Constraint Satisfaction Problem: Personalized Stock Portfolio

Devisha Arun, Matus Bernadovic, Julie Dawkins, Alexandra Dobre

Abstract

The rise of retail investors in financial markets has highlighted a need for personalized portfolio guidance. This paper presents a novel algorithm that uses a refined approach to the Constraint Satisfaction Problem (CSP) framework to optimize allocations of stocks based on individual preferences, such as risk tolerance, return expectations, and sustainability ratings. By modeling stocks as variables, allocations as domains, and financial goals as constraints, we ensure that portfolios align with investor objectives. Our approach bridges a gap in CSP literature by addressing portfolio optimization as a resource allocation problem, offering a practical tool for retail investors to make informed decisions and enhance financial inclusion.

Disclaimer

The content presented in this paper is for educational and illustrative purposes only. It does not constitute financial advice or a recommendation to buy, sell, or hold any securities or investments. Investments in financial markets involve risks, including the potential loss of capital.

1 Introduction

The prominence of retail investors, that is, private individuals, has grown significantly since the advent of the digital revolution and the increased accessibility of financial markets. Despite this rise in accessibility driven by the emergence of trading platforms such as eToro, Trading 212, and Binance, only a limited number of these platforms provide investors with practical guidance in constructing their investment portfolios.

Without such guidance, investors are often subject to behavioral biases, leading to irrational decisions that may not align with their actual preferences, such as risk tolerance, investment horizon, or available capital. Although some asset management firms offer personalized portfolio solutions, these often require lengthy and tedious questionnaires and remain largely inaccessible or unfamiliar to the general public.

Our objective is to design an algorithm that takes as input the investor-specific preferences expressed through financial indicators and ratios. These may include metrics such as volatility to represent idiosyncratic risk, 1-, 5-, or 10-year returns based on the desired investment horizon, or desire to invest in sustainable companies. Using these constraints, the algorithm will construct a personalized portfolio by allocating a percentage of the investor's capital to individual stocks or indices.

To solve this algorithm, we leverage the constraint satisfaction problem framework (CSP). CSPs were first developed in the field of computer science and artificial intelligence in the 1970s and have since been used to solve a variety of real-world challenges, such as scheduling conflicts and resource allocation Mackworth (1977). A CSP is constructed through variables, domains, and constraints; the goal is to assign values to the variables—where all possible values are defined in the domain—such that all constraints are satisfied Bulatov (2018).

Researchers have extensively explored CSPs, with a particular emphasis on algorithms and complexities, a focus that began in the 1990s with Feder and Vardi (1998) Bulatov (2018). Recent advances in CSP methodologies have enabled more efficient exploration of large and intricate solution spaces, making CSPs increasingly valuable in addressing practical challenges Ayadi et al. (2021).

Alongside these theoretical papers, a subset of the literature has tackled a wide variety of real-world challenges in a constraint satisfaction framework (for example, Åstrand et al. (2020) examines mine scheduling, while Gao et al. (2018) uses constraints to improve crowd counting in video form; Švaco et al. (2023) solves pallet loading, and Lakshmanna and Khare (2016) uses a group optimization algorithm for DNA sequence mining). However, we have found few papers that examine financial challenges using constraint satisfaction problems. Winkel et al. (2023) uses reinforcement learning alongside allocation constraints in order to optimize a portfolio. Bandi and Tulabandhula (2020) also takes a reinforcement learning approach with constraints, further illustrating the potential for a constraint satisfaction approach to portfolio optimization as a resource allocation problem. The vast majority of articles focused on artificial intelligence and financial investing center around machine learning and large language model frameworks to predict future stock movement and improve decision making, rather than working with individual-level preferences or CSPs Bi et al. (2024); Yang et al. (2023); Papasotiriou et al. (2024). As such, exploring investment through the lens of retail investor preferences and constraint satisfaction problems addresses a gap in the CSP literature.

The challenge of capital investment is well-suited to a constraint satisfaction problem because the preferences of an individual can be well defined via constraints in order to solve for the ideal investment in stocks. Moreover, this is a highly applicable topic and solves a strong gap: the knowledge gap for a retail investor who can now better understand where to best invest their funds. In this paper, we use a convex optimization framework and the GNU Linear Programming Kit (GLPK), an open-source optimization solver, to solve the constraint satisfaction problem. This allows us to generate the ideal investments for 3 simulated investors with varying levels of risk tolerance, desired return, and sustainability preferences. In this paper, we have created a novel tool to easily invest for a lay person where we begin by discussing the problem formulation and framework, outline a proposed solution, apply numerical experiments to the solution, and then finish with conclusions.

