https://doi.org/10.15407/scine21.02.101

SHOLOPAK, V. A. (https://orcid.org/0000-0001-7492-3578),
and TRETIAK, D. D. (https://orcid.org/0000-0002-1610-6905)

Taras Shevchenko National University of Kyiv,
64/13, Volodymyrska St., Kyiv, 01601, Ukraine,
+380 44 239 3329, info@knu.ua

MEAN-VARIANCE OPTIMIZATION: MODELING AN OPTIMAL INVESTMENT PORTFOLIO IN THE U.S. TECH SECTOR

Introduction. *Modern Portfolio Theory (MPT) provides a quantitative framework for making informed investment decisions. The highly variable and uncertain U.S. technology sector challenges traditional investment approaches, necessitating methods that better address its unique risk-return trade-offs.*

Problem Statement. *Traditional investment strategies frequently fail to capture the dynamic and volatile nature of the tech market. They rely on limited data and inefficient calculation processes, resulting in suboptimal asset allocation. One of the advanced methods for refining portfolio formation strategies tailored to the tech market is the mean-variance optimization (MVO) method.*

Purpose. *To optimize mean-variance optimization (MVO) to construct optimal portfolios for the U.S. tech sector, leveraging contributions from MPT, Sharpe's optimization techniques, and Tobin's asset allocation model.*

Materials and Methods. *Historical stock data serves as the basis for implementing MVO with Python to construct portfolios that include a risk-free asset, enabling the calculation of the Capital Allocation Line (CAL) and the upper Efficient Frontier. The geometric mean evaluates expected returns, improving long-term predictability and portfolio comparability, while daily returns enhance the model's sensitivity.*

Results. *The study has demonstrated that optimized portfolios achieve higher Sharpe ratios and superior risk-return characteristics, outperforming benchmarks through efficient computation.*

Conclusions. *The MVO is an effective investment tool for the tech sector, enabling informed asset selection and portfolio construction. This study has highlighted the importance of integrating iterative calculation processes and advanced computational techniques to adapt traditional investment strategies to the extensive data requirements of today's markets.*

Keywords: optimization, capital allocation line, efficient frontier, Python programming.

Citation: Sholopak, V. A., and Tretiak, D. D. (2025). Mean-Variance Optimization: Modeling an Optimal Investment Portfolio in the U.S. Tech Sector. *Sci. innov.*, 21(2), 101—114. https://doi.org/10.15407/scine21.02.101

In the realm of investment, particularly within the dynamic U.S. tech sector, constructing a portfolio that navigates the balance between risk and return is an ongoing challenge. The concept of mean-variance optimization emerges as a crucial framework in this endeavor, offering a systematic approach to achieving optimal portfolio construction. As markets continue to evolve, characterized by rapid technological advances and unpredictable fluctuations, the need for sophisticated, yet accessible, investment strategies become increasingly apparent.

This paper explores the mean-variance optimization (MVO) model as a tool for informed investment decision-making in a complex landscape. It simplifies navigating uncertainties by using prices as the sole input, which, while not detailing the reasons behind asset fluctuations, aggregates all influences into one comprehensive indicator, offering a streamlined approach to investment choices.

Portfolio investing is cross-border transactions and positions in debt, equity or derivative securities, where the allocation and management of diverse but combined investment objects are scientifically based to form a diversified portfolio with the highest level of expected return for a given level of risk.

The streamlined methodology for optimal investment portfolio analysis and modeling can be divided into the following stages:

1. Data Collection. Select a timeframe and gather data on asset price movements from reputable sources.

2. Return Calculation. Compute actual returns for each asset to assess past performance.

3. Expected Return Evaluation. Calculate expected returns using historical data to forecast future portfolio performance.

4. Risk Assessment. Measure asset and portfolio volatility, including covariance analysis to understand how asset values move in relation to each other, impacting portfolio diversification.

5. Optimal Asset Allocation. Employ economic-mathematical models and software to find the ideal asset weights, optimizing the risk-return balance.

6. Portfolio Evaluation and Selection. Compare the selected portfolios against benchmarks, focusing on risk, return, and the Sharpe ratio for a comprehensive performance evaluation.

Modern Portfolio Theory (MPT), or mean-variance analysis, is a mathematical framework for assembling a portfolio of assets such that the expected return is maximized for a given level of risk. It is a formalization and extension of diversification in investing, the idea that owning different kinds of financial assets is less risky than owning only one type. Its key insight is that an asset's risk and return should not be assessed by itself, but by how it contributes to a portfolio's overall risk and return. The variance of return (or its transformation, the standard deviation) is used as a measure of risk because it is tractable when assets are combined into portfolios [1].

Economist Harry Markowitz introduced MPT in a 1952 essay, which represents a mathematical optimization problem that is used to construct the Efficient Frontier in the context of portfolio selection.

Markowitz (1952, 1959) first introduced the mean-variance framework for portfolio selection, which determines optimal portfolio choices for individual investors given asset prices and payoff distributions [2]. Tobin (1958) expanded on this by incorporating a risk-free asset into the portfolio choice problem [3]. Building upon the foundational work of Markowitz and Tobin, Treynor (1961) and Sharpe (1963, 1964) independently developed a "general equilibrium" model that retained the mean-variance approach but allowed for endogenous asset pricing. Their contributions, along with the refinements by Mossin (1966) and Lintner (1965, 1969), culminated in what is now known as the Capital Asset Pricing Model (CAPM). This model is a key component of Modern Portfolio Theory and has become a cornerstone in the field of financial economics [4].

