MATH584 - Math for Algo Trading Homework 1

Niti Wattanasirichaigoon
A20406934

File Descriptions

File Name	Description
meanReturns.csv	1a) Estimated vector of mean returns (Annualized)
covMatrix.csv	1a) Estimated covariance matrix (Annualized)
mvp_weights.csv	1b) Weights of the minimal variance portfolio
omvp_weights.csv	1c) Weights of the optimal mean-variance portfolio
omvp_weights_robust.csv	1d) Weights of the optimal mean-variance portfolio in the robust setting
omf_weights.csv	1f) Weights of the market portfolio (Optimal mutual fund)
betas.csv	2a) 'beta' for each basic asset (CAPM model)
aibi.csv	2b) 'alpha' and 'beta' values calculated by using CAPM regression
aici.csv	2c) 'alpha' (intercept) and Ci (slope) of an additional factor (prior 5 days' mean), CAPM regression
aici2.csv	2d) 'alpha' (intercept) and Ci (slope) of an additional factor (Volume weighted rolling mean), CAPM regression
meanExcessReturns.csv	2e) Mean excess returns of each asset
p_values.csv	2e) Lowest p-values obtained from the optimal lambda for 64 assets (rows) and 2 factors (columns). (col0=2c, col1=2d)
meanHedged.csv	2e) Mean excess hedged returns of the sub-samples obtained from optimal lambda (64 assets, 2 factors)
sharpe.csv	2e) Sharpe ratios for each of the sub-samples (64 assets, 2 factors)

1. Constructing the Efficient Frontier

a. The daily returns are first calculated by the finding the relative change of adjusted closing prices for each day, thus the very first day does not have a value. The estimated vector of mean returns (sample mean returns) is calculated by taking the average of all returns for each asset and the estimated covariance matrix (sample covariance matrix) is calculated upon all available returns as well.

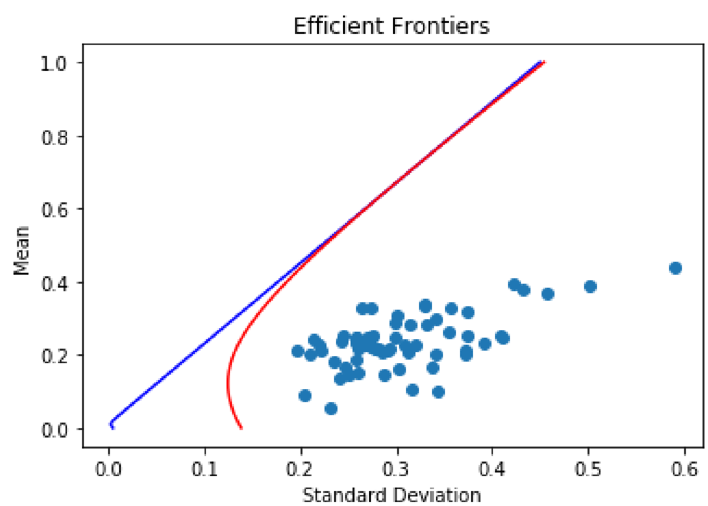
b. The weights of the minimal-variance portfolio is calculated using the *scipy.optimize.minimize* function. The objective function is the formula for calculating the portfolio variance and the constraint is that the sum of all weights add up to 1. The resulting weights are the weights that should yield us the portfolio with the least variance.

c. The objective function for the optimal mean-variance portfolio returns a negative value because *scipy* only has to ability to minimize (minimize a negative = maximize). The function itself consists of the portfolio mean minus the variance with a risk tolerance factor $\frac{1}{\gamma} = 1$. By plotting the weights, we observe that their values spread over the range -1.5 to +1.5.

d. The process in this step is similar to part 1c, but we also add a robustness factor to the objective function. The robustness term penalizes weights that have a high standard deviation (uncertainty of the mean). When plotted, most of the weights are reduced down to nearly zero, while the less risky assets have been scaled up. The upper and lower bounds of the mean are equidistant from the mean in part 1c, but now the variance of the weights are much smaller.

e. The efficient frontier is then created by adding another constraint to the minimal variance portfolio-a target mean for the portfolio. We are then able to compute the weights that correspond with each value of the mean from 0 to 1, and use those weights to calculate the corresponding portfolio standard deviation. Plotting the mean-SD values gives use the efficient frontier. We observe that our assets points lie to the right of the frontier because the standard deviations for each individual asset will always be higher than the portfolio's standard deviation.

f. We add a riskless asset to the market and compute the weights for the optimal mutual fund by solving the system of linear equations using *numpy.linalg.solve*. We then plot the 'extended' market's efficient frontier (here I treated the riskless asset as an extra asset to compute the weights, but the weight of the riskless asset is no used to calculate the portfolio's standard deviation). The extended market's efficient frontier (in blue) tangents the efficient frontier in part 1e at a point in the upper middle half of the graph. This point of tangent is exactly our optimal mutual funds.



2. Regression interpretation of CAPM

a. The 'beta' values for each basic asset is calculated as the ratio of the covariance between the i^{th} asset and the market portfolio to the variance of the market portfolio.

b. We regress the excess returns of each individual basic asset onto the excess returns of the market. The resulting regressions return an intercept and a slope for each asset. The intercepts (a_i) are the mean returns of the hedged assets and the slopes (b_i) are our 'beta' values. The intercepts have values very close to zero and our 'beta' values are very close to that obtained in part 2a.

c. Using the 'beta' we obtained, we calculate the daily return of hedged assets and regress them on an additional factor. The factor here is the prior 5-day average return (mean return of the 5 days before that day). The results give us a new set of intercepts (a_i) and slopes (coefficient for the additional factor (c_i)). We observe that 27 out of the 64 assets have p-values less than 0.05.

d. The process is similar to that of part 2c, but the factor is changed to the volume-weighted prior 5-day return instead. Hence, the factor becomes puts an emphasis on returns that also had higher volumes on their days. Here, 22 of the 64 assets have p-values less than 0.05.

e. We modify parts 2c and 2d by adding a coefficient lambda (**λ**). For each asset and factor (f1 = factor from 2c and f2 = factor from 2d), we select a sub-sample of days where the factor's value is above the threshold $\mu + \boldsymbol{\lambda\sigma}$ (for factors), and perform least squares regression of the hedged returns on the factors. We then define the optimal lambda as the one that gives us the least p-value, and save those p-values for each asset and factor. We also take the days that correspond to our 'optimal' sub-samples and calculate the mean excess returns of each hedged asset.

   Comparing these to the non-hedged mean excess returns, they are much higher. This is because we filtered out days where the factors' significance was explained the most, so the hedged mean excess return of that day should also be high. Additionally, the f2 factor has a smaller Sharpe ratios than the f1 factor which makes sense because the factor captures more information (a_i closer to zero). Therefore the hedged mean excess returns of f2 is smaller than f1, making its Sharpe ratios smaller.