2 Problem Formulation

The optimization of a personalized stock portfolio can be effectively modeled as a Constraint Satisfaction Problem (CSP), which allows systematic representation and resolution of multiple interconnected requirements. In the following section, we define the problem using the CSP framework and elaborate on its components and significance.

2.1 CSP Framework for Portfolio Optimization

A CSP is defined by three core components: variables (x), domains (D), and constraints (C). Each component plays a critical role in ensuring that the solution satisfies the investor's preferences, financial requirements, and market realities.

2.1.1 Variables (X):

The variables in the CSP, $X = \{X_1, X_2, \ldots, X_n\}$ represent the individual stocks in the portfolio. Each variable corresponds to a decision: What percentage of the total investor capital should be allocated to a specific stock? For example:

- X_1: Allocation percentage for stock A.
- X_2: Allocation percentage for stock B.
- X_3: Allocation percentage for stock C.
- ...
- X_n: Allocation percentage for stock n.

The primary goal is to determine a valid and optimal allocation for each stock.

2.1.2 Domain (D):

The domains, $D = \{D_1, D_2, \ldots, D_n\}$, define the set of possible values each variable X_i can take.

- D_i represents the percentage of the total investment capital allocated to stock X_i.

Each domain is constrained within the range $\in [0, 10]$, ensuring that allocations are non-negative and do not exceed 10%. This guarantees the portfolio is diverse and avoids a solution where only a few stocks are allocated non-zero weights.

2.1.3 Constraints (C):

The constraints $C = \{C_1, C_2, \ldots, C_n\}$ are critical to ensuring the portfolio meets the investor's financial goals, preferences, and market expectations. Each constraint imposes a specific requirement that the allocation percentages must satisfy. The following are the key constraints in this problem:

C1: Sum of Weights Constraint

The total allocation of capital across all n stocks must be equal to 100% (or 1 in normalized form). This ensures the entire capital is utilized, leaving no funds unallocated. Mathematically, if w_i is the allocation weight of a single stock:

$$\sum_{i=1}^{n} w_i = 1$$

This constraint also prevents the short-sell of securities, as retail investors rarely engage in this practice. If short-selling is desired, the constraint could be modified to allow any w_i to take negative values.

C2: Maximum Allocation Constraint

The allocation for any single stock must not exceed 10% of the total capital. This constraint promotes diversification and prevents over-concentration in any one stock. Mathematically:

$$w_i \leq 0.10, \forall i \in \{1, 2, \ldots, n\}$$

This ensures a balanced portfolio by limiting excessive exposure to individual stocks.

C3: Minimum Return Constraint

The expected return of the portfolio, calculated as the weighted average of individual stock returns, must meet or exceed the minimum return required by the investor. This establishes that the portfolio generates satisfactory performance. Numerically:

$$\text{Portfolio Expected Return} = \sum_{i=1}^{n} (w_i \cdot \text{Expected Return}_i) \geq \text{Minimum Required Return}$$

For example, if the minimum required return is 15%, the portfolio must allocate capital to stocks in such a way that the overall return meets this threshold.

C4: Maximum Volatility Constraint

The portfolio's total idiosyncratic volatility, calculated as a weighted sum of individual stock idiosyncratic volatilities, must not exceed the maximum risk tolerance specified by the investor. Mathematically:

$$\text{Weighted Average Idiosyncratic Volatility} = \sum_{i=1}^{n} (w_i \cdot \text{Idiosyncratic Volatility}_i) \leq \text{Maximum Risk}$$

This constraint minimizes the risk of the portfolio while adhering to the investor's risk preferences. For more information on how idiosyncratic volatility was calculated, refer to Section 4.

C5: Minimum ESG Constraint

The portfolio's weighted average ESG score must meet or exceed the minimum threshold set by the investor. This ensures alignment with sustainable and ethical investment practices. Mathematically:

$$\text{Weighted Average ESG Score} = \sum_{i=1}^{n} (w_i \cdot \text{ESG}_i) \geq \text{Minimum ESG Score}$$

This constraint incorporates sustainability considerations into the portfolio optimization process.