This paper will utilize a focused set of models for portfolio construction and analysis, specifically drawing on the foundations of Harry Markowitz's Modern Portfolio Theory and the optimization techniques proposed by William Sharpe. Additionally, it will incorporate James Tobin's asset allocation model to explore the interplay between risk-free assets and the portfolio's risky asset mix within the mean-variance optimization (MVO) framework.

G. Markowitz, the founder of MPT, proposed to evaluate the random return of a deterministic investment portfolio by two indicators — mathematical expectation and variance, and to choose the best investment portfolio from the set of efficient portfolios of a two-criteria problem with the criteria of maximizing the expected return and minimizing the variance of the return. Mathematically, the task of optimizing the investment portfolio is a linear objective function with quadratic nonlinear constraints [5].

Cumulative portfolio risk, H. Markowitz divided into two parts. To the first part, he referred to the systematic risk, which is caused by the economic, psychological and political situation in the country, which simultaneously affects all assets equally. The second is the specific risk that each specific asset has, which can be eliminated by managing the assets portfolio.

The market portfolio is made up of individual stocks, so why doesn't its variability reflect the average variability of its components? The answer is that diversification reduces variability. Diversification works because prices of different stocks do not move exactly together. Statisticians make the same point when they say that stock price changes are less than perfectly correlated [6].

The risk that potentially can be eliminated by diversification is called specific risk. Specific risk stems from the fact that many of the perils that surround an individual company are peculiar to that company and perhaps its immediate competitors. The risk that we can't avoid, regardless of how much we diversify is generally known as the systematic risk or according to [6] known as market risk.

To measure the performance of a risky investment over a specific period the following formula is used:

$$R_{it} = \frac{A_t - A_{t-1}}{A_{t-1}}, \quad (1)$$

Where R_{it} is the simple return for i-th risky asset in a particular period t; A_t, A_{t-1} are current and previous risky asset price values, respectively.

The formula for the arithmetic mean R_i of returns R_{it} over a period T would be:

$$R_i = \frac{1}{T}\sum_{t=1}^{T} R_{it}, \quad (2)$$

where T is the total number of time periods over which the returns of the risky asset are considered; R_{it} represents the return of the i-th risky asset in the t-th time period.

The expected return of the risky asset portfolio R_p is defined as the weighted average return of the assets constituting the portfolio:

$$R_p = \sum_{i=1}^{n} w_i R_i, \quad (3)$$

w_i is the weight of the i-th risky asset in the portfolio; n is the number of risky assets that make the investment portfolio; R_i is the risky asset expected return.

The risk of one risky asset is:

$$\sigma_i = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(R_{it} - R_i)^2}. \quad (4)$$

The total portfolio risk is:

$$\sigma_p = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j p_{ij}\sigma_i\sigma_j} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \operatorname{cov}(R_i, R_j)}, \quad (5)$$

where σ_p is the investment portfolio risk level; w_i, w_j are the weights of the i-th and j-th risky assets in the portfolio; σ_i, σ_j are the standard deviations of returns for the i-th and j-th risky assets; p_{ij} is the coefficient of correlation between the returns of the i-th and j-th risky assets; cov (R_i, R_j) is the covariance between the returns of the i-th and j-th risky assets. The formula for covariance is given by:

$$\operatorname{cov}(R_i, R_j) = \frac{\sum_{t=1}^{T}(R_{it} - R_i)(R_{jt} - R_j)}{T}, \quad (6)$$

where T is the number of past observations.

The theory of H. Markowitz allows investors to measure the level of risk and determine effective portfolios.

The following represents the mathematical formulation of Harry Markowitz's portfolio optimization model, where the portfolio risk is constrained not to exceed a specified value σ_p:

$$\begin{cases} \sum_{i=1}^{n} w_i R_i \to \max \\ \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \operatorname{cov}(R_i, R_j)} \le \sigma_p \\ \sum_{i=1}^{n} w_i = 1 \\ w_i \ge 0, \forall i \in \{1, \dots n\} \end{cases}. \quad (7)$$

The following is the mathematical formulation of Harry Markowitz's inverse portfolio optimization problem, which seeks to minimize risk while ensuring the portfolio return is not less than a specified minimum R_p:

$$\begin{cases} \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \operatorname{cov}(R_i, R_j)} \to \min \\ \sum_{i=1}^{n} w_i R_i \ge R_p \\ \sum_{i=1}^{n} w_i = 1 \\ w_i \ge 0, \forall i \in \{1, \dots n\} \end{cases}. \quad (8)$$

The formula represents a mathematical optimization problem that is used to construct the Efficient Frontier in the context of portfolio selection.

In practice, both formulations can be used to trace out the entire Efficient Frontier by varying the parameters (σ_p in the first formula, R_p in the second) and solving the respective optimization problem for each parameter value. The points obtained through either method would be the same, representing the optimal combinations of risk and return that an investor can choose from based on their risk tolerance and return expectations.

James Tobin later extended Markowitz's work by incorporating a risk-free asset into the portfolio, leading to the separation theorem. This theorem implies that the optimal portfolio choice can be decomposed into two independent tasks: the selection of a portfolio of risky assets (as if the risk-free asset does not exist) and then deciding the mix between the risky portfolio and the risk-free asset. The integration of a risk-free asset simplifies the opportunity set, as indicated in the second text, by forming a straight line — known as the Capital Allocation Line (CAL) — in the risk-return space, which emanates from the risk-free rate and is tangent to the Efficient Frontier at one point.

