PYTHON FOR THE FINANCIAL ECONOMIST, RETAKE EXAM 2021
VOLATILITY AND HISTORICAL SCENARIOS

Copenhagen Business School
28. February 2022

2 weeks, home assignment

The home assignment is to be answered in groups of two students (maximum of 25 A4-pages) or individually (maximum of 15 A4-pages). The students must individualize the assignment.

The take-home assignment should take the form of an academic report written in either Word or Latex converted to pdf format. It is expected that students present relevant formulas, present results using visualizations and tables, include relevant references, etc. The overall impression of how the results are presented will count in the assessment.

The analysis should be performed using Python. Please attach code / Jupyter notebooks.

If you think that you do not have all the necessary information to answer a problem, make the necessary assumptions in order to proceed and state these assumptions in the solution.

Good luck!

The GARCH(1,1) model - A simple model for time-varying volatility

Note: It is expected that the students implement relevant functions, optimizations, etc. themselves and not use e.g. the ‘arch‘ package.

Volatility in financial markets is typically time-varying and tends to cluster. We consider the daily returns of the S&P 500 index and apply the Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) model for time-varying volatility.

First, use e.g. *pandas-datareader* to obtain index values of the S&P 500 index from the beginning of year 2000 (note that the ticker is "$^\wedge GSPC$" when using Yahoo Finance). Calculate the daily log-returns using the adjusted closing prices. Plot the returns and the squared returns. We can think of the squared returns as a proxy or signal of volatility. Comment on the figures.

The GARCH(1,1) model introduced by Bollerslev (1986) can be written as

$$r_t = \mu + \varepsilon_t = \mu + \sqrt{h_t} z_t \tag{1}$$

$$h_t = \omega + \beta h_{t-1} + \alpha \varepsilon_{t-1}^2 \tag{2}$$

where $z_t \sim N(0, 1)$ is i.i.d. Equation (1) shows that the daily log-returns r_t can be decomposed in its expected value μ and an error term $\varepsilon_t = \sqrt{h_t} z_t$ with time-varying conditional volatility. Note $\mathrm{E}[r_t] = \mu$. The return variance conditional on the information at time $t - 1$ is given by

$$\mathrm{Var}_{t-1}[r_t] = h_t \mathrm{Var}_{t-1}[z_t] = h_t$$

Equation (2) defines the particular specification of the conditional variance. We can think of the specification as a filter that puts the weight β on the previous estimate of volatility and the weight α on the proxy of volatility we observe. The constraints $\omega > 0, \beta > 0$ and $\alpha > 0$ will ensure positivity of the variance. Furthermore, we will assume that $\alpha + \beta < 1$ such that the unconditional variance exists

$$\mathrm{Var}[r_t] = \mathrm{Var}[\varepsilon_t] = \frac{\omega}{1 - \alpha - \beta}$$

Implement a function that calculates the conditional variances, $h_t, \quad t = 1, ..., T$, for given values of μ, ω, α and β and a vector of log-returns, $r_t, \quad t = 1, ..., T$. Assume that the initial variance is equal to the unconditional variance.[1] Calculate and plot the conditional variances when $\mu = 0$, $\omega = 0.02$, $\alpha = 0.13$, $\beta = 0.86$ and returns are in percent (the calculated log-returns are multiplied with a factor 100).

[1] Implementation example assuming an expected return of zero.

The parameters, used above, are assumed and not estimated. To estimate the parameters, it is possible to use maximum likelihood estimation. The log-likelihood function can be written as[2]

$$\mathcal{L}(\mu, \omega, \alpha, \beta) = \sum_{t=1}^{T} \left[-\ln h_t - \frac{\varepsilon_t^2}{h_t} \right]$$

Estimate the parameters using data since the beginning of 2000 and enforce the relevant constraints. Present the results and plot the conditional variances using the estimated parameters.

Describe how to calculate the standard errors of the MLE in this specific case (using numerical scores and Hessian will suffice). Compute the standard errors. Are the parameters significant? At what significance level?