C6: Beta Range Constraint

The portfolio's beta, representing its systemic risk, must fall within a specified range determined by the investor. This range is defined by the ideal beta and a tolerance level. Mathematically:

$$\text{Weighted Average Beta} = \sum_{i=1}^{n}(w_i \cdot \beta_i) \geq \text{Ideal Beta} - \text{Beta Tolerance},$$
$$\text{Weighted Average Beta} = \sum_{i=1}^{n}(w_i \cdot \beta_i) \leq \text{Ideal Beta} + \text{Beta Tolerance}.$$

This ensures the portfolio achieves the desired level of market sensitivity while adhereing to the investor's specified tolerance level.

3 Proposed Solution

The proposed solution is to utilize a computational optimization framework to construct a personalized stock portfolio that aligns with the investor's preferences. The objective is to maximize the total portfolio return subject to a set of constraints designed to ensure the portfolio meets the investor's requirements. These constraints include the full allocation of the investor's capital, diversification through a cap on individual stock allocations, risk management through maximum portfolio idiosyncratic volatility, and alignment with sustainability goals by enforcing a minimum weighted ESG score. Additionally, the portfolio's beta, representing sensitivity to market movements, is constrained to remain within a specified range. The integration of these constraints creates a structured approach to solving the portfolio optimization problem.

In order to tackle the issue of computational complexity due to our continuous domain, we implemented a CSP solution via the `CVXPY` library and its GLPK solver. It provides an efficient platform for solving convex optimization problems. As demonstrated in Section 2, the CSP of portfolio creation can be considered in this framework. Both the objective function (return maximization) and the constraints have convex representations as linear (in)equalities.

We ensured that the optimization framework is designed to balance all our constraints simultaneously, ensuring a feasible solution that satisfies the investor's preferences. Once the optimization problem is solved, we use the optimal weights to calculate key portfolio metrics, such as total return, volatility, beta, and ESG score. These metrics are validated to confirm that the portfolio adheres to all constraints. We then ranked the stocks by their allocation weights, providing the investor with a clear strategy for capital allocation. This helps us in ensuring that the portfolio delivers the desired return while maintaining acceptable levels of risk and sustainability.

4 Numerical Experiments

The problem solution has been tested on real-world financial data obtained from the Bloomberg database, focusing on equities (stocks) traded on the London Stock Exchange with a reference point of 1 January 2025. After data cleaning and outlier detection, the dataset contains parameters for 1,211 distinct equities. This approach ensures the data is both robust and representative of the underlying real-world market conditions.

For each stock in the dataset, we obtained or calculated the following key parameters: 1-, 5-, and 10-year annualized returns, 20-year annualized idiosyncratic volatilities (standard deviations of daily logarithmic returns), Bloomberg ESG scores and Betas. By incorporating returns across various timeframes (1, 5, and 10 years), we have the ability to address a range of investor preferences and investment outlooks, catering to both short- and long-term strategies. The 20-year annualized idiosyncratic volatility is used to quantify stock-specific risk that cannot be diversified away, providing a measure of inherent risk that aids in precise portfolio construction. We define Beta as the 1-year Beta of daily logarithmic returns, reflecting its most common real-world application. Beta is instrumental in evaluating a stock's sensitivity to market movements, representing the degree of systemic risk, or the extent to which a stock's returns are expected to converge with the broader market. Furthermore, we operationalize the sustainability of a stock using the Bloomberg ESG rating, which ranges from 1 to 10, with 10 being the best. This score integrates environmental, social, and governance factors, adding an essential dimension for socially conscious investment strategies.

Notably, since the 20-year annualized idiosyncratic volatilities were not directly available in the Bloomberg database, it was calculated using the following relationship between the total variance of the stock, its Beta, and the market variance:

$$\sigma_{\text{idiosyncratic}} = \sqrt{\sigma_{\text{total}}^2 - \left(\beta^2 \cdot \sigma_{\text{market}}^2\right)}$$

After data cleaning and importing, we explored various methods for defining the variables and domains, experimenting with different Python modules and classes to identify the most effective and efficient solution to the problem.

To evaluate the different solutions, we considered two primary performance measures: (1) runtime and computational efficiency, and (2) the ability to return a valid solution given the constraints. Initially, we attempted to solve our CSP using the solver introduced in class, `python-constraint`, as well as the CSP class from the AIMA library. While these approaches were theoretically correct after discretizing the domain, they struggled with scalability due to their reliance on brute-force or neighbor-based solutions, which were computationally prohibitive given the number of stocks and constraints in our problem. We ultimately adopted a convex optimization approach, which is well-suited for constrained problems with continuous variables and where constraints form a convex region Russell and Norvig (2022). Because our stock allocations are continuous variables and our constraints are linear (in)inequalities, this was an appropriate solution. Moreover, convex optimization scales efficiently, resolving the computational complexity issue and making it the best choice for addressing this portfolio optimization challenge.

