Project Report

Indian Institute of Technology Indore

Report for Solving the k-Asset Portfolio Selection Problem

Course Instructor: Dr. Kapil Ahuja

Team Members:
Jilagam Dinesh Venkat Kumar (210001026)
Mahendraker Gaurav (210001036)
Mullapudi Surya Karthikeya (210001041)

Contents

1 Introduction

1.1 Background and Overview

Portfolios are fundamental for investors, offering diversified collections of financial assets like stocks, bonds, and securities. They spread risk and optimize returns by allocating investments across various classes, industries, and regions. This diversification mitigates the impact of poor performance by a single asset, enabling a balance between high-risk, high-return and stable, conservative options, aligning with investors' risk tolerance and financial goals. Portfolios not only spread investments but also offer potential for capital growth, income, and preservation.

Portfolios are essential for managing risk and maximizing returns, offering a strategic approach to wealth management and financial planning by diversifying investments. In dynamic financial markets, they provide flexibility to adapt to changing economic conditions and fluctuations, allowing investors to adjust risk exposure and hedge against uncertainties. This strategic approach helps shield against unforeseen events that could negatively impact a singular investment. Portfolios play a pivotal role in a well-rounded investment strategy, enabling individuals to tailor their financial holdings to their specific risk appetites and investment goals. ‘

1.2 Necessity of the problem

Portfolio optimization is pivotal in investment strategies, being the linchpin for risk management and maximizing returns. Leveraging quantitative models and algorithms, it constructs portfolios to maximize gains while minimizing risk exposure. It enables effective diversification across various asset classes, industries, or regions, ensuring a well-balanced investment approach. Optimization techniques aid in identifying the most efficient asset allocation, considering historical performance and future projections. This structured method not only enhances potential gains but also mitigates losses, aligning with an investor's risk tolerance and financial objectives.Ultimately, portfolio optimization steers investors toward a more informed and balanced investment strategy.

2 Problem

2.1 Problem Description

Within the context of an initial portfolio comprising n assets, the challenge is optimising for minimal risk within defined limitations. The restricted asset portfolio optimisation problem is centred on tailoring the investor's holdings to a limited set of k assets. The objective involves minimising risk exposure while considering multiple constraints. This process involves meticulously analysing correlations, anticipated returns, and risk factors inherent in each asset. The aim is to construct a well-diversified portfolio that adheres to the specified constraints. This tailored approach ensures that the investment strategy remains both efficient and secure despite the constraints imposed by the limited asset pool. By integrating a strategic selection of assets based on their risk-return profiles, the investor can navigate the complexities of the market, aiming for an optimal portfolio mix that aligns with their risk tolerance and financial objectives.
The Limited Asset Management Problem comes under the class of NP-Hard problems as it transforms to Mixed Integer Quadratic Program, which comes under the class of NP-Hard Problems.

$$\text{Minimize} \quad \sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}x_i x_j \tag{1}$$

$$\text{subject to} \quad \sum_{i=1}^{n} r_i x_i = R \tag{2}$$

$$x = x^0 + n \in F \tag{3}$$

$$\sum_{i=1}^{n} x_i = 1 \tag{4}$$

$$x_i <= 1 \ \forall x_i \in x \tag{5}$$

$$\sum_{i=1}^{n} |supp(x)| <= K \tag{6}$$

2.2 Two Stage Solution Approach

To improve the limited asset portfolio optimisation problem, we solve the problem in two stages. The initial method employs an iterative algorithm to compute minimum costs, constructing sets of growing sizes in each iteration. Given the exponential complexity of the first approach, the second method employs a heuristic approach to reduce the problem space. It achieves this by considering only subsets surpassing a specific threshold and discarding the rest before advancing to the next iteration. This modification enables the algorithm to operate within polynomial time. However, it's important to note that this approach serves as an approximation to the primary solution, offering an accuracy of up to two decimal places.

2.3 Increasing Set Algorithm

Let $N = \{1, \dots, n\}$. Every face of Δ has the form $\Delta_I = \{x \in \Delta : \sum_{i\in I} x_i = 1\}$, where $I \subseteq N$ is a subset of indices. Furthermore, the dimension $\dim(\Delta_I)$ of Δ_I coincides with the cardinality $|I|$ of I. Let I_K denote the family of all subsets of N with cardinality at most K. Then the cardinality constrained StQP can be reformulated as:

$$\min_{x\in\bigcup_{I\in\mathcal{I}_K}\Delta_I} f(x) = x'Qx$$

For $I \subseteq N$, let Q_I denote the submatrix of Q formed by those elements with row and column indices in I. When Q_I is positive definite, the unique global minimizer of the quadratic form $x^T Q_I x$ on the hyperplane $\sum_{i\in I} x_i = 1$ is attained at the point $x_I^* = \frac{1}{\lambda} Q_I^{-1} e$, where $\lambda = 1/(e^T Q^{-1} e)$.
The above equations for λ and x_I^* can be derived from KKT conditions as follows:

$$\begin{aligned}
\sum x_i &= 1\\
e^\top x &= 1\\
\nabla f(x) &= \lambda \nabla a(x)\\
Qx &= \lambda e\\
x &= \lambda Q^{-1} e\\
e^\top x &= \lambda e^\top Q^{-1} e\\
1 &= \lambda e^\top Q^{-1} e\\
\lambda &= \frac{1}{e^\top Q^{-1} e}
\end{aligned}$$