William Sharpe furthered this line of inquiry by developing the Capital Asset Pricing Model (CAPM), which provides a mechanism to assess the expected return on an asset by relating it to its risk relative to the market as a whole, through the beta coefficient. The CAPM demonstrates that the expected return on a security is a function of the risk-free rate, the security's sensitivity to the market portfolio (beta), and the expected return of the market portfolio. This model underpins much of financial asset pricing theory and has profound implications for investment practice.

The contributions of these economists, rooted in the foundational work of Markowitz, effectively bridged the gap between theoretical finance and practical investment strategies. Markowitz's Efficient Frontier concept highlighted the trade-off between risk and return for multiple risky assets. Tobin's introduction of a risk-free asset into the portfolio mix, and the resulting CAL, simplified the selection process for investors. Sharpe's CAPM provided a pivotal relationship between an asset's risk, its expected return, and the overall market, helping investors make more informed decisions.

The Capital Allocation Line (CAL) is a concept from portfolio theory that describes the risk-return trade-off of portfolios that combine a risk-free asset and a risky asset. The CAL formula includes a component of risk (standard deviation or volatility) to account for the risk-return trade-off of combining a risk-free asset with a risky portfolio.

Calculating the capital allocation line is done as follows:

$$R_p = R_f + (R_i - R_f)\frac{\sigma_p}{\sigma_i}, \quad (9)$$

where R_p is the expected return of the portfolio on the CAL; R_f is the risk-free rate; R_i is the expected return of the risky portfolio (the portfolio of risky assets only); σ_p is the standard deviation (risk) of the portfolio on the CAL; σ_i is the standard deviation (risk) of the risky portfolio.

The term $(R_i - R_f)$ represents the risk premium of the risky portfolio over the risk-free rate. The ratio σ_p/σ_i indicates how much risk is being taken in the portfolio compared to the risk of the risky portfolio. The product of these two terms is then added to the risk-free rate to give the expected return of the portfolio on the CAL.

The slope of the CAL, which is the risk premium per unit of risk, can be expressed as:

$$\text{Slope of CAL} = (R_i - R_f)/\sigma_i. \quad (10)$$

The slope of the CAL represents the Sharpe ratio of the portfolio [7], indicating the additional amount of return an investor can expect for each additional unit of risk assumed by moving from the risk-free asset to a portfolio that includes a combination of risky assets. This measure provides a comparative basis for assessing the performance of the portfolio against the risk-free rate, taking into account the volatility of the portfolio's returns.

Risky asset is any asset that has a significant level of uncertainty in its future returns, such as stocks or index funds. These assets have the potential for higher returns compared to risk-free assets but come with the risk of losing value.

Risk-free asset is an asset that is assumed to have a certain return in the future, with no risk of financial loss. Typically, short-term government securities, like Treasury bills, are considered risk-free assets because they are backed by the full faith and credit of the issuing government.

When a risk-free asset with a portfolio of risky assets is combined, the CAL becomes tangential to the Efficient Frontier at the point representing the portfolio with the highest Sharpe ratio. This tangency portfolio is the optimal risky portfolio because it offers the best risk-return combination that can be achieved by investing in risky assets only.

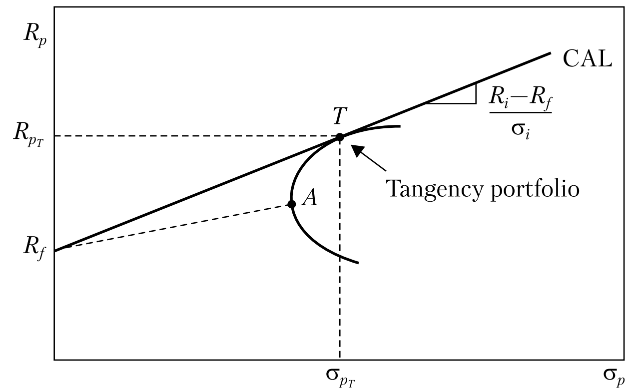

Fig. 1. Capital Allocation Line and the Efficient Frontier

The graph (Fig. 1) illustrates the concept that by combining a risk-free asset with the market portfolio [8], investors can achieve different levels of expected return for different levels of risk by choosing a point along the CAL.

Here's a breakdown of the figure 1:

- The vertical axis (R_p) represents the expected return of the portfolio.
- The horizontal axis (σ_p) represents the standard deviation of the portfolio's returns, which is a common measure of risk.
- The point R_f on the vertical axis indicates the risk-free rate, which is the return of an investment with no risk of financial loss.
- The curve represents the upper Efficient Frontier for all possible risky portfolios (without including the risk-free asset). This frontier shows the maximum expected return for a given level of risk for portfolios composed entirely of risky assets.
- The point labeled T is known as the Tangency Portfolio or the market portfolio. It's the portfolio on the Efficient Frontier with the highest Sharpe ratio, i.e., it offers the best possible risk-return combination.
- The solid line extending from R_f to T and beyond is the Capital Allocation Line. This line represents portfolios that optimally combine the market portfolio T with the risk-free asset R_f.
- Any portfolio on the CAL is considered to be optimally diversified because it combines the

market portfolio with the risk-free asset in some proportion.

- The slope of the CAL $(R_i - R_f) / \sigma_i$ represents the market price of risk or the “price of risk reduction” [9, p. 781], which is the additional expected return per unit of risk that an investor can obtain by moving from the risk-free asset to an investment in the market portfolio *T*.
- Point *A* on the curve of the Efficient Frontier represents a portfolio of risky assets that is not as optimally diversified as the market portfolio *T*.

