)-20)){
  sd.roll[i] <- sd(sp500[i:(i+20)],na.rm = TRUE)*sqrt(252)
}
sd.roll %>% ts() %>% autoplot(colour = "blue") +
  labs(title = "Rolling 1-month volatility",
       subtitle = "Annualized", y= bquote(sigma[t]))
```



The observations above form the motivation for the GARCH Model framework.

Section 2

GARCH Models

Conditional vs Unconditional Variance

The **unconditional variance** is the standard measure of the variance.

$$var(x) = \mathbf{E}(x - \mathbf{E}(x))^2$$

The **conditional variance** is the **true measure of our uncertainty** about a variable given a model and the information set Ω

$$\text{cond. } var(x) = \mathbf{E}(x - \mathbf{E}(x|\Omega))^2$$

We will discuss the information matrix in detail shortly.

Motivation

From our calculations earlier, using the standard deviation σ as our volatility measures is backward looking. That is, we are using past data to understand what the volatility **was**.

GARCH models are a bit more flexible and allow us to predict future volatility as well.

Useful Notations

Ω_{t-1}: The full set of information known at time t-1. For example, all the return values known (observed) at time t-1. These would include $R_{t-1}, R_{t-2}, R_{t-3}, \ldots$

$\mu_t = \mathbf{E}(R_t|\Omega_{t-1})$: This tells us that the prediction of the returns in time t is the expected value of R_t conditional on the information 1 period earlier, $t-1$.

Prediction error: $e_t = R_t - \mu_t$

Predicted variance:

$$\sigma_t^2 = var(R_t|\Omega_{t-1})$$
$$\sigma_t^2 = \mathbf{E}(R_t - \mu_t|\Omega_{t-1})$$
$$\sigma_t^2 = \mathbf{E}(e_t^2|\Omega_{t-1})$$
$$\sigma = \sqrt{\sigma_t^2}$$

Modeling the Mean

When we are coding, we will need to decide on a formula to replace the expectation formula above.

We could find μ_t using a rolling mean model where

$$\mu_t = \frac{1}{m}\sum_{i=1}^{m} R_{t-i}$$

- We could also achieve this using our ARMA models from class.

Modeling the Variance

In the case of the variance, you can take the average of the m most recently observed squared prediction errors.

That is,

$$\sigma_t = \frac{1}{m}\sum_{i=1}^{m} e_{t-i}^2$$

Note

All m observations are equally weighted in this approach regardless of when they are observed.

We would expect, however, that the future variance is more affected by the more recent events than by those in the distant past. Therefore, we can achieve a higher forecasting accuracy by giving more weight to the most recent observations (think of an exponential smoothing type of approach).

This obvious shortcoming motivates the use of an ARCH model.

Autoregressive Conditional Heteroskedasticity (ARCH)

ARCH(p) Model Specification

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i e_{t-i}^2$$

In an ARCH equation, the predicted variance is the sum of a constant and a weighted sum of `p` lagged observed squared prediction errors.

