Optimizing Stock Portfolio Through Linear Programming

by Jordan Pauzie

May 1, 2024

Abstract

This paper presents an integer linear programming model that optimizes portfolio investment in the stock market for maximizing expected return. The model considers risk tolerance, expected dividends and sets limits on the quantity of capital investment and shares per company. In this project, the model was applied to optimize a potential portfolio investment in shares of technology companies in the United States of America. It was successful in identifying an optimized portfolio selection of companies that was diversified and yielded significant returns all whilst staying below the given risk index threshold. Thus the linear programming model's success shows that this approach is viable for future portfolio selections, and given added dimensions, could tackle higher-complexity problems in the future.

Problem

Investing in the stock market can appear daunting due to the high volatility and uncertainty present within the market, and its sensitivity to external factors. Furthermore, given the gravity of large capital investments, there is always a not-insignificant deal of risk involved. In order to alleviate this risk, portfolio diversification is tantamount because it helps avoid dependence on a single company.

In general, portfolio selection is an important investment problem for both individuals and organizations alike, and can essentially be boiled down to finding the optimal allocation of resources such that expected return is maximized and market volatility is minimized. Framed as such, portfolio selection can consequently be viewed as a linear programming problem. Thus the project will address the issue of portfolio selection by introducing a linear programming model that will maximize expected return whilst minimizing risk.

Literature Review

The problem of optimizing portfolio selection has been well-studied for decades in various fields such as economics and finance. In 1952, Markowitz introduced the concept of portfolio diversification by proposing the Modern Portfolio Theory (MPT), a mean-variance portfolio model. Furthermore he mathematically proved that a diversified portfolio was less volatile than the total sum of its individual parts, even if each asset was quite volatile. He defined risk as the standard deviation of returns, and expressed the variance of the whole portfolio's return as a function of the individual variances of the returns of selected stocks. With the Markowitz's model, investors could maximize return for a given risk or minimize risk for a given return. Over the years, there have been numerous extensions, variations, and applications of Markowitz's model, as shown in papers such as Markowitz (1999), Rubinstein (2002), and Zhang et al. (2018). For example, Zhang et al. (2018) explored several types of improvements of Markowitz's mean-variance model including dynamic, robust and fuzzy portfolio optimization.

In the literature, portfolio diversification to reduce investment risk is usually only considered across firms i.e buying shares from various companies or across geographical regions i.e. buying assets from various countries (Luigi and Jappelli, 2008). Several papers have been released regarding international diversification such as Hui (2005), which used factor analysis to analyze which international markets to invest in, and Topaloglou et al. (2008), which developed a stochastic dynamic programming model to optimize an international portfolio selection problem.

Linear programming (LP) has been frequently employed to solve such portfolio selection problems. Mansini et al. (2014) surveyed and classified LP-based portfolio optimization techniques since 1994 such as mixed integer linear programming (MILP) techniques, used after linearizing Markowitz's quadratic programming model. Sawik (2013) reviewed linear and mixed integer programming techniques for portfolio selection optimization with multiple objectives, classified into weighting methods, lexicographic methods, and reference point methods. Papahristodoulou and Dotzauer (2004) used linear programming to compare three variations of the portfolio selection problem; these variations being the classical quadratic programming model, maximin model, and minimum mean absolute deviation model. Xidonas et al. (2018) integrated three objectives and several real-world constraints in their portfolio linear programming model, which they consequently applied to the European stock market.

Model

The linear programming model that will be utilized to solve the optimized portfolio selection is shown below:

Indices

i Company index ($i = 1, 2, ..., I$), where I is the total number of companies

Input Parameters

R_i Return of buying one share in company i

C_i Cost of buying one share in company i

D_i Dividend of one share in company i

S_i Risk of buying one share in company i

B Total starting capital

d Investor minimum dividend income, as a percentage of the investment

s Investor maximum expected risk per share

p Maximum cost proportion of each company out of total cost of the portfolio

Decision Variables

TC Total cost of the investment

TR Total return of the investment

X_i Number of i shares to buy

Objective Function

$$Maximize\ TR = \sum_{i=1}^{i} R_i X_i \quad (1)$$

Constraints

The objective function (1) will be subject to the following constraints $(2) - (7)$. Constraint (2) defines the total cost TC as the sum of purchase costs of all company shares. Constraint (3) sets the upper bound for purchase costs to be the available starting capital B. Constraint (4) guarantees that the total income from dividends is at least equal to a given proportion d of the purchase cost. Constraint (5) helps mitigate for risk by assuring that the average risk index per share does not exceed a given threshold s. Constraints (6) ensure that the investment in each individual company does not exceed a given proportion p of the total investment. Constraints (7) sets non-negative bounds for the decision variables.