Thus the quadratic form $f(x)$ has a global minimizer on Δ in the relative interior $\mathrm{rint}(\Delta_I)$ of a face Δ_I where f is strictly convex only if $x_I^* \in \mathrm{rint}(\Delta_I)$.
To every subset $I \subseteq N$ we associate the (nonlinear) weight.

$$w(I) = \min\{f(x) : x \in \Delta_I\}$$

$$\min_{x\in\bigcup_{I\in\mathcal{I}_K}\Delta_I} x'Qx = \min_{I\in\mathcal{C}_K} w(I) = \min_{I\in\mathcal{C}_K} f(x_I^*) = \min_{I\in\mathcal{C}_K} \frac{1}{4} e' Q_I^{-1} e,$$

$$C_K = \{I \in I_K : Q_I \text{ is positive definite and } x_I^* \in \mathrm{rint}(\Delta_I)\},$$

where C_K is the subset of I_K defined by the conditions specified.

For $j \leq K$, let $C_j = \{I \in C_K : |I| = j\}$. Then it is ensured that for any I^* minimizing $w(I)$ on C_K, there exists a sequence $I_1 \subseteq I_2 \subseteq \ldots \subseteq I_h = I^*$, such that $I_j \in C_j$ for all $j = 1, \ldots, h$. Thus we can apply the following Increasing Set Algorithm algorithm to solve the cardinality-constrained StQP or, equivalently, to minimize $w(I)$ on C_K.

Note that, at any iteration j, $MIN(j)$ contains the minimum value of $w(I)$ among all sets in C_j. Furthermore, if $C_{j+1} = \emptyset$ in step 5 of the pseudo code in the section 2.4, then, again by the above theorem, C_h must be empty for all $h \geq j + 1$. Hence the algorithm correctly stops with the global minimizer in C_K. In fact, at each iteration, j, the Increasing Set Algorithm provides in $MIN(j)$ the solution to the StQP problem with cardinality constraint $|\mathrm{supp}(x)| \leq j$.

So in summary, we are converting the LAM Problem into a Standard Quadratic Program by considering various sets of assets with *cardinality* $<= k$. We are solving this Standard Quadratic Program as a subproblem using KKT conditions.

2.4 Pseudo-code

First, we define MIN(i) as the minimum risk over all sets of cardinality i.The minimum risk corresponding to the set is calculated by considering only those assets whose indices belong to the set, for this corresponding set the problem is reduced to a Standard Quadratic Problem with equality constraint, this sub-problem can be solved using the KKT conditions to find the minimum risk. C_j corresponds to the collection of all the feasible sets of cardinality j.

While constructing the sets in C_{j+1} by appending elements to set C_j we check if the weights corresponding to the minimum risk are all positive and if the Q matrix formed is positive definite. We stop the Algorithm if the C_{j+1} is empty with no feasible sets and we return MIN(j) as the minimum risk for the given LAM problem.

Increasing Set Algorithm()

1 Set $\mathcal{C}_0 \leftarrow \emptyset,\ \mathcal{C}_1 \leftarrow \{\{i\}, i \in N\}$

2 $\mathrm{MIN}(1) \leftarrow \min_{I \in \mathcal{C}_1} w(I) = \min_{1 \leq i \leq n} q_{ii}$

3 **for** $j = 1$ to K

4   **do** construct $\overline{\mathcal{C}}_{j+1}$ by increasing, if possible, all elements in $\overline{\mathcal{C}}_j$

5   **if** $\overline{\mathcal{C}}_{j+1} = \emptyset$

6   **then** $\mathrm{MIN}(h) \leftarrow \mathrm{MIN}(j)$ for $h = j+1, \ldots, K$, **return** $(\mathrm{MIN}(K))$

7   **else** $\mathrm{MIN}(j+1) \leftarrow \min\{\mathrm{MIN}(j),\ \min_{I \in \overline{\mathcal{C}}_{j+1}} w(I)\}$

8   **return** $(\mathrm{MIN}(K))$

2.5 Heuristic Approach

Based on the suggestions given in the research paper we designed our own algorithm by implementing a heuristic approach where after constructing C_{j+1} we will sort the entries based on risk and select only the top p elements for the next iteration. This step will imply that the size of C_{j+1} for the next iteration is bounded by p thus making the time complexity of the algorithm as polynomial.

