MIE377 – Financial Optimization Models

Lab 3 (03–Feb–2023)

Factor Models

In this laboratory we will implement a single-factor model and use it to estimate the input parameters required for portfolio optimization. Specifically, we will implement the Capital Asset Pricing Model (CAPM).

Single-factor model

The CAPM is used to estimate the *excess* return of an asset, i.e., its rate of return in excess of the risk-free rate. Recall the structure of CAPM

$$r_{it} - r_f = \alpha_i + \beta_i(f_{mt} - r_f) + \varepsilon_{it}$$

where

- $r_{it} - r_f$: Excess return of asset i at time t (random variable).
- r_f: Risk-free rate.
- α_i: Active return (i.e., the intercept of regression).
- β_i: Factor loading (for a single-factor model, it can be thought of as the slope of the regression line).
- $f_{mt} - r_f$: excess market return at time t (random variable).
- $\varepsilon_{it} \sim \mathcal{N}(0, \sigma^2_{\varepsilon_i})$ is the stochastic error term (pertains to the idiosyncratic risk of the asset).

Data pre-processing

We will use historical asset prices for $n = 50$ assets, and we will use the S&P 500 index as the market factor. The first step is to calculate the returns of both assets and the index as we did in Lab 1,

$$r_{it} = \frac{P_{it} - P_{i,t-1}}{P_{i,t-1}}.$$

Since CAPM requires the excess returns, we must adjust our returns by subtracting the risk-free rate. Assume a constant risk-free rate of 0.01% per week (the raw data provided is weekly).

Estimate the expected excess market return, $\bar{f}_m - r_f$, by using the geometric average of the historical returns. Finally, we can also estimate the factor variance, σ^2_m, from its historical returns.

Beta estimate

Financial factor models are typically estimated using ordinary least squares (OLS) regression. To get a better understanding of this process, we will estimate the factor loading β_i for each asset i using the following three methods:

a) Use the CAPM definition of 'beta', $\beta_i = \frac{\sigma_{im}}{\sigma^2_m}$. Use the Matlab function $cov(x, y)$ and take the off-diagonal element to calculate σ_{ij}.

b) Use the Matlab function $regress(y, B)$. Refer to the Matlab website for further details

   https://www.mathworks.com/help/stats/regress.html

c) Use the closed-form solution,

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{\alpha}^T \\ \boldsymbol{V} \end{bmatrix} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{R}$$

where $\boldsymbol{R} \in \mathbb{R}^{T\times n}$ is the matrix of all asset returns and $\boldsymbol{X} = \begin{bmatrix} \mathbf{1} & \boldsymbol{f} \end{bmatrix} \in \mathbb{R}^{T\times(p+1)}$ is the matrix of predictors. The matrix $\boldsymbol{X}$ is composed of the p factors plus a single column of ones to account for the intercept, yielding $\boldsymbol{\alpha} \in \mathbb{R}^n$ and $\boldsymbol{V} \in \mathbb{R}^{p\times n}$. $\boldsymbol{V}$ is a matrix where each element, β_{ik}, is the factor loading of asset i corresponding to factor k. For a multi-factor model, $k = 1, ..., p$. For a single-factor model, $p = 1$.

After calculating the CAPM 'beta' using these three methods, take a look at the three values. What can you conclude?

Portfolio optimization

We will now estimate our input parameters for optimization: the expected returns $\boldsymbol{\mu} \in \mathbb{R}^n$ and covariance matrix $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$,

$$\begin{aligned} \mu_i &= \alpha_i + \beta_i(\bar{f}_m - r_f) + r_f, \\ \sigma_i^2 &= \beta_i^2\sigma_m^2 + \sigma_{\varepsilon_i}^2, \\ \sigma_{ij} &= \beta_i\beta_j\sigma_m^2. \end{aligned}$$

We should notice we are missing one parameter: the idiosyncratic risk, $\sigma_{\varepsilon_i}^2$. We can estimate this value as the unbiased estimate of the residual variance. The residual of asset i at time t is

$$\varepsilon_{it} = (r_{it} - r_f) - \Big(\alpha_i + \beta_i(\bar{f}_{mt} - r_f)\Big).$$

To calculate the unbiased estimate of the residual variance, we must calculate the sum of squared residuals and divide it by the corresponding number of degrees of freedom (DOF). The DOF corresponds to the number of observations minus the number of estimated parameters. If we have p factors, then we have $T - p - 1$ DOF. Thus, the residual variance of asset i is

$$\sigma_{\varepsilon_i}^2 = \frac{1}{T-p-1}\sum_{t=1}^{T}(\varepsilon_{it})^2.$$

Now we can estimate the required input parameters for MVO. We can reduce the number of lines in our Matlab code if we use linear algebra to estimate μ and Q,

$$\begin{aligned} \boldsymbol{\mu} &= \boldsymbol{\alpha} + \boldsymbol{V}^T(\bar{f}_m - r_f), \\ \boldsymbol{Q} &= \boldsymbol{V}^T\sigma_m^2\boldsymbol{V} + \boldsymbol{D}, \end{aligned}$$

where $\boldsymbol{D} \in \mathbb{R}^{n\times n}$ is a diagonal matrix where the diagonal elements are the asset residual variances, $\sigma_{\varepsilon_i}^2$.

In a similar fashion to Lab 1, use $\boldsymbol{\mu}$ and $\boldsymbol{Q}$ to construct an optimal portfolio with a target excess return of $R_{\text{target}} = 2.5\%$. Short-selling is allowed.