Probabilistic Portfolio Optimization with CVaR



Table of Contents

Abstract	3
Introduction	4
Background Study	4
Literature Review	4
Model Implementation	5
Description of the Models	5
Handling Randomness	7
Related Applications	8
Numerical Experiments	8
Problem Description	8
Data Description	
Algorithms for Numerical Solution and Interpretation	12
Conclusion	18
Limitations	20
References	

Abstract

The objective of constructing an effective financial portfolio is to devise an investment strategy that maximizes returns while minimizing risk exposure. Traditional portfolio optimization methods often rely on deterministic models that require precise inputs for expected returns, risk assessments, and correlation matrices among assets. However, these deterministic approaches must adequately capture the uncertainties in financial data and market dynamics. Recognizing this limitation, our study introduces an innovative portfolio optimization framework incorporating probabilistic constraints to better manage uncertainty. By employing probabilistic modeling techniques, this approach integrates a range of potential outcomes into the decision-making process, thereby allowing for a more robust and resilient investment strategy.

This research contributes significantly to the field by showcasing how probabilistic bounds can be applied to investment decisions, utilizing mathematical models to quantify and mitigate financial risks more effectively. The methodology facilitates the accommodation of various risk preferences and constraints, enabling investors to define acceptable levels of risk through metrics such as Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). By prioritizing not only the optimization of returns but also the minimization of potential losses, our method aligns with the principle of creating a balanced and risk-aware investment portfolio.

Advanced computational methods, including stochastic programming, are employed to identify optimal investment strategies under uncertainty. This technique accounts for both the expected performance and the variability of returns across different asset portfolios. Empirical analysis,

using accurate financial market data, demonstrates the superiority of probabilistic constraint-based portfolio management over traditional optimization models, especially in terms of risk management and uncertainty mitigation. Our findings suggest that incorporating probabilistic considerations into portfolio optimization processes significantly enhances the resilience and performance of investment portfolios in the face of market volatility.

Keywords: Portfolio Optimization, Risk Management, Stochastic Programming, Financial Markets, Volatility, Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR).

Introduction

Background Study

It is important to have a diversity of finance investments. The aim is to create a portfolio with just the right amount of risk and potential for profit to match what investors need. Traditional methods of handling investments involve estimating how to maximize the project and how the value of investments can change. It emphasizes the shift from traditional to modern approaches that accommodate the unpredictable nature of financial markets. Predicting how investments will perform in finance is difficult because of uncertainties. This research aims to improve the field of investment portfolio management by proposing a probability method that can enhance decision-making processes. Making it more accessible to inform the decision for investors where to invest their money by considering how much they earn and how prices change.

Literature Review

The study about portfolio optimization looks at different ways to pick the best combination of investments. It all began with Harry Markowitz's development of the idea of mean-variance

optimization. Now, more methods like robust optimization and stochastic programming have been added. Traditional methods do not always think about possible issues. New research suggests making an investment less likely to succeed in portfolio planning models (Fox et al., 2019). Many studies have examined different ways to show uncertainty, like using different situations, making educated guesses, and using strong optimization methods. Moreover, Conditional Value at Risk (CVaR) extends its application beyond the original scope. It demonstrates the versatility of the CVaR approach functions, such as expected returns, by incorporating them within the constraints of CVaR. This structure allows for integrating multiple CVaR constraints reflecting various confidence levels, thus enabling loss distribution to align with decision-making risk preferences (Rockafellar & Uryasev, 2000). The "Risk Tomography" by (Prékopa, A., & Lee, J. 2018, pp. 161-165) introduces novel multivariate risk measures that assess risk along lines through a reference point in a multidimensional Euclidean space, using Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) as foundational elements. These measures take maximum or mixture approaches with respect to lines lying in cones, offering a nuanced perspective on risk that accommodates the multidimensional nature of financial portfolios.

Model Implementation

Description of the Models

When planning to improve a portfolio within certain limits, we will use a strong method to handle uncertain situations. In a market that is hard to predict, the model tries to make as much money as possible while avoiding Risk. Here are the main points:

Convexity means the problem will be easy to solve, and the solutions will work well and be reliable. The aim is to find the best way to divide investments between profit-making and reducing the chance of loss.

Objective Function

$$\operatorname{Max}\left(\sum_{i=1}^{n} \mu i * x i\right)$$

Where:

- xi represents the proportion of portfolio invested in asset i
- μi represents the expected return of asset i
- \blacksquare *n* is the number of assets in the portfolio.

Subject to:

The portfolio return should meet the probabilistic constraints for managing risk.

$$P(R < R_{min}) \le 1 - \alpha \tag{1}$$

Where:

- *R* represents the return on the portfolio
- \blacksquare R_{min} represents the minimum acceptable return, also known as the threshold return.
- α confident interval level, a number between 0 and 1 that represents the degree of certainty the investor requires.
- The probability $P(R \le R_{min})$ is the likelihood that the return on the portfolio will fall below this threshold return R_{min} .

