Project "Advanced Topics"

Nicolas Gaussel

March, 2024

Instructions

This project will serve to evaluate students for the "Advanced Topics in Financial Modelling" class. Only students who have signed at least one presence sheet before March 22^{nd} can submit.

The project can be done individually or by pair, but not more. Your project shall be made available to me by April 14^{th}, midnight. I highly recommand the output to be a published as a Jupyter Notebook[1] or a R Markdown[2]. As a fallback solution, you can send it to me as a pdf: it will be accepted but less well rated.

As an indication, you might need around 5 hours on the "financial aspects" (reasoning, writing, etc.) and 5 hours on IT related issues (retrieving data, computing statistics, creating graphs, creating notebook, ...)

---

1 Problem

1.1 Data

You will use the following data:

- US_stocks_histo_clean.csv, a dataset containing 5 years of daily returns of stocks, representing 96% of the SP500 index.
- US_stocks_info_clean.csv, another dataset containing relevant information about those stocks.

1.2 Initial formulation

Let n be the number of available stocks and $x_{i\leq n}$ the fraction of wealth invested in each stock. Define $x = \begin{pmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{pmatrix}$, $|x| = \begin{pmatrix} |x_1| \\ |x_2| \\ .. \\ |x_n| \end{pmatrix}$ and $I_n = \begin{pmatrix} 1 \\ 1 \\ .. \\ 1 \end{pmatrix}$. The

[1] e.g. https://jupyterbook.org/en/stable/start/publish.html
[2] e.g. https://rpubs.com/YaRrr/markdownbasics

net and gross exposure (netExpo and grossExpo) are defined as

$$netExpo := \sum x_i = I_n^t x$$
$$grossExpo := \sum |x_i| = I_n^t\, |x|$$

Let Σ the empirical covariance matrix. You want to compute the minimum variance portfolio defined as

$$\begin{aligned} &\min_x \; x^t \Sigma x \\ &s.t. \left\{ \begin{array}{c} netExpo = 1 \\ grossExpo \leq e_{max} \end{array} \right. \end{aligned} \tag{1}$$

This program is not directly a quadratic program because of the constraint on the sum of the absolute values.

1.3 Reformulation as a quadratic program

Quadratic programs with n-dimensional control x, which include absolute values of x in the objective function or in the constraints, can be reformulated as plain 2n dimensional quadratic programs. The trick consists in using a $2n$ control $X := \begin{pmatrix} x \\ |x| \end{pmatrix}$ and adding $2n$ constraints

$$0 \leq |x| - x$$
$$0 \leq |x| + x$$

Question 1. - Along those lines show that the previous program can be written as

$$\begin{aligned} &\min_x \; X^t P X \\ &s.t. \; l \leq AX \leq u \end{aligned} \tag{2}$$

with $P = \begin{pmatrix} \Sigma & 0_{nn} \\ 0_{nn} & 0_{nn} \end{pmatrix}$, $A = \begin{pmatrix} I_n^t & 0_n^t \\ 0_n^t & I_n^t \\ -I_{nn} & I_{nn} \\ I_{nn} & I_{nn} \end{pmatrix}$, $l = \begin{pmatrix} 1 \\ -\infty \\ 0_n \\ 0_n \end{pmatrix}$, $u = \begin{pmatrix} 1 \\ e_{max} \\ +\infty_n \\ +\infty_n \end{pmatrix}$

and

$$I_{nn} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & ... & 0 \\ 0 & ... & ... & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \; 0_{nn} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & ... & ... & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Matrixes dimensions are : $P \to 2n \times 2n$, $A \to (2n+2) \times 2n$, $l \to (2n+2) \times 1$, $u \to (2n+2) \times 1$.

2 Computing minimum variance portfolios for several emax

2.1 Use of OSQP solver

We want to make use of the OSQP library to solve the above program. To check for robustness, you will compute the minimum variance portfolio based on data before $t_{in-sample} = 25/03/2022$ and analyse its behavior based on data after that date.

Question 2. Read the historical data csv file, compute log returns and compute the historical covariance matrix of daily returns based on data before $t_{in-sample}$.

-

Question 3. Create matrixes P, A, and l. Create u for emax $= 10$. Solve for the minimum variance portfolio associated with emax $= 10$.

-

Question 4. Solve for the minimum variance for emax in $[1; 1.5; 2; ...; 10]$ and store the results in a data frame with the following columns: ticker, emax, weight.

-

2.2 In sample analysis

Question 5. Compute the in_sample volatility of each solution. Plot it. What do you observe?

-

Question 6. Compute the number of stocks for each solution. Plot it. What do you observe?

2.3 Out of sample analysis

Question 7. Compute the covariance matrix of ouf of sample data. Compute the out of sample annualized volatility of each solution. Create a graph whith emax values on the x-axis and the in_sample and out of sample volatility as the y-axis. Comment.

3 Analysis of the minimum variance portfolio with emax$=1$

Question 8. You eventually pick the solution for emax=1. Thanks to the "info" dataframe, create a dataframe which contains, for every stock, the minimum variance portfolio weight, its PE_ratio, its sector, and its market capitalisation.

-

Question 9. What is the gain of volatility reduction as compared to the least volatile stock ?

-

Question 10. Do you observe a sector bias in the minimum variance portfolio ?

-

Question 11. Same question for PE_ratio and market caps