To test our constraint framework, we designed experiments to explore how hypothetical retail investors might shape their preferences and the resulting portfolios. During these experiments, we realized that less-experienced retail investors could easily create infeasible combinations of constraints, causing the constraint satisfaction problem to fail. To address this, we developed an input system for our hypothetical investors that determines specific constraint values based on a user's responses on time horizon, risk, sustainability, and market preferences. For example, an amateur investor might select unrealistically low volatility for their desired return, despite these two preferences being inherently linked. We examined the underlying distribution of the stock data and determined realistic boundaries for these ranges: for example, the minimum volatility for a low-risk investor; the smallest acceptable range around beta; and the maximum ESG value a portfolio can attain. Ultimately, we developed **three example investors** with varying time horizons, risk preferences, sustainability, and market preferences, to demonstrate the adaptability of our framework.

Below we present the optimal stock allocation for an investor with a long investment horizon, who wants to minimize risk, values sustainability, and wants a portfolio that is not strongly tied to market movement. The other two investors' results, with different preferences, can be found in the attached Colab notebook.

Metric	**Value (Constraint)**
Portfolio Volatility	1.00 (Max Allowed: 1)
Portfolio ESG	6.00 (Min Required: 6)
Portfolio Beta	0.75 (Range: ±0.05: 0.8)
Portfolio Return	35.27 (Min Required: 6)

Table 1: Portfolio Metrics and Constraints

Stock	**Weight**	**Return**	**ESG**	**Volatility**	**Beta**
GABILNEquity	0.100000	70.54	5.55	0.86	0.41
GOO2LNEquity	0.100000	61.71	6.34	1.98	0.80
1AMZLNEquity	0.100000	44.98	5.35	0.95	0.87
AOFLNEquity	0.100000	24.40	6.95	0.57	0.45
DECLNEquity	0.100000	37.62	4.67	1.09	1.30
IGLNLNEquity	0.100000	7.88	6.81	0.24	0.52
SPALLNEquity	0.100000	0.88	6.20	0.65	1.26
EJFILNEquity	0.100000	31.07	5.43	0.34	0.35
GBSLNEquity	0.082736	7.65	6.99	0.67	0.51
1TSLLNEquity	0.067510	65.22	6.39	2.97	1.06
CNEGLNEquity	0.027910	11.54	5.17	0.45	1.03
ADVTLNEquity	0.021844	91.73	5.31	2.89	0.57

Table 2: Optimal Weights and Stock Characteristics

5 Limitations and Further Research

With the main focus of this paper being the technical illustration of how the CSP framework can be applied to the stock allocation problem, there is significant potential to implement additional portfolio optimization techniques in future work.

To make our findings comprehensible to a wider audience, we chose to use the most widely recognized risk measure: the annualized standard deviation of the daily logarithmic returns. However, in modern investment risk management, this measure is insufficient, as it lacks both coherence and convexity. To properly reflect an investor's risk appetite in the portfolio, more robust measures such as Value at Risk (VaR) or Expected Shortfall (ES) should be employed. Additionally, to better isolate idiosyncratic volatility, sector, country, and industry-specific effects should be taken into account. In a real-world application, diversification effects should be obtained via consideration of covariances of individual stocks. Finally, the approximation of portfolio or stock returns could be improved by employing

Stochastic Discount Factor models, which offer a more theoretically grounded approach compared to relying solely on historical data.

The solution we propose in this paper is flexible, and allows for personalization and adjustment of the constraints to fit the specific demands of the investor. Factors such as required dividend yield, sector orientation, return on assets, or other financial ratios and metrics can be incorporated into the model. This is particularly relevant for existing investment platforms that already have access to detailed data on individual stocks. If our portfolio construction algorithm were to be implemented, these platforms could easily tailor the constraints to align with their users' preferences and requirements. Furthermore, the approach is not limited to stocks. Provided that the respective securities are traded on the platform, the algorithm could be extended to include other investment options in the domain, such as indexes, bonds, or financial derivatives. This would significantly improve the flexibility and applicability of the model, offering investors a wider range of customized portfolio optimization opportunities. The only possible caveat lies in the case where the desired constraints do not have a convex representation, as convex optimization relies on this assumption to solve the CSP.