Upon examination of the Capital Allocation Line, we observe that portfolios represented by points below the tangency portfolio *T*, such as point *A*, are suboptimal within the framework of the MPT. This suboptimality arises from the fact that these portfolios do not fully utilize the diversification benefits afforded by the market portfolio. Additionally, points extending beyond the tangency portfolio *T* on the CAL imply the use of leverage, which may involve short-selling risk-free assets to finance the purchase of additional units of the market portfolio. In practical investment scenarios where short-selling is either restricted or undesired, these leveraged portfolios are not feasible.

Therefore, to construct an effective model that aligns with these practical constraints, we propose restricting our analysis to the segment of the CAL that extends from the risk-free rate R_f to the tangency portfolio *T*, inclusive of these endpoints. Also, the part of the Efficient Frontier that is above the tangency point *T* represents portfolios that are composed of only risky assets and are highly efficient in terms of their risk-return profile included. The portion of the Efficient Frontier that lies above the tangency point is typically steeper, reflecting a higher reward for each incremental unit of risk. These practical constraints would lead to better performance even though Python is not the fastest language due to its high-level nature and interpreter-based execution.

The economic-mathematical portfolio model with maximizing return is as follows:

$$\begin{cases} w_0 R_f + \sum_{i=1}^{n} w_i R_i \to \max \\ \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \operatorname{cov}(R_i, R_j) \le \sigma_p} \\ w_0 + \sum_{i=1}^{n} w_i = 1 \\ w_i \ge 0, w_0 \ge 0, \forall i \in \{1, \dots n\} \end{cases}, \qquad (11)$$

where R_f is expected value of risk-free asset return; w_0 is the weight of a risk-free asset in a portfolio.

The computation of the mathematical model for the Capital Allocation Line and the upper segment of the Efficient Frontier by employing a mathematical model whose objective function seeks to minimize portfolio risk is done as follows:

$$\begin{cases} \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \operatorname{cov}(R_i, R_j) \to \min} \\ w_0 R_f + \sum_{i=1}^{n} w_i R \ge R_p \\ w_0 + \sum_{i=1}^{n} w_i = 1 \\ w_i \ge 0, w_0 \ge 0, \forall i \in \{1, \dots n\} \end{cases}. \qquad (12)$$

CAL and the part of the Efficient Frontier that is above the tangency point T can be found separately. If w_0 is allowed to vary without restriction but to be greater than 0, the solution to the optimization problem will give the CAL, which will include portfolios on the straight line from the risk-free rate up to the tangency portfolio. If leverage is not desired, w_0 must be constrained to be non-negative, which would exclude the leveraged portion of the CAL above the tangency portfolio.

To find the part of the Efficient Frontier that lies above the tangency point, one would simply set $w_0 = 0$ and resolve the optimization problem without the inclusion of the risk-free asset and set the expected return R_p to a value that is greater than the return of the tangency portfolio R_T. This would trace out the upper part of the Efficient Frontier, where investors are only taking positions in risky assets and not mixing in the risk-free asset, thus not utilizing the CAL.

It should be noted that the correctness of calculations is achieved only if the data follow a normal distribution law. This requirement is usually not met for asset prices, but it is met for return logarithms. But, for further calculations of geometric mean that will eventually be used for the calculation of the portfolio return simple returns are used because it's suitable for the purpose of calculating the geometric mean while log returns are not. This approach better reflects the compounded nature of investment returns over time.

The geometric mean, also known as the compounded annual growth rate (CAGR) for multiple periods, is often more appropriate than the arithmetic mean for datasets that involve compounding, such as cumulative growth rates, returns on investments over time, or other scenarios involving percentage changes. For projections over a year or similar periods, especially when using daily returns as the basis, the geometric mean is better suited than the arithmetic mean. It aligns with the principle that long-term investment returns should reflect the compounded growth rate, giving a realistic view of what an investor might expect their portfolio to achieve.

The formula for the geometric mean R_i of returns R_{it} over a period T would be:

$$R_i = \left(\prod_{t=1}^{T}(1+R_{it})\right)^{1/T} - 1. \qquad (13)$$

In the practical application of MPT, the choice between using log returns or simple returns for calculating the covariance matrix of risky asset returns depends on several factors, including the time horizon of the investment and the specific requirements of the analysis being conducted.

The formula for log return is:

$$R_{it} = \ln\left(\frac{A_t}{A_{t-1}}\right), \qquad (14)$$

where ln denotes the natural logarithm.

We're focusing on long-term planning and using daily returns as our data basis, using log returns to calculate the covariance matrix for assessing portfolio risk is a suitable approach since it can better handle volatility over long periods [10].

The risk-free rate is typically represented by the yield on government securities, such as U.S. Treasury bills. The rate obtained from API is quoted on an annual basis, so it is necessary to convert it to the same frequency as our risky asset returns i.e. daily.

To represent the process of converting an annualized risk-free rate to a daily rate suitable for comparison with daily risky asset returns in a mathematical formula, we can use the following approach. When converting an annualized risk-free rate to a daily rate, it's important not to use an arithmetic mean. Instead, the conversion should account for the compounding effect that occurs over the course of a year:

$$R_f = \left(\prod_{i=1}^{n}(1+R_{annual,i})^{\frac{1}{d_{year}}}\right)^{\frac{1}{n}} - 1, \qquad (15)$$

where $R_{annual,i}$ is the annual risk-free rate for day i; d_{year} is the number of trading days in the year, which is typically 252; n is the number of periods (days) over which the geometric mean is calculated.

