CO370 Final Project (Fall 24) - Portfolio Selection

Dami Peng (dkpeng@uwaterloo.ca)
Cassidy Yan (c8yan@uwaterloo.ca)
Hyunji Ryoo (h4ryoo@uwaterloo.ca)
Yeonseo Ko (y22ko@uwaterloo.ca)

December 2024

1 Introduction

The OR problem we will address in this paper will be optimally selecting from a set of assets and maximizing the net return. This problem is one of the most well-known finance optimization models introduced by the American economist Harry Markowitz. Also modernly referred to as "Modern Portfolio Theory" (MPT), this groundbreaking framework was introduced by Harry Markowitz in his seminal 1952 essay, Portfolio selection[1]. This essay on portfolio selection provided a mathematical foundation for understanding how investors can construct portfolios to maximize expected returns while minimizing their exposed risk.

The core premise of MPT lies in diversification and the relationship between the risk and return of the underlying assets. By analyzing the statistical properties of asset returns (expected returns) and covariances, Markowitz demonstrated how combining assets into an optimal portfolio could reduce overall risk through diversification.

This paper will address a similar problem, focusing on optimizing market returns. To account for the unpredictable nature of asset returns, we will model the problem similar to robust programming.
With this change, instead of having a quadratic program, we can adapt Soyster's Robust Model and the Bertsimas-Sim Model and develop this problem as a linear program.

2 Problem Description

Assume the following: We are given a set of assets along with its:

1. Compounded Annual Growth Rate (CAGR) of each underlying asset
2. Expected Implied Volatility (IV)/Variance of each underlying asset
3. Price/Cost of each underlying asset
   - (Prices used will be dated as of December 9, 2024)

Given an initial budget $n \in \mathbb{R}$, our goal is to determine which assets to purchase and in what quantities to maximize our net return, while also preserving portfolio diversity (in our case, we do not want more than 15% of our portfolio to consist of only 1 asset). However, this is not as straightforward as merely choosing the asset with the highest average compound return.

We must also factor in the overall volatility of the stock market. Each underlying asset in the stock market (specifically the S&P 500) comes with its own market volatility/risk. When it comes to portfolio selection, this additional volatility and uncertainty make it challenging to predict market behaviour and accurately forecast the future price of a stock.

This is a common dilemma faced by many investors when deciding whether to purchase a specific security. The unpredictability of the market means that any investment carries a significant level of uncertainty and risk. An investor could either achieve substantial gains, often referred to as "hitting big" or face considerable losses, essentially "going home" empty-handed.

To minimize the uncertainty in estimating the variance of returns for each asset, we will approach this problem using a robust formula. Robust Optimization focuses on accounting for the worst-case scenario for each stock's return within the given range of their average 52-week implied volatility (IV).
For this problem, we will adapt 2 formulations:

1. Soyster's Robust Model
2. Bertsimas-Sim Model

The ideology of the formulations are using the ideas of Soyster and Bertsimas Sim to create an LP that accounts for variation in the returns.

3 Formulation

We will primarily use the idealogy from Soyster's Robust Model to formulate this problem.
First let us have a set of assets, $S = \{1, 2, 3, ..., 20\}$. Each of these assets will have:

1. a_i, Average Compound Return
2. Δ_i, Expected Volatility/Variance
3. p_i, Price/Cost
4. $\bar{a}_i$, Actual return

Our LP will have the following formulation (1.1):

$$maximize \;\; c^T x$$

s.t.

$$Ax \leq b$$

$$x \geq 0$$

Where b denotes the total initial budget we have.
As we have mentioned before, it is very difficult to actually predict $\bar{a}_i$, since the market is unpredictable in nature. As such, we will have an uncertainty set $A^{Soy} := \{\bar{a} : \bar{a} \in [a_i - \Delta_i, a_i + \Delta_i] \; , \; \forall i \in S\}$
That is, the actual return on each asset can be represented by the average compound return of the stock $\pm$ variance/volatility of the stock. Then we can rewrite (1.1) as (1.2):

$$maximize \quad \sum_{i \in S} \bar{a}_i p_i x_i - p_i x_i \qquad \forall i \in S, \forall \bar{a} = (\bar{a}_i)_{i=1}^{20} \in A^{Soy}$$

s.t.

$$\sum_{i \in S} p_i x_i \leq b$$

$$x \geq 0 \quad \forall i \in S$$

Since A^{Soy} is an infinite set, it is essentially impossible to formulate this problem without compromise. In Soyster's model, he suggests taking the worst case for the constraint $Ax \leq b$.