Set $C_0 \leftarrow \emptyset$, $C_1 \leftarrow \{\{i\}, i \in \mathrm{N}\}$
$MIN(1) \leftarrow min_{I \in C_1} w(I) = min_{1 \leq i \leq n} q_{ii}$
for $j = 1$ *to* K
    do construct C_{j+1} by increasing, if possible all elements in C_j
    sort C_{j+1} based on the corresponding risk associated
    select the first p elements from C_{j+1}
    if $C_{j+1} = \emptyset$
        then $MIN(h) \leftarrow MIN(j)$ for $h = j + 1, \ldots, K$ **return** $(MIN(K))$
        else $MIN(j + 1) \leftarrow min\{MIN(j), min_{I \in C_j'} w(I)\}$
return $(MIN(K))$

2.6 Complexity analysis

The algorithm described in the section 2.4 runs for exponential complexity as it involves generating sets from the previous set by appending elements and then discarding the elements which do not satisfy the condition. This is done for k iterations and the minimum risk over all the iterations is taken for the optimal asset allocation. This algorithm is similar to the breadth first search in a tree. In the worst case we would be generating all the 2^k subsets so the time complexity would be exponential. On the contrary, in the heuristic approach since we limit the size of C_{j+1} by p after every j^{th} iteration, the heuristic approach runs in a polynomial time. So the runtime complexity of Increasing Set Algorithm is $O(2^n)$ and that of heuristic algorithm is $O(n * k * p)$.

2.7 Convergence

The algorithm converges to minimum if the initial Q matrix (covariance matrix) is positive definite this is because the standard quadratic subproblem converges to minimum value if Q is positive definite (by applying the sufficient conditions(KKT) on the standard quadratic subproblem). Solution of the original problem is the minimum over all subproblems. The value $x = \lambda Q^{-1} e \ with \ \lambda = \frac{1}{e^\top Q^{-1} e}$ is minima of the subproblem only if Q is positive definite.

3 Datasets and Results

3.1 Datasets

An important issue for evaluating computational results for a class of problems is the availability of benchmark data sets, possibly with solutions, that can be used by researchers to compare the efficiency of their algorithms, and the quality of the solutions obtained in the case of heuristics.We have used the dataset containing weekly price data of real world markets, available on OR Library.We have run our heuristic algorithm on this dataset and compared its results with a Portfolio MIQP Solver available on MATLAB by modifying it to solve the LAM Problem.

3.2 Computational Results

N = 225	Solver		Our algorithm	
K	Time	Minima	Time	Minima
10	>15min	-	195s	0.33624
40	68s	0.324913	219s	0.33286
100	58s	0.324913	225s	0.33285
225	57s	0.324911	225s	0.33285

The above results are obtained by running both the MATLAB Solver and our algorithm on the port5 dataset of OR Library.

The complexity of the exact Increasing Set Algorithm is **exponential** in time complexity as in worst case we might have to compare all sets of assets with *cardinality* $<= k$. So by bounding the size of C_j in our heuristic algorithm, we have achieved a **Polynomial time algorithm**. As the table indicates, with an increase in the value of k, the corresponding minima values consistently decrease, aligning with our expectations. **Additionally, noteworthy is the observation that the minima values achieved by our algorithm closely approximate those attained by the solver. From the table it is clearly visible that the minima for lesser values of k for the solver will take more than 15min but on the other hand our algorithm finds the optimal value in less than 4min.**

3.3 Code Implementation

- Our Implementation of both Increasing Set and Heuristic Approach in python.
- Matlab solver for LAM problem.
  Above are our implementations for the Increasing Set Algorithm, Heuristic Approach, and the Solver.

4 Acknowledgment

We thank Dr. Kapil Ahuja for teaching us the course: OPTIMIZATION TECHNIQUES AND ALGORITHMS with utmost finesse. As a computer science student this course is crucial for any career path we wish to pursue, and to have learnt it with this level of clarity is our pleasure. Moreover, this project, along with many other lab assignments, have enhanced not only our theory application but also has given hands-on coding experience with real-life task-based questions. So, once again we show our gratitude towards our professors as well as our TAs, who made sure our learning went on smoothly.

5 References

- Francesco Cesarone, Andrea Scozzari‡ , Fabio Tardella Universit‘a di Roma “La Sapienza” Dip. Matematica per le Decisioni Economiche, Finanziarie ed Assicurative
- Hybridising Local Search With Branch-And-Bound For Constrained Portfolio Selection Problems.
- H. M. Markowitz, Portfolio selection, Journal of Finance 7 (1952), no. 1, 77–91.
- Proof that MIQP is NP Complete.
- J.E. Beasley, Or-library: Distributing test problems by electronic mail, Journal of the Operational Research Society 41 (1990), 1069–1072..