While Treasury bills are typically quoted on a simple interest basis with a 360-day year for discount securities, when integrating such data into a portfolio optimization model that also includes assets, it is needed to synchronize the data frequency. Assets are traded on days when the markets are open, which is generally 252 days a year. Therefore, aligning the Treasury bill data to this same frequency by deannualizing to a daily rate over 252 trading days is a practical approach. It facilitates direct comparison and combination of risky asset and risk-free returns in the portfolio optimization process.

Also, after the formation of a portfolio with daily return and risk, we will convert them into yearly basis for better comparison between portfolios. The conversion process called annualization. To make annualization of the portfolio that consist the asset return and the risk-free return with different frequency is a mistake.

The annualized return is a form of geometric mean that accounts for the compounding over the period. The formula to convert daily portfolio return to an annualized return is:

$$\text{Annualized Return} = \left(\prod_{t=1}^{n}(1+R_p)\right)^{\frac{d}{year}} - 1,$$

where R_p are the daily portfolio return.

Using daily returns makes the model more sensitive to changes, and the annual format makes it suitable for comparing performance expectations across different time horizons and asset classes.

Annualizing risk, typically measured as the standard deviation of returns (volatility), involves scaling the risk measure from a shorter period (like daily or monthly) to an annual basis. This process is essential for comparing the risk levels of investments with different time frames or for presenting a standardized risk measure for portfolio performance projections. The formula to annualize volatility depends on the square root of time scaling rule, which is based on the assumption that returns are independently and identically distributed (i.i.d.) over time.

The formula to annualize risk is:

$$\text{Annualized Risk} = \sigma_p \sqrt{N}, \tag{17}$$

where N is the number of those periods in a year (252 for days, 12 for months, 52 for weeks, etc.).

In academic or professional papers, having a detailed algebraic representation of the methodology is necessary to convey the approach taken. Also, a precise formulation can be directly translated into code for computational purposes. It serves as a blueprint for the algorithms used in software implementations.

To represent the process for calculating the Capital Allocation Line and the upper part of the Efficient Frontier, we would extend the formula (12) to include the iterative process of varying the target return of the portfolio. Here is how this can be mathematically structured:

Given:

- Σ: Covariance matrix of risky asset returns;
- R: A column vector of expected returns for each asset;
- R^T: The transpose of R, which converts it from a column vector into a row vector;
- R_f: Risk-free rate;
- R_{p_i}: The target returns for the portfolio at the i^{tn} iteration;
- w: A column vector of asset weights;
- $w_{1:n}$: Denoting the subvector containing elements 1 through n, assuming n is the number of risky assets and the risk-free asset is the $(n+1)^{th}$ element;
- n: Number of assets;
- m: Number of portfolios to generate to form CAL and the upper Efficient Frontier.

We define the iterative process as follows:

1. Set the range for the target returns:

$$R_{\min} = R_f. \tag{18}$$

$$R_{\max} = \max(R). \tag{19}$$

$$step = \frac{R_{\min} - R_{\max}}{m}. \tag{20}$$

2. Iterate to find the optimal weights for each target return:

For each from 0 to m, do:

$$R_{p_i} = R_{\min} + i \cdot step. \tag{21}$$

3. Solve the optimization problem for each R_{p_i}:

$$\begin{cases} \sqrt{w_{1:n}^T \sum w_{1:n}} = \sigma_{p_i} \to \min \\ R^T w \ge R_{p_i} \\ 1^T w = 1 \\ w_i \ge 0, w_0 \ge 0, \forall i \in \{1, \dots n\} \end{cases}. \tag{22}$$

The result of each optimization will give us a point on CAL and the upper Efficient Frontier, characterized by the pair (R_{p_i}, σ_{p_i}), where σ_{p_i} is the minimum volatility corresponding to the return R_{p_i}.

The symbol T in the notation $w_{1:n}^T \Sigma w_{1:n}$ stands for the transpose of a vector or matrix. In the context of portfolio theory:

- w is a column vector representing the weights of different assets in the portfolio. If we have $n$assets, then w is an $n \times 1$ column vector.
- Σ is the covariance matrix of the asset returns, which is an $n \times n$ square matrix. Each element

of Σ represents the covariance relationships between the returns of the assets in the portfolio.

When we multiply $w_{1:n}^T \Sigma w_{1:n}$ we are effectively calculating the portfolio variance. Here's how it breaks down:

- $w_{1:n}^T \Sigma$ is the matrix multiplication of the transpose of the subvector $w_{1:n}^T$ with the covariance matrix Σ. This results in a $1 \times n$ vector where each element is the weighted sum of the covariances of one asset with all other assets in the portfolio.
- Finally, is a scalar value resulting from the dot product of the $1 \times n$ vector $w_{1:n}^T \Sigma$ with the $1 \times n$ vector $w_{1:n}$. This scalar is the portfolio variance, which represents the risk of the portfolio's returns.

The product $R^T w$ is the product of the vector of expected returns with the vector of weights, which results in the expected return of the portfolio. The inequality $R^T w \geq R_{p_i}$ ensures that the expected return of the portfolio is greater than or equal to a specified target return R_{p_i}.

In programming, specifically when using libraries such as NumPy in Python, we often don't explicitly write the transpose for a one-dimensional array since the dot product operation inherently handles the orientation of the vectors correctly. However, in more formal mathematical notation, especially when documenting a model or within academic papers, the transpose is used to clarify that we're performing a matrix multiplication that respects the dimensions of the vectors, i.e., a $1 \times n$ row vector multiplied by an $n \times 1$ column vector. This formal notation is important for mathematical correctness and to avoid ambiguity in the representation of vector and matrix operations.

