Final project report
Portfolio optimization

Juha Reinikainen

May 18, 2022

1 Problem

Portfolio optimization involves deciding how to use the available investment budget to maximize the total value of the investment and minimize its risk [6]. The investment budget is allocated to assets which can be for example stocks, gold, foreign exchange, real estate, bonds and cryptocurrencies [3]. For simplicity only stocks are considered in this project.

The problem is difficult because there is a large number of possible assets to include in the portfolio and even larger number of ways to divide the budget among them. Investing also has a lot of uncertainty as stock prices are affected by real world events which are hard to capture in the model [1].

2 Data

Data-driven approach to this problem involves predicting expected return and risk based on historical time-series data of stock prices over given time range, possibly years. Stock price can be sampled e.g. daily, weekly, monthly or quaterly. The stocks that are included in the data-set need to be selected.

The stock price data was collected from Yahoo Finance by calling its api with different stocks and merging these into one dataset. For example the weekly prices between 2021-05-09 20:55:46 and 2022-05-09 20:55:46 can be queried using following url.

`https://query1.finance.yahoo.com/v7/finance/download/ADS.DE?period1=1620582946&period2=1652118946&interval=1wk&events=history&includeAdjustedClose=true`

Where perdiod1 and period2 are the start and end times of the range as unix timestamps.

The stocks included in the problem where the stocks in eurostoxx 50 index.

3 Modelling

The problem can be modelled as a two objective optimization problem maximing expected return of the investment and minimizing its risk.

3.1 Variables

The decision vector consist of proportions of total budget allocated to n stocks $w = (w_1, w_2, ..., w_n)$ and binary variable for each stock which indicates whether the corresponding stock is included in the portfolio $y = (y_1, y_2, ..., y_n)$.

3.2 Constraints

It's assumed that the whole budget is used so the sum of weights should add up to one for the weight where the corresponding boolean flag is 1.

Boundary constraint requires that the weights of each stock is between w_{min} and w_{max}. Maximum limit makes it so that all budget is not allocated to too small number of stocks leading to diverse portfolio. Too small weights typically have little impact on the performance and weak liquidity and can be costly in respect to brokerage fess or monitoring costs [2].

Cardinality constraint requires that the number of stocks included in the portfolio is between some two numbers C_{min} and C_{max}.

3.3 Objectives

Popular way to model these objectives is Markowitz model which is also known as mean-variance model [5].

Given time series data of n stocks for time period of 0...T with stock price p(t, i) of stock i at time t. prices of stock i are a series x_i.

$x_i = (p(0,i), p(1,i), ..., p(T,i))$

Return of an investment between time t-1 to t for stock i is calculated by.

$roi(t,i) = \frac{p(t,i)-p(t-1,i)}{p(t-1,i)}$

Expected return for stock i is calculated as mean of each individual roi.

$roi(i) = \frac{roi(1,i)+roi(2,i)+...+roi(T,i)}{T-1}$

Expected return for n stocks is calculated as weighted sum of all individual returns.

$er(w,y) = \Sigma_{i=1}^{n}(w_i * y_i * roi(i))$

Picking stocks where the price changes a lot is risky. This problem is increased by picking stocks where this happens similarly. This can be modelled using covariance of roi values [5]. n x n square covariance matrix Cov contrains covariances of each roi(i) values. Given weights and inclusion flags, the risk for portfolio is given by.

$risk(w,y) = (w^T * y)Cov(w * y)$

3.4 Problem

Putting all this together the whole problem is.

$$\begin{aligned} minimize\{risk(w,y), -er(w,y)\} \\ \Sigma_{i=1}^{n}(w_i * y_i) = 1 \\ w_{min} \leq w_i \leq w_{max}, i = 1...n \\ C_{min} \leq \Sigma_{i=1}^{n} y_i \leq C_{max} \\ y_i \in \{0,1\}, i = 1...n \end{aligned} \tag{1}$$

4 Algorithm and settings

NSGA-II algorithm is used to solve the multiobjective constrained mixed-integer problem.

For the first n real variables simulated binary crossover and polynomial mutation are used. For the last n binary variables two-point binary crossover and bitflip mutation are used. Crossover probability is set to 1, mutation probability is set to $\frac{1}{50}$. Distribution index for real mutation and crossover are set to 3.

The constraint that the weight need to add up to one can be enforced using repair method [4]. The weights are first clamped to w_{min} w_{max} range. Then weight vector is element-wise multiplied by vector y to remove unselected stocks. Then each weight is divided by sum of all weights. It's easy to see that now weights add up to one. Special case where all weights are zero can be handled by assigning all weights with value $\frac{1}{n}$.

$$\begin{aligned}
&\frac{w_1}{w_1 + w_2 + ... + w_n} + \frac{w_2}{w_1 + w_2 + ... + w_n} + \frac{w_n}{w_1 + w_2 + ... + w_n} \\
&= \frac{w_1 + w_2 + ... + w_n}{w_1 + w_2 + ... + w_n} \\
&= 1
\end{aligned} \tag{2}$$

5 Results

References

[1] Ningning Du, Yankui Liu, and Ying Liu. "A new data-driven distributionally robust portfolio optimization method based on wasserstein ambiguity set". In: *IEEE Access* 9 (2020), pp. 3174–3194.

[2] Okkes Ertenlice and Can B Kalayci. "A survey of swarm intelligence for portfolio optimization: Algorithms and applications". In: *Swarm and evolutionary computation* 39 (2018), pp. 36–52.

[3] Muhammad Azfar Faizan. “Multiobjective portfolio optimization including sentiment analysis”. In: (2019).

[4] Massimiliano Kaucic, Mojtaba Moradi, and Mohmmad Mirzazadeh. “Portfolio optimization by improved NSGA-II and SPEA 2 based on different risk measures”. In: *Financial Innovation* 5.1 (2019), pp. 1–28.

[5] Petter N Kolm, Reha Tütüncü, and Frank J Fabozzi. “60 Years of portfolio optimization: Practical challenges and current trends”. In: *European Journal of Operational Research* 234.2 (2014), pp. 356–371.

[6] TL van Zyl, Matthew Woolway, and Andrew Paskaramoorthy. “ParDen: Surrogate Assisted Hyper-Parameter Optimisation for Portfolio Selection”. In: *2021 8th International Conference on Soft Computing & Machine Intelligence (ISCMI)*. IEEE. 2021, pp. 101–107.