- If there is an ARCH effect, it can be tested by the statistical significance of the estimated coefficients.
- If they are significantly different from zero, we can conclude that there is an ARCH effect.

Generalized ARCH (GARCH)

In practice, we often use a GARCH(1,1) model for our empirical analysis of market and returns volatility.

$$\sigma_t^2 = \omega + \alpha e_{t-1}^2 + \beta \sigma_{t-1}^2$$

Parameter Restrictions

- $\omega, \alpha, \beta > 0$: this ensures that the variances are positive at all times.
- $\alpha + \beta < 1$: this ensures that we have stability in the system. A shock to the system (through α) will die out over time. That is the predicted variance, σ_t^2 will always return to its long run mean.
    - This implies that our variance is mean reverting

$$\text{LR var} = \frac{\omega}{1 - \alpha - \beta}$$

Assessing the LR Variance

Suppose

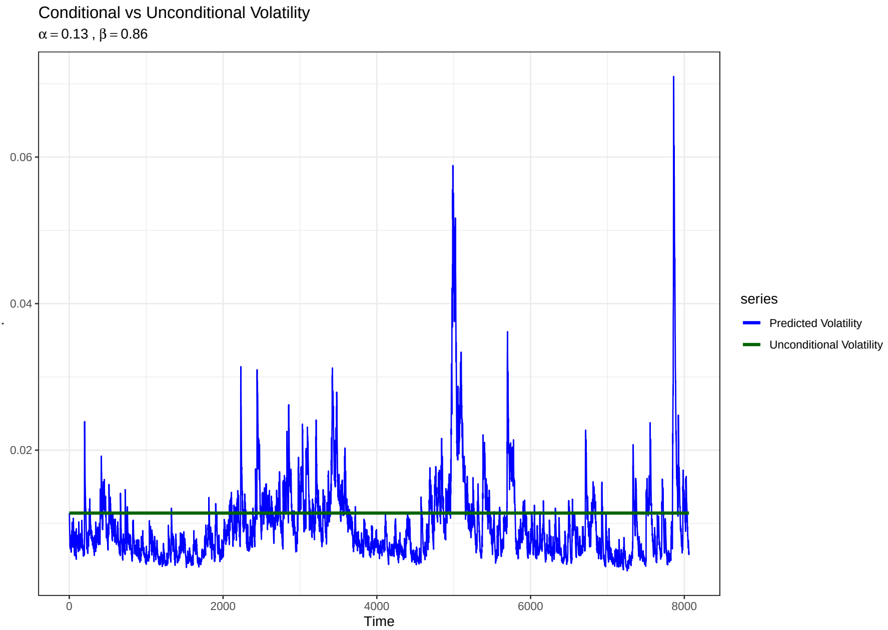
$$\sigma_t^2 = 0.000002 + 0.13e_{t-1}^2 + 0.86\sigma_{t-1}^2$$

The LR variance is

$$\frac{\omega}{1-\alpha-\beta} = \frac{0.000002}{1-0.13-0.86} = 0.0002$$

hence the LR volatility per day is $\sqrt{0.0002} \times 100 = 1.4\%$.

LET US PRACTICE IN R!


Conditional vs Unconditional Volatility
$\alpha = 0.13$, $\beta = 0.86$
0.06
0.04
0.02
0
2000
4000
6000
8000
Time
series
Predicted Volatility
Unconditional Volatility

Section 3

GARCH Models in R

Estimating a GARCH Model

Take the GARCH (1,1) model

$$
\begin{aligned}
R_t &= \mu + \varepsilon_t \\
\varepsilon_t &\sim \mathbf{N}(0, h_t) \\
h_t &= \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}
\end{aligned}
$$

- Notice that I replaced σ^2 from earlier with h_t since that is how many books define the variance of the residuals.
- There are 4 parameters to be estimated in this model: μ,ω,α, and β.
- A popular approach is to use the method of Maximum likelihood (MLE) to calculate the values of these parameters.

Estimating a GARCH Model (MLE)

Probability density function (pdf) for a normal distribution is:

$$f(x_i; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

The likelihood function is therefore:

$$\mathcal{L}(\mu, \sigma) = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right]$$

Taking the logs:

$$\log \mathcal{L} = -\frac{n}{2} log(2\pi) - \frac{n}{2} log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Specific to our model, we have:

$$\log \mathcal{L} = -\frac{T}{2} log(2\pi) - \frac{T}{2} \log h_t - \frac{1}{2h_t}\sum_{i=1}^{T}\varepsilon^2$$

where $\varepsilon = y_t - \alpha - \sum_{i=1}^{p} \phi_i y_{t-1}; \quad h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \theta_j h_{t-j}$

For our purposes, we will use the package `rugarch` for our model estimations.

rugarch in Action

The Workflow is as follows:

1. Use the function **ugarchspec** to specify the desired GARCH model

