EE 215 - Coding Assignment III

Stock market can be seen as a random process. The goal of this homework is doing data analysis on simple stock market data to develop intuitions towards random processes. We will consider five companies Amazon (AMZN), Intel (INTC), Google (GOOGL), Apple (AAPL), and Qualcom (QCOM) for the two year duration from 2015 to 2016. The homework dataset can be found under ilearn assignments tab (see "dataset for coding homework 3").

Explanation of the dataset: Dataset is a CSV file with $N = 504$ lines. Each line represents consecutive trading days from 01/01/2015 to 12/31/2016. The columns represent the daily returns corresponding to that company (e.g. Amazon returns are under column AMZN). Denoting the company price on day i by p_i, the return r_i is the relative price increase from day $i-1$ to day i and is calculated as

$$r_i = \log \frac{p_i}{p_{i-1}}.$$

For instance, a negative return means, the company stock price decreased on that day. Note that, the log-price ratio from day i to day j ($j > i$) can be easily calculated by returns as follows:

$$\log(\frac{p_j}{p_i}) = \sum_{k=i+1}^{j} r_k. \tag{1}$$

Your tasks are described next.

1 Assignment

1. **Autocorrelation:** Intuitively, we expect that, today's return should not be correlated with future returns. If there is a significant correlation, it would be too easy to make money. Let us verify this. Pick GOOGL column. Calculate the empirical autocorrelation $R(\tau)$ (which is $R_{\text{GOOGL}}(\tau)$) and autocovariance $C(\tau)$ from day-offset $\tau = 0$ to $\tau = 125$ and plot it. For instance, $R(\tau)$ can be calculated as follows.

$$R(\tau) = \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} r_i r_{i+\tau}. \tag{2}$$

   For $C(\tau)$ you need to subtract the empirical mean first. Plot $R(\tau)$ from -125 to 125 (using $R(\tau) = R(-\tau)$. Brief comment on $R(0)$ (or $C(0)$) compared to other values.

2. **Power spectrum:** Take the (discrete) Fourier transform of the autocorrelation $C(\tau)$ and plot it from -125 to 125. Brief comment on the spectrum.

3. **Stationarity:** Let us see if stocks are stationary. One way to test this is checking whether the average return changes over time (i.e. whether expectation changes). Pick GOOGL again. Let us split 2 years into 8 quarters of $T = 63$ days each. Each quarter runs from day $(k-1)T+1$ to kT for $k = 1$ to 8. Find the average return in each quarter i.e.

$$\mu_k = \frac{1}{T} \sum_{i=(k-1)T+1}^{kT} r_i.$$

   Create a bar plot where x axis runs from 1 to 8 and y axis is the average return. Is this stationary?

4. **Cross correlation:** Calculate the empirical cross-correlation between Amazon and Google i.e. find empirical $R_{XY}(\tau)$ where X = Google and Y = Amazon by modifying (2). Briefly comment on the plot.

   Now, let us look at the daily correlations. Create a 5×5 correlation coefficient matrix where (j,k)th entry is the correlation between the daily returns of (j,k)th companies. Mathematically, you need to pick columns j,k and calculate the good old correlation coefficient

$$\rho(i,k) = \frac{1}{N}\sum_{i=1}^{N}\frac{(r_{j,i}-\mu_j)(r_{k,i}-\mu_k)}{\sigma_j\sigma_k}$$

   where μ_j, σ_j are mean/standard deviation. Which two companies are most correlated?

5. **Overall performance:** As we discussed in (1) one can find the log-price from summation of the returns. Hence, the price is simply a sum-process obtained from the returns! Let us see which company has the best performance during 2015/16. Create the (exponentiated) sum-process

$$q_i = \frac{p_i}{p_0} = \exp(\sum_{j=1}^{i} r_j).$$

   q_i is the normalized price i.e. it represents the stock price of the company if its stock price was 1 at the start of our analysis which is 2015/01/01. Plot all 5 companies in the same figure. The x axis runs from 1 to $N = 504$ and y axis is the normalized price. Which company stock is the top performer during 2015/16?