6 Conclusion

In this report, we present a robust and scalable approach to the personalized stock portfolio optimization using a CSP framework. By incorporating key investor preferences such as risk tolerance, expected returns, and sustainability goals, our methodology effectively bridges the gap between traditional portfolio theory and modern retail investor needs. Using the `cvxp` library and its GLPK solver, we were able to solve the portfolio optimization problem with convex constraints, ensuring efficient solutions for large datasets. Our approach demonstrates that CSP is a viable and powerful approach to solving complex investment problems by systematically balancing multiple objectives and constraints through the integration of constraints, such as the sum of weights, maximum allowable volatility, minimum portfolio return, etc. Our proposed framework aligns portfolios with diverse investor preferences.

The numerical experiments validate the computational efficiency of our proposed approach, demonstrating that the model runs in a reasonable time and yields feasible optimized portfolios that satisfy all constraints. Our approach uses concepts from portfolio theory that can and should be amended in the event of integration into a trading platform. It relies on standard deviation for idiosyncratic risk measurement and historical data for returns, which should be re-evaluated. Future research could incorporate robust measures such as Value at Risk (VAR) or coherent ones such as Expected Shortfall (ES), to better express portfolio's risk. Expanding the model to allow short-selling and include securities such as bonds or derivatives could further enhance its versatility for real-world applications.

In conclusion, our research contributes to academic and practical discussions on personalized portfolio optimization. By using CSP methodologies and modern computational tools, the study offers a flexible and accessible solution for trading platforms to satisfy retail investors seeking tailored investment strategies, paving the way for further innovation in algorithmic portfolio construction.

References

Zouhayra Ayadi, Wadii Boulila, and Imed Riadh Farah. A hybrid apm-cpgso approach for constraint satisfaction problem solving: Application to remote sensing. In *International Conference on Knowledge-Based Intelligent Information & Engineering Systems*, 2021. URL `https://api.semanticscholar.org/CorpusID:235377285`.

Nymisha Bandi and Theja Tulabandhula. Off-Policy Optimization of Portfolio Allocation Policies under Constraints. Papers 2012.11715, arXiv.org, December 2020. URL `https://ideas.repec.org/p/arx/papers/2012.11715.html`.

Shuochen Bi, Wenqing Bao, Jue Xiao, Jiangshan Wang, and Tingting Deng. Application and practice of ai technology in quantitative investment. *Information Systems and Economics*, 5, 2024.

Andrei A Bulatov. Constraint satisfaction problems: Complexity and algorithms. *ACM SIGLOG News*, 5(4):4–24, 2018.

Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic snp and constraint satisfaction: A study through datalog and group theory. *SIAM Journal on Computing*, 28(1):57–104, 1998.

Liqing Gao, Yanzhang Wang, Xin Ye, and Jian Wang. Crowd counting considering network flow constraints in videos. *IET Image Processing*, 12(1):11–19, 2018.

Kuruva Lakshmanna and Neelu Khare. Constraint-based measures for dna sequence mining using group search optimization algorithm. *International Journal of Intelligent Engineering and Systems*, 9(3):91–100, 2016.

Alan Mackworth. Consistency in networks of relations. *Artificial Intelligence*, 8(1):99–118, 1977.

Kassiani Papasotiriou, Srijan Sood, Shayleen Reynolds, and Tucker Balch. Ai in investment analysis: Llms for equity stock ratings. In *Proceedings of the 5th ACM International Conference on AI in Finance*, 2024. URL `https://arxiv.org/abs/2411.00856`.

Stuart Russell and Peter Norvig. *Artificial Intelligence A Modern Approach: Fourth Edition*. Pearson, 2022.

David Winkel, Niklas Strauß, Matthias Schubert, and Thomas Seidl. Simplex decomposition for portfolio allocation constraints in reinforcement learning. In *ECAI 2023 - 26th European Conference on Artificial Intelligence*, 2023. URL `http://arxiv.org/abs/2404.10683`.

Yi Yang, Yixuan Tang, and Kar Yan Tam. Investlm: A large language model for investment using financial domain instruction tuning. *arXiv preprint*, abs:2309.13064, 2023. URL `http://arxiv.org/abs/2309.13064`.

Max Åstrand, Mikael Johansson, and Alessandro Zanarini. Underground mine scheduling of mobile machines using constraint programming and large neighborhood search. *Computers and Operations Research*, 123:105036, 2020.

Marko Švaco, Filip Šuligoj, Bojan Šekoranja, and Josip Vidaković. Solving pallet loading problem with real-world constraints. *arXiv preprint*, abs:2307.11531, 2023. URL `https://arxiv.org/abs/2307.11531`.