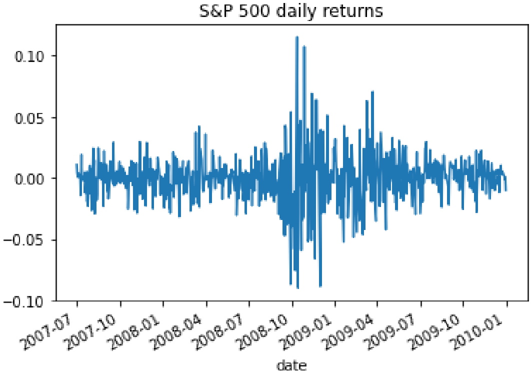
Volatility modeling

William Mann

EMORY
GOIZUETA BUSINESS SCHOOL

Motivation: Time-varying, persistent volatility



- Many datasets, especially financial data, exhibit time-varying and persistent volatility. We'd like to model this for many reasons, especially risk management.
- Note that it is really the *persistence* that is important to model.

Outline

- Engle's ARCH-LM test for persistence in volatility.
- Modeling persistence in volatility with GARCH.
- Application: Value-at-risk calculations.

A general theme is to use r_t^2 as an estimate of σ_t^2, and adapt the tools that we already know from prior weeks around this idea.

Engle's ARCH-LM test for persistence in volatility

When σ_t^2 is high, is σ_{t+1}^2 likely to also be high?

- This is like checking for time-series effects in any dataset.
- Could we just apply one of our existing tests to data on σ_t^2?
- The major problem is, you can't directly measure σ_t^2.
- Instead, focus on r_t^2 as an indirect measure.

The standard test is known as *Engle's test* or the *ARCH-LM* test.

- Run the following regression: $r_t^2 = \alpha + \beta r_{t-1}^2 + \varepsilon_t$
- Under the null of no persistence in σ_t^2, we have $N \times R^2 \sim \chi^2$.
- Implemented in statsmodels with the *het_arch* command.
- You can also extend the regression to include p lags of r_t^2.
  Then the test statistic is asymptotically $\chi^2(p)$.
  By default, het_arch uses a lag of $\sqrt{T}$, rounding down.

Modeling persistent volatility: Moving average

One popular way to measure volatility over time is through a moving average. This works much like the moving average from week 1.

But, a complication is that you cannot actually measure volatility at any point in time. Instead, the moving-average estimate of volatility is based on an average of r_t^2 terms:

$$\sigma^2_{t,CMA} \quad = \quad \frac{1}{m}\left(\frac{1}{2}\,r^2_{t-m} + r^2_{t-m+1} + \cdots + r^2_{t+m-1} + \frac{1}{2}\,r^2_{t+m}\right)$$

This is a fast and easy way to describe and plot volatility over time.

Modeling persistent volatility: EWMA

- EWMA stands for Exponentially Weighted Moving Average.
- The basic idea is like exponential smoothing from week 2.
- You choose a smoothing parameter α, and an initial σ_0^2, which is often the sample value from the first few observations.
- Then update the values of σ_t following the formula:

$$\sigma_{t+1}^2 = \alpha \times r_t^2 + (1-\alpha) \times \sigma_t^2$$

- Higher α moves you close to just estimating tomorrow's volatility with today's r^2. Lower α gives smoother results that react more slowly to changes, for better and for worse.

Smoothing vs model-based approaches

- We can make some of the same comments here that we did after we covered Holt-Winters forecasting:
- Smoothing-based approaches give an easy and effective way to describe the data and build accurate forecasts.
  If that is our only goal, they are likely to be good enough.
- But these approaches will break down as we move our attention to other goals like building confidence intervals, or understanding relationships between different series.
- At some point we have to adopt the *model-based* approach that is more standard in statistics.
- Before, this led us to ARMA. Here it leads us to a set of volatility models that will be our main focus.

Modeling persistence in volatility: ARCH

The simplest and oldest volatility model is called ARCH (which stands for "**A**uto**R**egressive **C**onditional **H**eteroskedasticity")

$$y_t = \mu + u_t \quad , \quad u_t \sim \mathcal{N}(0, \sigma_t^2)$$
$$\sigma_t^2 = \omega + \alpha u_{t-1}^2$$

- The first line could be replaced with any ARMA model for y_t. However, in practice volatility modeling is most important in situations where there probably is no predictability in y_t.
- The important thing is the second line: When today's (squared) error is large, tomorrow's error has larger variance, through α.
- More generally, an ARCH(q) model includes q lags of u_t^2. The above model is an ARCH(1).
- ARCH models are typically estimated by MLE.