$$TC = \sum_{i=1}^{i} C_i X_i \tag{2}$$
$$TC \leq B \tag{3}$$
$$\sum_{i=1}^{i} D_i X_i \geq dTC \tag{4}$$
$$\sum_{i=1}^{i} S_i X_i \leq s \sum_{i=1}^{i} X_i \tag{5}$$
$$C_i X_i \leq pTC \quad \text{for } (i = 1, 2, \ldots, I) \tag{6}$$
$$TC \geq 0, X_i \geq 0 \quad \text{for } (i = 1, 2, \ldots, I) \tag{7}$$

Assumptions

Several assumptions and intentional decisions were made in order to simplify the problem of stock portfolio selection. The largest one was regarding the scope of the linear programming model's application. In order to simplify the scope, the model will be applied for portfolio selection of investment in the stock market of the United States of America. In particular it will be limited to the ten largest companies in the technology sector by market share. Furthermore, it will be assumed that the trading commission rate is 0 partly in order to simplify calculations and partly due to the existence of various trading commission rates depending on what brokerage is used. Finally the data collected will be from a 10-year period (2014-2023) because the 2024 fiscal year is still in progress and annual stock returns haven't been determined yet. The time frame of 10 years is also due to the fact that many of the companies being analyzed first announced dividends during the early 2010's.

Initial Data

From the data that is publicly available, the expected annual returns R_i and expected annual dividends D_i were calculated as the mean returns and dividends during these ten years. For the unit stock prices C_i, the latest share prices as of May 2024 were used. To determine the expected risk for each company, S_i, the absolute difference between the expected and actual return for each year was averaged for the last ten years, with the average market risk for the U.S. stock market added. All of these input data values are shown in Table 1.

For the other parameters s, d, p and B, their values are flexible due to them being specific to the needs and desires of each individual investor. The values chosen for each parameter are given in Table 2.

Table 1. Technology sector companies and their input data values

i	Company Name	Abbreviation	C_i	D_i	R_i	S_i
1	Microsoft Corporation	MSFT	395.19	1.82	0.29	0.25
2	Apple Inc.	AAPL	169.41	0.69	0.30	0.36
3	Nvidia Corporation	NVDA	830.41	0.14	0.88	0.78
4	Broadcom Ltd.	AVGO	1,242.86	9.41	0.40	0.32
5	Oracle Corporation	ORCL	114.63	0.90	0.12	0.20
6	Salesforce Inc.	CRM	268.69	0.00	0.23	0.33
7	Advanced Micro Devices Inc.	AMD	144.19	0.00	0.72	0.84
8	Adobe Inc.	ADBE	469.39	0.00	0.31	0.41
9	Cisco Systems Inc.	CSCO	46.84	1.23	0.10	0.23
10	Accenture Plc.	ACN	298.66	3.17	0.19	0.26

Table 2. Given input values for an individual investor

s	d	p	B
40%	1%	30%	10 million

Results

Solving the linear programming model of the portfolio selection problem (code provided in Appendix A) with the values shown in Table 2 leads to the optimal solution shown in Table 3. For the optimal number of shares shown in Table 3, the corresponding objective function i.e. the expected return TR, is equal to 27,570.53. This diversified stock profile selects five technology companies, AAPL, ORCL, AMD, CSCO, and ACN to invest in. The income from dividends was exactly 1.00% of the total cost TC, while the average risk index per share was 36.65% which was less than the threshold of $s = 40\%$.

Table 3. Optimum number of shares to buy for each company

i	1	2	3	4	5	6	7	8	9	10
X_i	0	17,710	0	0	5,046	0	20,810	0	64,050	1,413

Conclusions

From the results obtained from the linear programming model, it can be seen that investing in the technology sector appears to be profitable albeit risky compared to other stabler sectors. In terms of the model itself, weaknesses do exist. Because the portfolio selection problem solely concerned the ten largest companies in the technology sector, despite the diversification of investment in the optimal solution, there is a possibility that a downturn in demand could negatively impact all of the companies at once. However there were definitely strengths with the model. Several practical constraints were incorporated in the linear programming model, including starting capital, dividends incomes, risk index, and limits on individual company's number of shares and proportion of the total investment.

In terms of applications, the portfolio selection problem that the linear programming model tackled highlights its ability to handle other potential scenarios. Furthermore, it can be seen as a base model for which additional modifications can be made in order to tackle more complex portfolio selections. From the results obtained, it can be seen that investing in the technology sector carries a relatively high risk but potentially decent returns. This high risk is validated empirically by the high levels of volatility present in the real-world technology sector.

References

Hui, T. K. (2005). Portfolio diversification: a factor analysis approach. Applied Financial Economics, 15(12), 821-834.