To forecast variance over multiple days, it should be noted that

$$\begin{aligned}
h_{t+1|t} &= \omega + \beta h_t + \alpha \varepsilon_t^2 \\
h_{t+2|t} &= \omega + \beta h_{t+1|t} + \alpha h_{t+1|t} = \omega + (\beta + \alpha) h_{t+1|t} \\
h_{t+3|t} &= \omega + \beta h_{t+2|t} + \alpha h_{t+2|t} = \omega + \omega(\beta + \alpha) + (\beta + \alpha)^2 h_{t+1|t} \\
&\vdots \\
h_{t+h|t} &= \omega + \omega(\beta + \alpha) + \cdots + \omega(\beta + \alpha)^{h-2} + (\beta + \alpha)^{h-1} h_{t+1|t} \\
&= \frac{\omega}{1 - \alpha - \beta} + (\beta + \alpha)^{h-1} \left(h_{t+1|t} - \frac{\omega}{1 - \alpha - \beta} \right)
\end{aligned}$$

where $h_{t+h|t} = \mathrm{E}_t[h_{t+h}]$ is the expectation using the information at time t. Therefore, $\mathrm{E}_t[\varepsilon_t^2] = \varepsilon_t^2$ (since the error term is known at time t) while $\mathrm{E}_t[\varepsilon_{t+h}^2] = h_{t+h|t}$. Thus, the forecast of the conditional variance will converge to the unconditional variance with the forecast horizon (given our assumption that $\alpha + \beta < 1$). Given the information at the last day in your sample, forecast the conditional variance for each of the next 100 days.

The GARCH(1, 1) model can be used for Value-at-Risk calculations. Remember that

$$r_{t+1} \sim N(\mu, h_{t+1})$$

conditional on the information at time t. Thus, the Value-at-Risk (VaR) for the next day can be found using the normal distribution. Calculate the one day ahead 5% VaR (for the log-returns) for each day in the sample. Plot the log-returns and the VaR in the same figure. Add red dots to visualize when the log-return is below the VaR-estimate.

[2] Derivation of GARCH log-likelihood. Note that the presented log-likelihood omits constants irrelevant for the optimization.

To obtain the VaR at longer horizons, one needs to apply simulations. Given the information at the last day in your sample, simulate and plot the distribution of cumulative log-returns for the next 100 days. Compare the variance of the simulate log-returns with the forecasted variance.

Historical scenarios and risk management

Using historical scenarios is an integral part of risk management, but one question arises: Are some scenarios more important than others or more relevant for a particular risk management application? Meucci (2010b) discusses how we can assign different probabilities to historical scenarios in order to enhance the flexibility of scenario based approaches.

Assume that r_t denotes the daily log-return of an equity index for a particular sample $t = 1, ..., T$. Typically, risk management is concerned with estimating the density function of r_t and analyzing distributional properties, e.g. tail properties such as Value at Risk. One can opt for a **parametric approach** where a particular distribution is assumed or a **scenario based approach** where the empirical distribution function is used.

The empirical distribution function of r_t can be written as

$$F(r) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}_{r_t \leq r} = \sum_{t=1}^{T} \frac{1}{T}\mathbf{1}_{r_t \leq r}$$

Basically, assigning a constant probability $p_t = 1/T$ to each observed scenario. To extend the flexibility of the scenario-based approach, we can modify the probabilities p_t associated with the scenarios allowing for more weight on more recent observations, etc. We define the generalized empirical distribution function as

$$F(r) = \sum_{t=1}^{T} p_t \mathbf{1}_{r_t \leq r}$$

First, use e.g. *pandas-datareader* to obtain index values of the S&P 500 index from the beginning of year 2000 (note that the ticker is "$^\wedge GSPC$" when using Yahoo Finance). Calculate the daily log-returns using the adjusted closing prices. Calculate mean, standard devation, skewness, kurtosis and the fifth percentile of the daily log-return based on the full sample.