Modeling persistence in volatility: GARCH

In practice, ARCH models have generally been replaced by GARCH ("**G**eneralized **A**uto**R**egressive **C**onditional **H**eteroskedasticity")

$$y_t = \mu + u_t \quad , \quad u_t \sim \mathcal{N}(0, \sigma_t^2)$$
$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2$$

- The new thing compared to ARCH is the $\beta\sigma_{t-1}^2$ term.
  This extra flexibility greatly improves the model performance.
- This general category of models is called GARCH(p,q), where
  - p is the number of lags of σ_t^2,
  - q is the number of lags of u_t^2,

  The above model is a GARCH(1,1).
  In this class we will only look at ARCH(1) and GARCH(1,1).
- Again, GARCH models are typically estimated by MLE.

Finding a good model for volatility in the data

- Decide on an ARMA model for y (often just $y_t = \mu + u_t$). In the first step, don't yet model any volatility effects.
- Estimate the model, apply Engle's ARCH-LM test to the sample residuals $\hat{u}_t$, and examine ACF and PACF plots of $\hat{u}_t^2$.
- If the test rejects, or you see important patterns in the plots, add a simple ARCH or GARCH specification to your model.
- Apply Engle's test to the *standardized* residuals $\hat{u}_t / \hat{\sigma}_t$, where $\hat{\sigma}_t$ are the model's estimates of volatility over time (labeled as conditional_volatility in Python's ARCH results). And plot the ACF and PACF of the squared values $(\hat{u}_t / \hat{\sigma}_t)^2$.
- Add just enough ARCH and GARCH terms, so that the test fails to reject and you don't see important patterns in the plots.

Application: Value-at-risk (VAR)

- The **value-at-risk** (VAR) of a portfolio is a percentile of its return distribution over some specified amount of time.
- For example, the 10-day, 5% value-at-risk is the 5th percentile of the distribution of returns over the next 10 days.
  - To be more precise, it is the *absolute value* of this number, because we are usually thinking about low percentiles (losses).
- Example: If the 10-day, 5% VAR is 8%, we would say you have a 5% chance of *losing more than 8%* over the next 10 days.
  - It is also common to express VAR in dollar terms:
    If your portfolio is worth \$1m, and the 10-day 5% VAR is 8%, then you have a 5% chance of losing more than \$80k.

Calculating VAR from a normal distribution

- Suppose you have daily return data ending at time T, and you want to know the α % VAR over the next K days.
- Let's use $r_{T \to T+K}$ to label the return over the next K days. The calculation is easy, if we know the *distribution* of $r_{T \to T+K}$.
- We will use log returns, so $r_{T \to T+K} = r_{T+1} + r_{T+2} + \ldots + r_{T+K}$.
- Assume that each day's log return follows a normal distribution.
  (This is not critical, but it also is not a bad approximation.)
- Then $r_{T \to T+K}$ is also normal (a sum of normal variables). If we can find its mean $\mu_{T \to T+K}$ and its volatility $\sigma_{T \to T+K}$, then the VAR is just $\mu_{T \to T+K}$ minus a multiple of $\sigma_{T \to T+K}$.
  - A 5% VAR is approximately $\mu_{T \to T+K} - 1.645 \times \sigma_{T \to T+K}$.

Calculating $\mu_{T \to T+K}$ from daily data

- We typically assume there is no predictability in stock returns.
- Then $\mu_{T \to T+K} = K \times \mu$, where μ is the average *one-day* return.

  This follows from:

$$\mathbb{E}[r_{T \to T+K}] = \mathbb{E}[r_{T+1}] + \mathbb{E}[r_{T+2}] + \ldots + \mathbb{E}[r_{T+K}] = \underbrace{\mu + \mu + \ldots + \mu}_{K \text{ times}}$$

- If you *do* think there is predictability in returns, you can just fit a model such as ARMA and forecast in the standard way.
  But in practice, this is not likely to give you meaningful results.

Calculating $\sigma_{T \to T+K}$, ignoring predictability in σ

- As in the previous slide, let's start with the simplest case, where we just assume that σ is completely unpredictable.
  - (But to be clear, the whole point of this topic is that this is *not* the appropriate approach, and will give us misleading results.)
- In this case, $\sigma_{T \to T+K} = \sqrt{K} \times \sigma$, where σ is one-day volatility. This follows from

$$\begin{aligned} \mathrm{Var}(r_{T \to T+K}) &= \mathrm{Var}(r_{T+1} + r_{T+2} + \ldots + r_{T+K}) \\ &= \mathrm{Var}(r_{T+1}) + \ldots + \mathrm{Var}(r_{T+K}) = \underbrace{\sigma^2 + \ldots + \sigma^2}_{K \text{ times}} = K \times \sigma^2 \end{aligned}$$