That is, the set $\bar{a}_i = a_i + \Delta_i$ when $x_i \geq 0$ and $\bar{a}_i = a_i - \Delta_i$ when $x_i < 0$.
In our specific problem, the worst case scenario will always be to set $a_i - \Delta_i$. This makes it so that given the asset volatility, we will always assume the worst. That is, for any asset, we will assume it will fall to the lowest price.
This gives us an objective function of:

$$maximize \sum_{i \in S} (a_i - \Delta_i) p_i x_i - p_i x_i$$

Now we also want to preserve portfolio diversity, thus we will make it such that an asset cannot be more than 15% of your portfolio:

$$p_i x_i - 0.15 \sum_{i \in S} p_i x_i \leq 0$$

Finally, we have the (Soy-LP):

$$maximize \sum_{i \in S} (a_i - \Delta_i) p_i x_i - p_i x_i$$

s.t.

$$\sum_{i \in S} p_i x_i \leq b$$

$$p_i x_i - 0.15 \sum_{i \in S} p_i x_i \leq 0$$

$$x \geq 0 \quad \forall i \in S$$

$$S = \{1, 2, .., 20\}$$

4 Formulation (Less Conservative Approach)

The formulation above considered the worst case in terms of volatility. That is, the optimal solution for the formulation above assumed that the best solution will avoid the asset with the biggest difference in Average Compound Return and Expected volatility, $|a_i - \Delta_i|$.
What if we wanted to solve the same problem, however, the client is not as risk adverse. What if we did not assume that every possible asset assumes the worst case possible? In order to model this problem, we will take inspiration from the model proposed by Bertsimas and Sim.
Let us introduce a new data variable, Γ, which will denote where at most $\bar{a}_i$ can deviate from a_i. Note that, the objective function will be

$$maximize \;\; z$$

where z = total net return on the investment. Therefore, we are now considering the following uncertainty sets for i = 1,...,20.

$A(\Gamma) := \{\bar{a} : \bar{a}_i \in [a_i - \Delta_i, a_i + \Delta_i], \forall i \in S,$ and at most Γ of the $\tilde{a}_i$ values may deviate from $a_i\}$

Then our objective function becomes:

$$c^T = \sum_{i \in S} (a_i - \Delta_i z_i) p_i x_i - p_i x_i$$

where $z_i \in \{0, 1\}$ is a binary decision variable that represents if $\bar{a}_i = a_i$. It will also follow that we need to constrict z_i to Γ:

$$\sum_{\forall i \in S} z_i = \Gamma$$

Notice that the uncertainty set is featured in our Objective function instead of the constraints (similar to the first formulation). As such our actual formulation will differ greatly from Bertsimas-Sim's robust model, as our formulation is easier to linearize.

To linearize the objective function c^T, we will provide an auxiliary viarable, $y_i = z_i x_i$ as such our objective function becomes:

$$maximize \;\; \sum_{i \in S} (a_i) p_i x_i - \Delta_i p_i y_i - p_i x_i$$

and since Δ_i, p_i are data, we have that this objective function is linear.

Now we need to set a few constraints s.t. z_i, y_i behaves as intended.
We have:

$$y_i \leq M z_i$$

$$y_i \leq x_i$$

$$y_i \geq x_i - M(1 - z_i)$$

where M is a data variable that we can set to a very high number (i.e. 10^6), s.t we constraint y_i

Finally we have the complete LP inspired by Bertsimas-Sim's robust LP: s.t.

$$maximize \quad \sum_{i \in S} (a_i) p_i x_i - \Delta_i p_i y_i - p_i x_i$$

s.t.

$$\sum_{i \in S} p_i x_i \leq b$$

$$p_i x_i - 0.15 \sum_{i \in S} p_i x_i \leq 0$$

$$\sum_{\forall i \in S} z_i = \Gamma$$

$$y_i \leq M z_i$$

$$y_i \leq x_i$$

$$y_i \geq x_i - M(1 - z_i)$$

$$x \geq 0$$

$$z \in \{0, 1\}$$

$$y \geq 0$$

$$c, y, z \in \mathbb{R}$$

$$S = \{1, 2, \ldots, 20\}$$

5 Data and Computation

An example list of stocks from the S&P 500 index fund:

i	Underlying Asset (STOCKS)	Price (USD)[1]	Compounded Annual Growth Rate (CAGR) (in %)[2]	52 week Implied Volatility (in %)[3]
1	NVDA	138.81	92.07	49.2
2	AAPL	246.75	30.26	21.4
3	TSLA	389.79	76.94	52.3
4	MSFT	446.02	24.26	19.8
5	GOOG	177.10	21.71	24.4
6	META	613.67	25.17	29.3
7	AMZN	226.09	20.66	24.5
8	COST	987.86	27.08	18.4
9	WMT	93.83	18.77	15.5
10	AXP	296.72	20.17	20.5
11	V	308.30	11.24	15.9
12	XOM	112.90	10.41	20.6
13	JNJ	149.60	1.42	15.3
14	PG	170.79	6.49	14.3
15	MA	522.82	12.70	15.7
16	HD	429.18	14.85	21.2
17	PEP	159.47	3.03	16.1
18	JPM	243.81	12.94	19.1
19	AVGO	178.94	41.81	42.9
20	UNH	560.62	14.88	22.7

[1] Pricing is as of Dec 9, 2024
[2] This will be used as average return
[3] This will be used as variance

For both the Soy-LP and Bertsimas-Sim models, we use the above data to obtain the optimal solution to maximize ROI with an initial budget of $10,000. The first column indicates the number of stocks (i = 1,...,20). The second column lists the names of the stocks, the third column shows the prices of the stocks (p_i), the fourth column presents the average rate of return (a_i), and the last column indicates the volatility (Δ_i).

We substitute these values into the models identified in the previous section and execute the code to determine the optimal allocation of stocks. The code details will be provided in separate files, and the analysis of the solutions is presented in the next section.

6 Solution Analysis

Soyster

After the optimization of Soyster's Robust Model, the solution, the optimal objective value (net return), resulted in \$1388.40, which is the result of optimizing the investment portfolio within a budget from investors. According to the model, the optimal allocation of shares that clients can get is 15.99% from Walmart, 10.81% from Nvidia, 6.08% from Apple, 3.85% from Tesla, 3.37% from AXP, 3.36% from Microsoft, and 1.52% from Costco when they invested their money. Based on the above analysis, the current portfolio results from our focus on pursuing the maximum returns clients can earn. One of the constraints we used is the budget constraint, which meant that the budget should not exceed the total investment. The second constraint was to preserve portfolio diversity, so we set the 15% limitation of the portfolio. More detailed solution analysis related to adjusting the gamma (Γ) value will be illustrated by the Bertsimas-Sim model.

Bertsimas-Sim

For the Bertsimas-Sim model, the solutions vary depending on the choice of Γ. Notice that when $\Gamma = 20$, this means that we do not constrict the uncertainty set, as such we will have our "Bertsimas-Sim" = "Soyster" Model. That is, if $\Gamma = 20$, our object value will be \$1388.40 with the following solution:

$x_1 = 10.81, x_2 = 6.08, x_3 = 3.85, x_4 = 2.24, x_6 = 2.44, x_8 = 1.52, x_{19} = 8.38$, and all other x_i's are 0.

That is, when there is the least amount of robustness, the client must buy 10.81 shares of NVDA, 6.08 shares of AAPL, 3.85 shares of TSLA, 3.36 shares of MSFT, 1.52 shares of COST, and 8.38 shares of AVGO. The expected ROI in this case is \$4642.55.

Decreasing Γ indicates lower level of uncertainty in the market. Therefore, the expected ROI increases, as we do not assume the worst case. For example, with $\Gamma = 0$, the objective value is \$4642.55 while the optimal solution remains the same.

We may interpret this as for a lower Γ, the investor's preferences would be more of a risk-taker, as to being risk-adverse.

Note that the Bertsimas-Sim model provides a more balanced allocation, as it is a less conservative approach. This results in a higher optimal value compared to the (Soy-LP)[5].

This demonstrates that as the uncertainty in the stock market increases, the expected rate of return decreases, even if the client invests in the best possible option to maximize their return. Therefore, it is recommended to start investing when the level of market uncertainty is reasonable.

7 References

1. Markowitz, H.M. (March 1952). "Portfolio Selection". The Journal of Finance. 7 (1): 77–91. doi:10.2307/2975974.

2. Comprehensive investment analysis tools: Market chameleon. MarketChameleon.com. (2024). https://marketchameleon.com/

3. Financecharts.com - powerful research for serious investors. (2024). https://www.financecharts.com/

4. Yahoo! (2024). Yahoo Finance - Stock Market Live, quotes, Business & Finance News. Yahoo! Finance. https://ca.finance.yahoo.com/

5. Bertsimas, D., Sim, M. (2004). The price of robustness. Operations Research, 52(1), 35–53. https://doi.org/10.1287/opre.1030.0065