Project proposal – Mathematical Modelling Techniques for Data Analytics

Authors: Yun Mai

# Portfolio optimization

**Abstract:**

Making a decision on what stocks to invest is a paramount important problem an investor facing when investing the stock market. In this project, I will use mathematical models as a base for making a decision and investigate portfolio optimization with the goal of maximizing the expected return on the investment while keeping the risk under the control.

**Data:**

The dataset for the portfolio optimization from the class website of Stanford University course: Introduction to Matrix Methods (IMM) (http://stanford.edu/class/ee103/portfolio.html). The datasets include the daily market close price, daily market volumes, and daily market returns. An alternative data source is the datasets from LSTM _stock_prediction-20170507 on Kaggle (https://www.kaggle.com/benjibb/lstm-stock-prediction-20170507/data). A sample shows as below:

| | X <fctr> | AGG <dbl> | DBC <dbl> | DFE <dbl> | DIA <dbl> | DXJ <dbl> | EEM <dbl> | EFA <dbl> |
|---|---|---|---|---|---|---|---|---|
| 1 | 2007-01-03 | 0.0000 | 0.00000 | 0.00000 | 0.00000 | 0.0000 | 0.0000 | 0.00000 |
| 2 | 2007-01-04 | 0.0021 | -0.02605 | -0.00091 | 0.00169 | 0.0121 | -0.0138 | -0.00259 |

**Problem:**

There are 54 stocks in the IMM dataset available for selection. The optimal portfolios will be generated for the period from 2006 to 2007 on a preselected 10 arbitrarily selected stocks. The optimal allocation weights for each stock under the constraint for risk and budget share need to establish.

**Mathematical methods:**

Two methods will be used in solving the problem.

1.Mean Variance (MV) optimization model

First, the mean returns will be calculated. Then from the mean return result to generate a covariance matrix. Generate an estimate of the portfolio variance from the covariance matrix. The portfolio variance was minimized to generate the expected return and calculate the final weights.

2.Game theory - the Minimax portfolio selection models

First, the mean returns will be calculated. The standard deviation and variances will be calculated for the returns data. The worst returns will be computed by subtracting standard deviations from mean returns. The maximum of the lowest worst return will be assigned as the weights of the assets.