The operation $1^T w$ is a matrix multiplication of a $1 \times n$ row vector with an $n \times 1$ column vector, which results in a scalar value. This scalar value is the sum of all the elements of the vector w. So, when we write $1^T w = 1$, we are stating that the sum of all the portfolio weights must equal 1, ensuring that the entire capital is fully allocated.

The constraint $w_i \geq 0$, $w_0 \geq 0$, $\forall_i \in \{1, ..., n\}$ is saying that every individual weight w of the portfolio must be non-negative, for each asset i from 1 to n and risk-free rate, ensuring that we're not shorting any assets and that the portfolio only consists of long positions.

In the context of the paper, the Sequential Least Squares Programming (SLSQP) method is employed to solve the portfolio optimization problem. It is an algorithm provided within the SciPy library's optimization module (scipy.optimize) for solving nonlinear optimization problems with both equality and inequality constraints.

The method iteratively solves a sequence of optimization subproblems. Each subproblem approximates the original nonlinear problem by a quadratic programming (QP) problem. The method uses the concept of a Lagrangian function, which incorporates both the objective function to be minimized (or maximized) and the constraints.

The iterative process looks like this:

1. Quadratic Approximation. At each iteration, SLSQP approximates the objective function by a quadratic function and the constraints by linear functions. This approximation is based on the current estimate of the solution and the derivatives (gradients) of the objective function and constraints with respect to the decision variables.

2. Solving Subproblems. The algorithm then solves a QP subproblem defined by the quadratic approximation of the objective function and the linear approximation of the constraints. The solution to this subproblem provides a direction along which the algorithm should move to improve the current estimate of the solution.

3. Line Search or Trust Region Step. After solving the quadratic subproblem, SLSQP either performs a line search or uses a trust region approach to determine the step size. This step is crucial for ensuring that the new estimate improves upon the current estimate in terms of the objective function value while satisfying the constraints.

4. Update and Repeat. The solution estimate is updated based on the results of the line search or trust region step. The algorithm then re-evaluates the objective function and constraints at the new point, updates the quadratic and linear approximations, and solves a new QP subproblem.

5. Convergence Check. After each iteration, SLSQP checks whether the solution has converged to an optimum. Convergence criteria typically involve the size of the step taken in the latest iteration, the change in the objective function value, and the satisfaction of the constraints. If the criteria are met, the algorithm stops; otherwise, it proceeds with another iteration.

SLSQP method is an optimization algorithm available within the SciPy library, a Python-based ecosystem of open-source software for mathematics, science, and engineering. Given its implementation in SciPy, researchers and practitioners have access to a robust and efficient tool for optimization tasks without the necessity to explore the underlying mathematical intricacies of the optimization process, such as the manual application of the method of Lagrange multipliers. The SLSQP method in SciPy abstracts this complexity, automating the solution of the constrained optimization problem and thereby streamlining the computational process.

The application of this methodology in a research paper involves using historical return data for the U.S. tech sector to populate the covariance matrix and the expected return vector, then applying these equations to calculate the optimal portfolio weights and the corresponding risk-return characteristics. This would allow for a detailed analysis of the investment opportunities within the U.S. tech sector under the mean-variance optimization framework.

Table 1. **Risk-Return Calculation Results of 11 Efficient Portfolios**

Portfolio	Daily		Yearly		
	σ, %	R_p, %	σ, %	R_p, %	Sharpe ratio
1	0	0.01	0	2.6	–
2	0.002	0.03	3.5	8.5	1.70
3	0.004	0.05	6.9	14.7	1.74
4	0.007	0.08	10.4	21.3	1.80
5	0.009	0.10	13.9	28.2	1.85
6	0.011	0.12	17.4	35.6	1.90
7	0.013	0.14	21.4	43.3	1.90
8	0.017	0.17	27.2	51.5	1.80
9	0.022	0.19	34.5	60.2	1.67
10	0.027	0.21	43.0	69.4	1.55
11	0.033	0.23	52.7	79.1	1.45

Utilizing Python, this study systematically retrieves and processes asset data, focusing on U.S. technology sector stocks from the NASDAQ 100 index and risk-free rate information, to construct efficient portfolios. By employing Python for data acquisition through Yahoo Finance API, we efficiently gathered daily closing prices and risk-free rate data for the period between March 1, 2021, and March 1, 2024, identifying 98 stocks with complete historical records.

The inclusion of major indexes — S&P 500, NASDAQ Composite, NASDAQ 100, Russell 2000, and Dow Jones Industrial Average — streamlines the analytical process by utilizing pre-calculated diversification, thereby serving as a foundational benchmark for asset allocation. As Myers, Allen, and Brealey articulate, astute investors maintain highly diversified portfolios, often beginning with the market itself as their foundational portfolio [6, p. 185]. This approach underscores the importance of incorporating a broad market index as a strategic basis for portfolio diversification.

The results of the risk-return analysis, summarized in Table 1, present the daily and annualized risk-return metrics for 11 efficient portfolios. These include the Sharpe ratios, illustrating the tradeoffs between risk and return at varying portfolio configurations.

We observe based on the Table 1 a systematic progression in both the expected daily return (R_p) and the associated daily risk (σ), from Portfolio 1 to Portfolio 11. To facilitate a meaningful comparison with established benchmarks, the daily figures have been annualized using standard formulas, a necessary conversion given that benchmark performances are typically reported on an annual basis.

The annualized data reveals that as we move from Portfolio 1 to Portfolio 11, there is an increase in both the expected annual return and annual risk, reflecting the fundamental trade-off at the heart of Modern Portfolio Theory — higher returns are accompanied by higher risk. Interestingly, the Sharpe ratio, which measures the risk-adjusted return, peaks at Portfolio 7 and then begins to decline. This suggests that Portfolio 7 represent the highest optimal balance between risk and return which also represent the tangency portfolio *T*.