Mansini, R., Ogryczak, W., & Speranza, M. G. (2014). Twenty years of linear programming based portfolio optimization. European Journal of Operational Research, 234(2), 518-535.

Luigi, G., & Jappelli, T. (2008). Financial Literacy and Portfolio Diversification. ECO 2008/31, European University Institute (EUI), Fiesole, Italy.

Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7(1), 77-91.

Markowitz, H. M. (1999). The early history of portfolio theory: 1600–1960. Financial Analysts Journal, 55(4), 5-16.

Papahristodoulou, C., & Dotzauer, E. (2004). Optimal portfolios using linear programming models. Journal of the Operational Research Society, 55(11), 1169-1177.

Rubinstein, M. (2002). Markowitz's "Portfolio Selection": A Fifty-Year Retrospective. The Journal of Finance, 57(3), 1041-1045.

Sawik, B. (2013). Survey of multi-objective portfolio optimization by linear and mixed integer programming. In Applications of Management Science (pp. 55-79). Emerald Group Publishing Limited.

Topaloglou, N., Vladimirou, H., & Zenios, S. A. (2008). A dynamic stochastic programming model for international portfolio management. European Journal of Operational Research, 185(3), 1501-1524.

Xidonas, P., Hassapis, C., Mavrotas, G., Staikouras, C., & Zopounidis, C. (2018). Multiobjective portfolio optimization: bridging mathematical theory with asset management practice. Annals of Operations Research, 267(1-2), 585-606.

Zhang, Y., Li, X., & Guo, S. (2018). Portfolio selection problems with Markowitz's mean–variance framework: a review of literature. Fuzzy Optimization and Decision Making, 17(2), 125-158.

Appendix

Appendix A

```python
from scipy.optimize import linprog
import numpy as np

# Objective function coefficients
c = np.array([.29, .3, .88, .4, .12, .23, .72, .31, .1, .19])

# Inequality constraint coefficients
A = np.array([
    [395.19, 169.41, 830.41, 1241.86, 114.63, 268.69, 144.19,
    469.39, 46.84, 298.66],
    [2.13, 1, 8.16, 3.02, .25, 2.69, 1.44, 4.69, -.76, -.18],
    [-.15, -.04, .38, -.08, -.2, -.47, .04, -.39, -.57, -.54],
    [276.63, -50.82, -249.12, -372.86, -34.39, -80.61, -43.26,
    -140.82, -14.05, -89.6],
    [-118.56, 118.59, -249.12, -372.86, -34.39, -80.61, -43.26,
    -140.82, -14.05, -89.6],
    [-118.56, -50.82, 581.29, -372.86, -34.39, -80.61, -43.26,
    -140.82, -14.05, -89.6],
    [-118.56, -50.82, -249.12, 869.3, -34.39, -80.61, -43.26,
    -140.82, -14.05, -89.6],
    [-118.56, -50.82, -249.12, -372.86, 80.24, -80.61, -43.26,
    -140.82, -14.05, -89.6],
    [-118.56, -50.82, -249.12, -372.86, -34.39, 188.08, -43.26,
    -140.82, -14.05, -89.6],
    [-118.56, -50.82, -249.12, -372.86, -34.39, -80.61, 100.93,
    -140.82, -14.05, -89.6],
```

```
    [-118.56, -50.82, -249.12, -372.86, -34.39, -80.61, -43.26,
    328.57, -14.05, -89.6],
    [-118.56, -50.82, -249.12, -372.86, -34.39, -80.61, -43.26,
    -140.82, 32.79, -89.6],
    [-118.56, -50.82, -249.12, -372.86, -34.39, -80.61, -43.26,
    -140.82, -14.05, 209.06],
    [-395.19, -169.41, -830.41, -1241.86, -114.63, -268.69,
    -144.19, -469.39, -46.84, -298.66],
    [-1, 0, 0, 0, 0, 0, 0, 0, 0, 0],
    [0, -1, 0, 0, 0, 0, 0, 0, 0, 0],
    [0, 0, -1, 0, 0, 0, 0, 0, 0, 0],
    [0, 0, 0, -1, 0, 0, 0, 0, 0, 0],
    [0, 0, 0, 0, -1, 0, 0, 0, 0, 0],
    [0, 0, 0, 0, 0, -1, 0, 0, 0, 0],
    [0, 0, 0, 0, 0, 0, -1, 0, 0, 0],
    [0, 0, 0, 0, 0, 0, 0, -1, 0, 0],
    [0, 0, 0, 0, 0, 0, 0, 0, -1, 0],
    [0, 0, 0, 0, 0, 0, 0, 0, 0, -1]
])

# Inequality constraint bounds
b = np.array([10000000, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0])

# Solve the linear programming problem
res = linprog((-1)*c, A_ub=A, b_ub=b)

# Display the result
print(res)
```