For instance, if $\alpha = 0.05$, which mean 95% confident level that return will likely to fall below $R_{min.}$ If the portfolio return is 4% mean within the risk tolerance level. If not, then it is too low for return and risky. Whether to consider the risk is acceptable than the tolerance level or not. In conservative approach, it is the common risk management constraint in portfolio optimization to control potential risk.

The portfolio return that risk measure f(x,r) exceeds threshold σ is less than or equal to $1 - \alpha$

$$P(f(x,r) \ge \sigma) \le 1 - \alpha \tag{2}$$

Where:

- f(x,r) function represents the risk measure (e.gm Value at Risk(VaR)) or Conditional Value at Risk (CVaR) of the portfolio return. That can identify expected loss over a given period.
- σ threshold represents the maximum level of risk that willing to accept.

Handling Randomness

Handling randomness is important in the field of financial risk management, where future events are uncertainty. By using this approach, it will become more stable and best for future expectations.

$$P(f(x, \mu i) \ge \sigma) \le 1 - \alpha \tag{3}$$

Where:

- Replace r with μi , the mean return of asset i. which can vary significantly over time due to market volatility,
- Replace σ with historical volatilities or another estimate. Which is measure of risk or how much portfolio return has been varied.

Related Applications

The portfolio optimization model with probabilistic constraints is used in finance and investment management for different purposes. **Asset allocation**, its deciding where to invest your money in different investments to maximize your returns. **Risk management** is to find investments that will protect you from losing money. Choosing a collection of investments that balances how much risk they have and how much profit they can make depends on what different investors are comfortable with. Managing **hedge funds** involves overseeing the investment portfolios to reach the desired financial gain level while minimizing potential losses (Olabode, 2020).

Numerical Experiments

Problem Description

The problem at hand involves the optimization of portfolio allocation, with the goal of maximizing the expected return while maintaining a predefined level of risk tolerance. Specifically, the task is to determine the optimal distribution of investments across a set of assets to achieve the highest possible expected return while ensuring that the portfolio's Conditional Value at Risk (CVaR) remains below a specified threshold.

Data Description



The plot illustrates the adjusted closing prices (USD) for fifteen different assets from March 1, 2023, to February 28, 2024. The assets represented are from three sectors (i.e., technology, healthcare, and financial services). It is diverse performance across different sectors.

First to calculate the daily return, annualized returns and annualized volatility for a financial asset. The daily return measures the percentage change in the price of an asset from one trading day to the next. It is calculated by taking the difference between today's price and yesterday's price, and then dividing that by yesterday's price. The actual calculation is based on only trading days, which is approximately 252 days in a year).

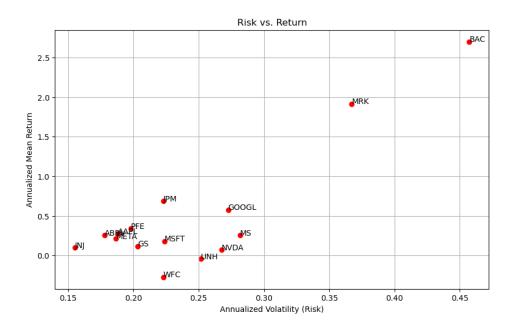
$$\label{eq:price_vesterday} \begin{aligned} \text{Daily Return} &= \frac{\text{Price}_{\text{today}} - \text{Price}_{\text{yesterday}}}{\text{Price}_{\text{yesterday}}} \end{aligned}$$

Calculating the Annualized Returned (with compound) is a way of expressing the return over a year by compounding the average daily return across the trading year (assuming there are 252 trading days in a year).

Annualized Return =
$$(1 + \text{mean daily return})^{252} - 1$$

Annualized volatility represents the standard deviation of the daily returns over a year. Volatility is a numerical measure of the scattering of returns and is often used as a measure of risk. To annualize the volatility, you multiply the standard deviation of the daily returns by the square root of the number of trading days in a year.

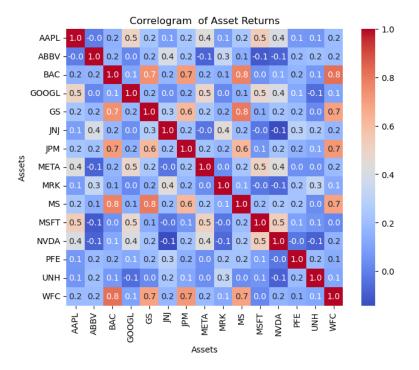




The plot represents the risk versus return portfolio of various assets. It's commonly referred to as a scatter plot in the context of financial analysis and is used to evaluate the trade-off between the expected return of an investment and its risk, as measured by volatility. It's useful to identify outliers on scatter plots, as they may indicate assets with an unusual risk-return profile. For instance, Bank of America (BAC) and Merck & Co., Inc. (MRK) could be considered outliers due to their significantly higher returns compared to the rest of the stocks, at their respective levels of risk. In financial terms, an annualized mean return of 1.0 implies a 100% return over the course of

a year. This means that the investment has doubled in value over the period being considered once the return is annualized. For example, if you had an investment that was worth \$1,000 at the beginning of the year, a 100% annualized return would suggest that it would be worth \$2,000 at the end of the year.

