Exercise set #1

Leopoldo Catania

The purpose of this exercise set is to get comfortable with R, financial returns, and GARCH models.
First start by opening R, create a new script and save it to your Hard Drive with the name: "Exercise1.R".

(1)

Consider the file DJITicker.xlsx which contains the Yahoo tickers associated to the Dow Jones Industrial Average (DJIA) index. Save it in csv and load it in R.

(2)

Your task is to download the time series of adjusted closing prices for all the assets in the DJIA index for the last year using the getSymbols function of the quantmod package of R. Compute the percentage logarithmic returns (i.e. $y_t = (\log(price_t) - \log(price_{t-1})) \times 100$) for each asset and collect the resulting series in a list called lRet.

(3)

Create a 30×7 matrix called DescStat. The rows name are the tickers and the columns name are: "mean", "median", "variance", "kurtosis", "skewness", "rho", "rho2". Fill each element of the matrix with the associated statistic where "rho"and "rho2" correspond to $cor(y_t, y_{t-1})$ and $cor(y_t^2, y_{t-1}^2)$, respectively.

(4)

The Capital Asset Pricing Model (CAPM) has been introduced by Jack Treynor, William F. Sharpe, John Lintner and Jan Mossin independently and is used to asses the required rate of return for a financial investment given it's risk level. Assuming that the risk free rate is zero, it simply takes the form of a linear regression as:

$$y_t = \alpha + \beta x_t + \varepsilon_t,$$

where $E[\varepsilon_t] = 0$, $E[\varepsilon_t^2] = \sigma^2$, $E[\varepsilon_t x_t] = 0$, and x_t is the market return. The β measures the asset's sensitivity to the market, the α measures the expected return of the asset (assuming $E[x_t] = 0$) and σ^2 it's variance. Assume that the market can be represented by the S&P 500 index:

- Compute x_t as the percentage logarithmic return of S&P 500. Download the series from Yahoo finance (the ticker is ^GSPC, note the ^ symbol).
- Estimate α, β, and σ^2 using the expression you know from your econometrics/statistics 1 course for each asset.
- Put the estimates in a 30×3 matrix.

(5): Simulation

Create a function that allows you to simulate T observations from a GARCH(1,1) model:

$$y_t = \sigma_t \epsilon_t, \quad \epsilon_t \overset{iid}{\sim} N(0, 1)$$
$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2$$

Note that, if $\beta = 0$ you have an ARCH(1) model. Initialize the GARCH recursion at its unconditional value, i.e. $\sigma_1^2 = E[\sigma_t^2]$.

(6): Estimation

- Create a function to evaluate the likelihood of an ARCH(1) model.
- Create a function to estimate an ARCH(1) model by Maximum Likelihood. (set σ_1^2 to it's unconditional value, i.e. $\sigma_1^2 = E[\sigma_t^2]$)

(7) Monte Carlo

You have a function to simulate from a GARCH(1,1) model and another function to estimate an ARCH(1) model. In this exercise we want to study the finite sample properties of the Maximum Likelihood estimator in the case of correctly specified and misspecified models. In a Monte Carlo exercise we usually simulate T observations from the true model and then we estimate the model. We repeat this operation B times and collect the estimated coefficients. We then plot the empirical density (see `help(density)`) of the estimated coefficient and check if the estimator is unbiased. Furthermore, as long as the sample size (T) increases, we should observe two facts: 1) the bias of the estimated coefficient ($\hat{\theta}$) from the true one (θ_0) reduces

$E[\hat{\theta}-\theta_0] \to 0$ when $T \to \infty$, and 2) the Mean Squared Error between the estimated coefficient and the true value reduces, $E[(\hat{\theta}-\theta_0)^2] \to 0$ when $T \to \infty$

- Perform a Monte Carlo exercise in the case of a correctly specified model. Repeat $B = 500$ times for $T = 200, 500, 1000$:
  - Simulate T observations from an ARCH(1) model with $\omega = 0.3$ and $\alpha = 0.7$.
  - Estimate ω and α maximizing the log likelihood of an ARCH(1) model. (Use the true values as starting values).
  - Collect the estimated coefficients $\hat{\omega}$ and $\hat{\sigma}$ in an array indexed by $b = 1, \ldots, B$, and T
- Perform a Monte Carlo exercise in the case of a misspecified model. Repeat $B = 500$ times for $T = 200, 500, 1000$:
  - Simulate T observations from a GARCH(1) model with $\omega = 0.3$, $\alpha = 0.1$, and $\beta = 0.8$.
  - Estimate ω and α maximizing the log likelihood of an ARCH(1) model. (Use the true values as starting values).
  - Collect the estimated coefficients $\hat{\omega}$ and $\hat{\sigma}$ in an array indexed by $b = 1, \ldots, B$, and T

Show that $\hat{\omega}$ and $\hat{\alpha}$ converge to their true value for the correctly specified model. Is is the same when the model is misspecified?

(8) Real data

Use the code from Exercise set #1 and compute the percentage logarithmic returns of the equity price of the 30 constituents of the Dow Jones Industrial Average index for the last **two** years.

- Estimate an ARCH(1) model using the code you wrote in the previous point on each time series. Compute the Average Bayesian Information Criteria:

  $$BIC = \frac{-2LL}{T} + \frac{2m}{T}, \tag{1}$$

  where LL is the log likelihood value at it's optimum, T is the sample size, and m is the number of estimated parameters. Collect the BIC values for each series. Note that models with lowest BIC are preferred!
- Install the rugarch package for R.

- Study the rugarch vignette you find in Blackboard: Introduction_to_the_rugarch_package.pdf. Specifically, look at the ugarchspec() and ugarchfit() functions.
- Estimate the GARCH(1,1), E-GARCH(1,1) and GJR-GARCH(1,1) models using the rugarch package and collect the BIC for each model/series. (If Fit is the output of ugarchfit(), then infocriteria(Fit) extracts the average BIC as well as other information criteria).
- Select for each series the best model between ARCH(1), GARCH(1,1), E-GARCH(1,1) and GJR-GARCH(1,1) according to the BIC criteria.