When we juxtapose the results of Table 1 against the indexed performances in Table 2, we extend our analysis over two distinct time horizons — 3 years and 10 years. The 3-year and 10-year performances refer to the average annual returns and risks over those respective periods. Here, we consider the risk (σ), expected return (*R*), and the Sharpe ratio of major indexes (Table 2).

The 3-year performance of benchmarks shows they carry higher risk and, in some cases, lower returns than efficient portfolios. Particularly, the Russell 2000 index experienced negative returns, evident in its negative Sharpe ratio. Over 10 years, however, benchmarks show improved returns and Sharpe ratios, suggesting better long-term risk-adjusted performance. This comparison confirms the strength of the MVO-based portfolio construction in achieving favorable risk-adjusted returns against traditional market benchmarks presented in Fig. 2.

The CAL is constructed from the risk-free rate through Portfolio 7, extending upwards and serving as a benchmark for efficient investing. The upper Efficient Frontier is formed by the curve connecting the series of portfolio options above the tangency point, highlighting the potential for higher returns at increased levels of risk beyond the best trade-off.

Analyzing the weight distribution across the portfolios provides insight into the asset allocation strategy and how it aligns with the Capital Allocation Line and the upper Efficient Frontier (Table 3). If some stocks always have a weight of 0%, they are excluded from the table to save space.

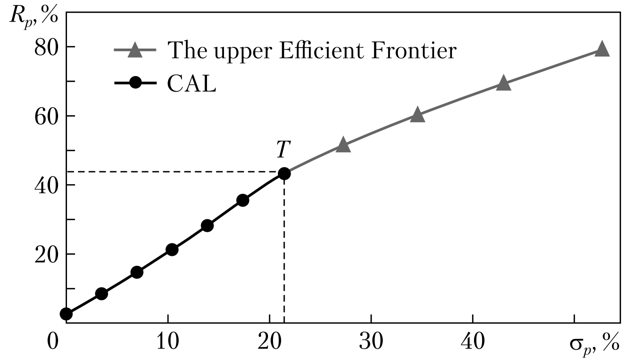

Fig. 2. The Capital Allocation Line and the upper Efficient Frontier of 11 efficient portfolios

Table 3 intriguingly shows the weight distribution across 11 efficient portfolios as percentages of the total portfolio, none of which incorporate the five indexes initially considered in the calculations. This exclusion underscores the indexes' risk-return profiles, which the algorithm deemed unsuitable for the portfolios it deemed most efficient. Despite analyzing 98 individual stocks, at most only nine are represented in any given efficient portfolio, emphasizing the selective nature of the optimization process and the critical role of asset performance and correlation in constructing an optimal portfolio.

The progression from Portfolio 1 through Portfolio 11 illustrates a dynamic allocation strategy, evolving from a risk-averse stance heavily weighted

Table 2. **Risk-Return Calculation Results of 5 Indexed over the Last 3 and 10 Years**

Index	3-year			10-year		
	σ, %	*R*, %	Sharpe ratio	σ, %	*R*, %	Sharpe ratio
S&P 500	17.4	9.3	0.39	17.8	10.7	0.53
NASDAQ Composite	23.3	5.8	0.14	21.1	14.2	0.61
Russell 2000	23.4	–3.3	–0.25	22.8	5.8	0.19
Dow Jones IA	14.8	7.3	0.32	17.5	9.2	0.45
NASDAQ 100	23.8	10.8	0.34	21.8	17.3	0.73

Table 3. **The Weight Distribution of 11 Efficient Portfolios, %**

№	AVGO	COST	FANG	NVDA	ORLY	PANW	PCAR	REGN	VRTX	R_f
1	–	–	–	–	–	–	–	–	–	100
2	0.6	2.4	2.5	1.8	6.6	0.2	0.7	2.7	1.7	80.7
3	1.1	4.3	4.9	3.8	13.9	0.1	1.5	5.7	3.2	61.5
4	0.9	8.0	7.6	5.8	20.0	0.2	3.3	8.0	4.0	42.3
5	1.3	11.5	9.4	7.9	26.1	0.7	4.1	10.2	5.6	23.2
6	2.3	14.9	11.5	8.4	30.9	0.6	5.8	14.5	9.9	1.3
7	–	10.9	15.3	22.7	37.9	–	–	10.2	3.0	–
8	–	1.0	16.7	39.3	41.3	–	–	1.7	–	–
9	–	–	18.0	58.0	24.0	–	–	–	–	–
10	–	–	18.6	77.2	4.3	–	–	–	–	–
11	–	–	–	100	–	–	–	–	–	–

in the risk-free asset (R_f) to a concentrated investment in a single equity (NVDA) by Portfolio 11.

Initially, Portfolio 2 displays a diversification among several equities with substantial allocation to the risk-free asset, suggesting a conservative investment approach. As we advance to Portfolio 7, we observe a significant shift towards equities like NVDA and ORLY, reflecting an increased risk appetite.

Strikingly, NVDA becomes the predominant holding in the latter portfolios, culminating in a single-asset strategy in Portfolio 11. This shift underscores the model's recognition of NVDA's performance within the analyzed period, positioning it as a high-impact asset within these optimized portfolios.

The table not only serves to exemplify the practical application of MVO in determining asset allocation but also showcases the range of investment strategies — from highly diversified to highly focused — that can be constructed under the MVO framework.