```
specs <- ugarchspec(mean.model = list(armaOrder = c(0,0)),
                    variance.model = list(model = "sGARCH"),
                    distribution.model = "norm")
```

2. Pass the arguments in Step 1 to the **ugarchfit** command.

```
fit.garch <- ugarchfit(spec = specs, data = sp500[-1,])
```

3. Use the **ugarchforecast** command make predictions about the future volatility of our returns series.

```
fore.fit <- ugarchforecast(fit.garch, n.ahead = 7)
```

Useful Commands

We can extract various results from our fitted model. A few useful commands are:

coef(): extracts the model coefficients

```
coef(fit.garch) %>% t() %>% format(digits = 3) %>% knitr::kable()
```

mu	omega	alpha1	beta1
6.01e-04	2.02e-06	1.05e-01	8.78e-01

uncvariance: delivers the **unconditional** Variance

```
unvar <- uncvariance(fit.garch)
cat("The Long run volatility of the S&P500 is",
    unvar %>% sqrt()*100,"%")
```

```
## The Long run volatility of the S&P500 is 1.099727 %
```

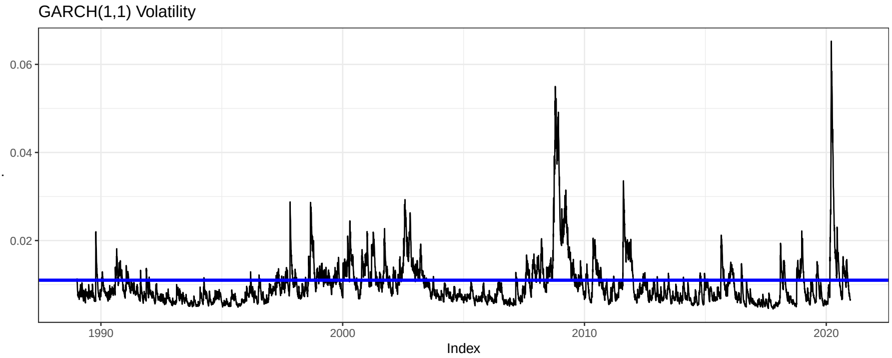
Useful Commands

fitted: extracts the predicted mean

```r
fitted(fit.garch)
```

sigma: extracts the predicted volatilities

```r
sigma(fit.garch) %>% autoplot() + geom_hline(yintercept = sqrt(unvar),
  col = 'blue', lwd = 1.2, series = "LR") + labs(title = "GARCH(1,1) Volatility")
```



Forecasting with GARCH Models

Example with a GARCH(1,1) model:

$$h_t = \omega + \alpha u_{t-1}^2 + \beta_1 h_{t-1}$$

with unconditional variance $\sigma^2 = \dfrac{\omega}{1 - \alpha - \beta}$

write:

$$h_t = \sigma^2 + \alpha\big(u_{t-1}^2 - \sigma^2\big) + \beta_1\big(h_{t-1} - \sigma^2\big)$$
$$h_{t+s} = \sigma^2 + \alpha\big(u_{t+s-1}^2 - \sigma^2\big) + \beta_1\big(h_{t+s-1} - \sigma^2\big)$$

Then the predicted h_{t+s} is:

$$h_{t+s} = \omega + (\alpha + \beta)\big(h_{t+s-1} - \sigma^2\big)$$

Section 4

A Quick Application

A Portfolio Allocation Problem

Scenario

Assume that you are an investor who is interested in investing in simple two asset portfolio.

- a risky asset (in this case the SP500) and
- a risk free asset such as the U.S. Tbills.

Assume further that, based on your risk tolerance, you would like to target a 6% annualized volatility in your portfolio.

Step 1: Compute the Annualized volatility, `vol.annual`, as implied by the GARCH(1,1) model.

Step 2: Compute the weights using the formula

$$w_{sp500} = \frac{\text{Target Volatility}}{\text{vol.annual}}$$

Section 5

Additional Notes

Digression: Modeling conditional means and variances

Given the model

$$Y = \alpha + \beta\varepsilon$$

if $\varepsilon \sim \mathbf{N}(0,1)$ then

$$\begin{aligned}\mathbf{E}(Y) &= \mathbf{E}\alpha + \beta\mathbf{E}(\varepsilon)\\ &= \alpha\end{aligned}$$

and

$$\begin{aligned}var(Y) &= var(\alpha) + \beta^2 var(\varepsilon)\\ &= \beta^2\end{aligned}$$

To model the conditional mean of Y_t given

$$X_t = \{X_{jt}\}^{j=1,2,\ldots,k}$$

Model Building

Building a volatility model consists of four steps:

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.