Then the K-day, 5% VAR would be $\quad K \times \mu \quad - \quad 1.645 \times \sqrt{K} \times \sigma.$

Calculating $\sigma_{T\to T+K}$, modeling predictability in σ

- Revisit the previous calculation. Now suppose σ is predictable.
- The first step from before is still correct:

$$\begin{aligned}\text{Var}(r_{T\to T+K}) &= \text{Var}(r_{T+1} + r_{T+2} + \ldots + r_{T+K}) \\ &= \text{Var}(r_{T+1}) + \text{Var}(r_{T+2}) + \ldots + \text{Var}(r_{T+K})\end{aligned}$$

  This is still correct because daily returns are uncorrelated.
- Now replace each $\text{Var}(r)$ with a *forecast* of that day's variance.
  - Before, we assumed that variance/volatility were unpredictable, so our forecast for each day was just the one-day variance σ^2.
  - Now, we instead use forecasts from a GARCH model.

How to build volatility forecasts from a GARCH model

- Suppose you have chosen and estimated a GARCH model, as described earlier. Then, you can forecast future values of σ^2 following a similar approach as with ARMA models.
- In statsmodels, the forecast() function will do this for you. Let's understand how it works.
- For example, suppose your model was GARCH(1,1). This means you specified that $\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2$, and you now have estimates of all these things up to time T.
- Your first forecast σ_{T+1}^2 is just $\hat{\omega} + \hat{\alpha}\, \hat{u}_T^2 + \hat{\beta}\, \hat{\sigma}_T^2$.
  - In statsmodels, $\hat{\omega}$, $\hat{\alpha}$, and $\hat{\beta}$ are reported in "params"; $\hat{u}$ is in "resid," and $\hat{\sigma}$ is in "conditional_volatility".
- What about the later forecasts? See next slide...

Forecasting volatility with a GARCH model (2)

What is your forecast of σ^2_{T+2}?

- The model says $\sigma^2_{T+2} = \omega + \alpha u^2_{T+1} + \beta\sigma^2_{T+1}$.
- We can use the same estimates of ω, α, and β as we did before, and for σ^2_{T+1} we substitute the forecast from the previous slide.
- But what about u^2_{T+1}? Although our best forecast of u_{T+1} is zero, our best forecast of its *squared* value is its variance σ^2_{T+1}. So we substitute the forecast from the previous slide there too!
- This means our forecast of σ^2_{T+2} is $\hat{\omega} + (\hat{\alpha} + \hat{\beta}) \times \hat{\sigma}^2_{T+1}$, where $\hat{\sigma}^2_{T+1}$ is the forecast calculated on the previous slide.
- All later forecasts follow the same logic.
  For example we forecast σ^2_{T+3} as $\hat{\omega} + (\hat{\alpha} + \hat{\beta}) \times \hat{\sigma}^2_{T+2}$, and so on.

Back to VAR calculations

We can finally complete the VAR calculation from earlier:

$$\begin{aligned}\mathrm{Var}(r_{T\to T+K}) &= \mathrm{Var}(r_{T+1} + r_{T+2} + \ldots + r_{T+K}) \\ &= \mathrm{Var}(r_{T+1}) + \mathrm{Var}(r_{T+2}) + \ldots + \mathrm{Var}(r_{T+K}) \\ &= \hat{\sigma}^2_{T+1} + \hat{\sigma}^2_{T+2} + \ldots + \hat{\sigma}^2_{T+K}\end{aligned}$$

and just substitute in the forecasts from the previous two slides. The square root of this sum is $\sigma_{T\to T+K}$, the volatility of $r_{T\to T+K}$.

Finally, we can calculate the VAR:

Start with $\mu_{T\to T+K}$, which is just $K \times \mu$, and subtract an appropriate multiple of $\sigma_{T\to T+K}$. For a 5% VAR, this multiple is 1.645.

Illustration: VAR during a crisis

- The notebook on Canvas illustrates today's concepts: You imagine doing a VAR calculation in April 2020.
- This basically comes down to estimating $\sigma_{T \to T+K}$.
- If you ignore predictability in σ, you will report a VAR that is much too optimistic. That is, you will understate your risk.
  - This approach implicitly assumes that the next K days will look like the "typical" day in the data.
  - This is not appropriate, because high volatility tends to persist, and volatility in Apr 2020 was at an all-time high!
- A GARCH model takes into account that (1) recent volatility is high, and (2) this means future volatility will also be high.
- When you use GARCH to forecast $\sigma_{T \to T+K}$, you report a VAR that is more realistic, and much larger.