Modern Portfolio Theory (MPT), developed by Harry Markowitz in 1952, with further contributions by James Tobin and William Sharpe, provides a quantitative framework for assembling a diversified investment portfolio that maximizes expected return for a given level of risk. By considering the covariances between assets alongside their individual expected returns and variances, we have delineated the efficient frontier — showcasing portfolios that yield the highest expected return for a predefined level of risk.

The use of geometric mean in evaluating expected returns enhances the model's sensitivity to daily return fluctuations, thereby enriching long-term forecast accuracy and ensuring comparability across portfolios. This method effectively captures the compound nature of investment returns.

Through the incorporation of a risk-free asset, as posited by Tobin's Separation Theorem, we have simplified the construction of the efficient frontier and established the Capital Allocation Line and the upper Efficient Frontier, facilitating investor decisions on the optimal trade-off between risk and return.

Leveraging Python's computational capabilities, including optimization libraries like SciPy's SLSQP method, we have significantly streamlined the process of data retrieval, portfolio statistics calculation, and the critical iterative optimization process. This computational proficiency, in tandem with MVO's theoretical rigor, empowers investors to construct portfolios that are well-diversified and primed to potentially deliver superior risk-adjusted returns, particularly within the volatile U.S. technology sector.

In conclusion, the integration of computational advancements with MVO not only substantiates its application in current market conditions but also equips investors with the methodology to execute rapid, sophisticated portfolio optimizations — thus, systematically achieving efficient asset allocation.

REFERENCES

1. Contributors to Wikimedia projects. Modern portfolio theory. URL: https://en.wikipedia.org/wiki/Modern_portfolio_theory (Last accessed: 25.01.2024).
2. Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77. https://doi.org/10.2307/2975974
3. Tobin, J. (1958). Portfolio Selection. Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies*, 25(2), 65—86. URL: https://www.jstor.org/stable/2296205 (Last accessed: 23.02.2024).
4. Balvers, J. R. (2001). *Foundations of asset pricing*. Virginia.
5. Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1), 13. https://doi.org/10.2307/1924119
6. Myers, S., Allen, F., Brealey, R. (2011). *Principles of Corporate Finance*. McGraw-Hill/Irwin.
7. Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3), 425. https://doi.org/10.2307/2977928
8. Ahmadi, S. A., Peivandizadeh, A. (2022). Sustainable Portfolio Optimization Model Using PROMETHEE Ranking: A Case Study of Palm Oil Buyer Companies. *Discrete Dynamics in Nature and Society*, 2022, 1—11. https://doi.org/10.1155/2022/8935213
9. Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), 768. https://doi.org/10.2307/1910098
10. Bauer, G. H., Vorkink, K. (2011). Forecasting multivariate realized stock market volatility. *Journal of Econometrics*, 160(1), 93—101. https://doi.org/10.1016/j.jeconom.2010.03.021

Received 15.03.2024
Revised 28.03.2024
Accepted 29.03.2024

В.А. Шолопак (https://orcid.org/0000-0001-7492-3578),
Д.Д. Третяк (https://orcid.org/0000-0002-1610-6905)

Київський національний університет імені Тараса Шевченка,
вул. Володимирська, 64/13, Київ, 01601, Україна,
+380 44 239 3329, info@knu.ua

ОПТИМІЗАЦІЯ ЗА ДИСПЕРСІЄЮ ТА СЕРЕДНІМ ЗНАЧЕННЯМ: МОДЕЛЮВАННЯ ОПТИМАЛЬНОГО ІНВЕСТИЦІЙНОГО ПОРТФЕЛЯ В ТЕХНОЛОГІЧНОМУ СЕКТОРІ США

Вступ. Сучасна портфельна теорія (СПТ) забезпечує кількісну основу для прийняття обґрунтованих інвестиційних рішень. Нестабільний та непередбачуваний технологічний сектор США ставить під сумнів традиційні інвестиційні підходи, що робить необхідним дослідження методів, які краще враховують його унікальні співвідношення ризику та дохідності.

Проблематика. Традиційні інвестиційні стратегії часто не здатні адекватно врахувати динамічну природу технологічного ринку, оскільки вони покладаються на обмежену кількість даних і неефективні процеси розрахунків, що призводить до неоптимального розподілу активів. Одним із прогресивних методів вдосконалення стратегій формування портфеля, адаптованих до технологічного ринку, є метод оптимізації за дисперсією та середнім значення (MVO).

Мета. Оптимізація MVO для формування оптимальних портфелів у технологічному секторі США з використанням внесків СПТ, оптимізаційних технік Шарпа та моделі розподілу активів Тобіна.

Матеріали й методи. Використовуючи історичні дані про акції, MVO реалізовано за допомогою Python для формування портфелів, які включають безризиковий актив для розрахунку лінії розподілу капіталу (CAL) та верхньої ефективної межі. Геометричне середнє використано для оцінки очікуваної дохідності, що підвищує довгострокову прогнозованість і порівнянність портфелів, тоді як щоденні дохідності підвищують чутливість моделі.

Результати. Оптимізовані портфелі продемонстрували вищі коефіцієнти Шарпа та кращі характеристики ризику та дохідності, перевершуючи бенчмарки завдяки ефективним обчисленням.

Висновки. MVO є ефективним інструментом для інвестування в технологічному секторі, дозволяючи проводити інформований вибір активів і формування портфеля. Дослідження також підкреслює важливість інтеграції ітеративних процесів розрахунків та сучасних обчислювальних технік для адаптації традиційних інвестиційних стратегій до великих обсягів даних у сучасних ринкових умовах.

Ключові слова: оптимізація, лінія розподілу капіталу, ефективний кордон, програмування на Python.