Probabilities as a function of time

A simple and often applied modification of the probabilities $p_t,\ t = 1, ..., T$ (known as **rolling window**) is to assume them to be constant over some time interval such that ($\propto$ means proportional to)

$$p_t \propto \begin{cases} 1 & \text{if } \underline{t} \leq t \leq \bar{t} \\ 0 & \text{otherwise} \end{cases}$$

The probabilities must be rescaled such that

$$\sum_{t=1}^{T} p_t = 1$$

Another alternative is **exponential smoothing**. We can simply set p_t equal to the exponential decay factor (we need to account for a scaling factor - divide by $\sum_{t=1}^{T} e^{-\lambda(T-t)}$)

$$p_t \propto e^{-\lambda(T-t)}$$

such that more recent scenarios are more likely.

Calculate the same metrics as before with constant probabilities for the case where we use a rolling window of the last 750 days and for the case where we use exponential smoothing with $\lambda = 0.0055$. Create a plot that compares the probabilities for the three cases so far mentioned.

Market conditioning

So far, we have defined "closeness" in terms of time such that we have selected scenarios based on when they have occurred. However, it may also be relevant to define scenarios based one macroeconomic state variables. For instance, we may want to select scenarios where a macroeconomic variable Y is inside some region $\mathcal{Y}$:

$$p_t \propto \begin{cases} 1 & \text{if } y_t \in \mathcal{Y} \\ 0 & \text{otherwise} \end{cases}$$

Again, we must normalize probabilities such that they sum to one. We may label this way of setting probabilities for **crisp conditioning** since we assign probability only if the state variable is inside some interval.

Obtain the VIX index from the beginning of year 2000 (note that the ticker is "$\wedge VIX$" when using Yahoo Finance) and plot it. Assign probability when the VIX is above 25. Calculate the same metrics as above. Visualize the probabilities.

The crisp selection above may be a bit arbitrary in the sense that scenarios with the macroeconomic variable just inside the boundary get assigned full probability while scenarios with the macroeconomic variable just outside the boundary get assigned zero probability. An alternative approach is to use **smoothed probabilities** by applying e.g. a normal kernel to define closeness to some target level $\overline{y}$

$$p_t \propto e^{-(y_t - \overline{y})^2/2s^2}$$

where s is the volatility or bandwidth which controls the level of smoothing. When $s \to \infty$ then $p_t \to 1/T$ which corresponds to the unconditional case. Again, we need to normalize the probabilities such that they sum to one.

Assign smoothed probabilities when assuming $\bar{y} = 35$ and $s = 5$. Calculate the same metrics as above. Visualize the probabilities.

In the above kernel smoothing approach, we used the Gaussian kernel to define closeness to some target $\overline{y}$. We could follow the same approach using different kernels and different choices of the bandwidth parameter.

An alternative approach is to specify probabilities using the **Fully Flexible Views approach** of Meucci (2010a) such that the expected value of the macroeconomic variable is equal to our target. Basically, we want to select a the vector of probabilities $\mathbf{p} = (p_1, ..., p_T)^\top$ such that

$$\mathbf{p}^\top \mathbf{y} = \sum_{t=1}^{T} y_t p_t = \overline{y}$$

while the distance measured with the relative entropy (also known as Kullback-Leiber divergence) from the reference or prior distribution defined by $q_t = 1/T$ (unconditional case) is minimized. The relative entropy is defined as

$$d(p, q) = \sum_{t=1}^{T} p_t \ln \frac{p_t}{q_t}$$

To sum up, we have the optimization problem

$$\mathbf{p}^* = \arg \min_{p} \, p_t \ln T p_t \;\; \text{st.} \; \sum_{t=1}^{T} y_t p_t = \overline{y}$$

which can be solved efficiently using the algorithm presented in Meucci (2010a).

Assign fully flexible probabilities when assuming $\bar{y} = 30$. Calculate the same metrics as above. Visualize the probabilities.

References

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Meucci, A. (2010a): "Fully Flexible Views: Theory and Practice," *SSRN working paper.*

——— (2010b): "Historical Scenarios with Fully Flexible Probabilities," *SSRN working paper.*