Math 584 (Math for Algo Trading), Spring 2020.

Homework 1

Due: Tue, Feb 11, 2020, NO LATER than 4:30pm.

For each ticker in the file 'TechTickers.scv', download its adjusted closing price for each business day between Jan 1, 2009, and Dec 31, 2019. This is your sample. All questions listed below must be answered using this sample. All mean returns and their variances must be annualized assuming 250 days in a year. Assume the existence of a riskless return 0.01. All exercises must be solved using only the modules listed in the file 'Template.ipynb'.

1. In this question, you construct the efficient frontier. For parts (a)–(e), you need to use only stock returns, without a riskless asset.

   (a) 3 pts Produce the estimated vector of mean returns (sample mean returns) and the estimated covariance matrix (sample covariance matrix) and save them in '.csv' files.

   (b) 5 pts Compute the weights of the minimal-variance portfolio and save them in '.csv' file.

   (c) 5 pts Compute the weights of the optimal mean-variance portfolio (i.e., maximizing a linear combination of mean and variance) with the risk tolerance $1/\gamma = 1$. Save the weights in a '.csv' file and plot them on a graph. Output the mean and variance of the resulting portfolio.

   (d) 6 pts Compute the weights of the optimal mean-variance portfolio in the robust setting, assuming that the mean returns of basic assets are within one standard deviation (the standard deviation is estimated from the sample) away from the mean (the mean is also estimated from the sample). Save the weights in a '.csv' file. Output the variance of the resulting portfolio, as well as its worst- and best-case mean (according to the chosen intervals of possible mean returns of the basic assets). Compare these means and the variance to those produced in part (c) and explain the difference between the two results. Plot the weights of the resulting optimal portfolio and compare this graph to the one produced in part (c).

   (e) 6 pts Compute the efficient frontier and plot it as a set $\{(f(\mu), \mu)\}$, where μ changes over a grid of 100 equidistant points in $[0, 1]$, and $f(\mu)$ is the standard deviation corresponding to the mean return μ. On the same plot, show the pairs $(\sqrt{\sigma_{ii}}, \mu_i)$ corresponding to the standard deviations and the means of the returns of individual basic assets. Comment on where the latter pairs lie relative to the efficient frontier and why.

   (f) 6 pts Add a riskless asset to the set of available ones. Compute the weights of the market portfolio (i.e., of the optimal mutual fund) and save them in a '.csv' file. Compute the efficient frontier for the extended market and plot it in the same coordinates, and for the same values of μ, as in part (e). Plot the efficient frontier from part (e) on the same graph and comment on the relationship between the two.

2. In this question, you investigate the regression interpretation of CAPM.

   (a) 3 pts Compute the "beta" for each basic asset, according to the CAPM model, and save the results in a '.csv' file.

(b) 6 pts Use the (least-square) linear regression model, to regress the excess returns of the individual basic assets on the excess return of the market portfolio and a constant. Recall the mean returns of the hedged assets, $\{a_i\}_{i=1}^{64}$ ("hedged" means that we subtract $\beta_i(R_M - R)$ from the return). Save the resulting $\{(a_i, \beta_i)\}$ in a '.csv' file. Comment on the magnitude of $\{a_i\}$ and compare the resulting $\{\beta_i\}$ to those obtained in part (a).

(c) 6 pts Use the (least-square) linear regression model, to regress the excess returns of the individual basic assets on the excess return of the market portfolio, a constant, and an additional factor. As the additional factor use the prior 5-day average return. Save the resulting regression coefficients, for all basic assets, in a '.csv' file. Comment on the significance of the new factor using the p-values.

(d) 6 pts Answer the same questions as in part (c) using the volume-weighted prior 5-day average return as the factor (do not use the additional factor used in part (c)).

(e) 8 pts Perform the regressions described in parts (c), (d), with the following modification. Choose a coefficient $\lambda \in (0, 1)$ and, for each factor, perform the regression over a subsample consisting only of those days on which the additional factor is above $\mu + \lambda\sigma$, where μ and σ are the mean and standard deviation of the factor (computed over the entire sample). Vary λ over an equidistant grid of 10 points in $(0, 1)$ and find optimal λ_{ij}, for each basic asset i and each factor j. Save in a '.csv' file the p-values of the associated regression coefficients, for each factor and each asset, using the associated optimal λ_{ij}. Finally, for each factor, compute and save in a '.csv' file the mean excess returns of each hedged asset ("hedged" means that we subtract $\beta_i(R_M - R)$ from the return) computed over the subsample determined by λ_{ij}. Compare these values to the mean excess returns of each basic asset computed over the entire sample (and not hedged). Compare the corresponding Sharpe ratios. Comment, from a practical point of view, on the sizes of the subsamples corresponding to each λ_{ij}.