A correlogram of asset returns matrix is a table that shows the correlation coefficients between sets of variables. The cell in the matrix represents the correlation between two variables. Correlation coefficients quantify the intensity and the direction of a linear association between a pair of variables, ranging from -1 to 1.



A correlation coefficient of 0.8 indicates a strong positive relationship between the returns of Bank of America (BAC) and Morgan Stanley (MS). Both Bank of America and Morgan Stanley operate within the financial services sector, specifically in banking and investment services. This means similar economic factors, such as interest rate changes, economic policies, and credit markets likely influence them. It's likely to impact both stocks in a comparable way. If the goal is to

diversify risk, one might consider pairing these stocks with others from different sectors or with lower correlations. A diversification standpoint, JNJ (Johnson & Johnson) and NVDA (NVIDIA), lower correlation at (-0.1). When one might have challenges, the other may not be affected.

Algorithms for Numerical Solution and Interpretation

Algorithms Overview: Algorithms for numerical solutions play a pivotal role in addressing the complex and dynamic challenges of financial portfolio optimization, enabling the handling of intricate models, uncertainty, and risk that are often beyond the reach of analytical solutions. By employing a variety of optimization and simulation algorithms, such as financial practitioners can formulate and solve numerical problems that aim to maximize returns while adhering to constraints like budget and risk tolerance. This approach not only facilitates the precise calculation of critical risk measures like Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) but also accommodates the stochastic nature of financial markets, allowing for more informed and resilient investment strategies. Through the application of these numerical algorithms, the field of finance benefits from enhanced capability to predict outcomes, manage risks, and optimize portfolio performance in the face of uncertainty, demonstrating the algorithms' indispensable value in modern financial analysis and decision-making. Algorithms implementation of a numerical solution using an optimization convex library like CVXPY in Python.

Data Retrieval: Historical data for fifteen assets across three sectors were retrieved using Yahoo Finance for the period between March 1, 2023, and February 28, 2024 (Figure-1).

Data Processing: Adjusted closing prices were resampled to reflect only trading days, thus aligning with the financial calendar (Figure-1).

```
tickers = ['AAPL', 'META', 'NVDA', 'GOOGL', 'MSFT', 'JNJ', 'PFE', 'MRK', 'ABBV', 'UNH', 'JPM', 'BAC', 'WFC', 'GS', 'MS']

# define the date range
start_date = '2023-03-01'
end_date = '2024-02-28'

# Fetch data using yfinance
data = yf.download(tickers, start=start_date, end=end_date)

# Only business days are used for returns calculation
data = data['Adj Close'].asfreq('B').ffill()
```

Figure -1

Return Calculations: Daily returns were computed, providing the necessary granularity to understand daily price movements. Annualized returns were derived by compounding daily returns, giving a long-term perspective on investment performance and daily volatility, a proxy for risk, was calculated and then annualized by scaling with the square root of trading days (Figure-2).

```
# Calculate daily returns
daily_returns = data.pct_change().dropna()
# Calculate mean return (annualized)
mean_return_daily = daily_returns.mean()
# Calculate the compounded annual growth rate (CAGR)
mean_return_annualized = ((1 + mean_return_daily) ** 252) - 1
# Calculate volatility (standard deviation)
volatility_daily = daily_returns.std()
# Calculate annualized volatility
volatility_annualized = volatility_daily * (252 ** 0.5)
# output
results_df = pd.DataFrame({
   'Assets': tickers,
   'Mean Return': mean_return_annualized.values,
   'Volatility': volatility_annualized.values
})
```

Figure -2

Assets	Mean Returns (Annualized)	Volatility (Annualized)
AAPL	0.277863	0.188076
META	0.214598	0.186713
NVDA	0.071834	0.267482
GOOGL	0.576825	0.272594
MSFT	0.180757	0.223845
JNJ	0.098123	0.155098
PFE	0.339525	0.198084
MRK	1.913803	0.367173
ABBV	0.255995	0.178144
UNH	-0.043549	0.251709
JPM	0.686191	0.222917
BAC	2.700409	0.457161
WFC	-0.271734	0.222909
GS	0.116618	0.203231
MS	0.253930	0.281661

Figure -2

Covariance Calculations: The covariance matrix is a critical component in portfolio optimization as it quantifies the extent to which the returns of two assets move together. A positive covariance between two assets implies that they tend to move together. In contrast, a negative covariance indicates that they move inversely. A zero covariance suggests that the asset returns are uncorrelated (Figure-3).

$$\mathrm{Cov}(X,Y) = rac{\sum_{i=1}^{N}(X_i - ar{X})(Y_i - ar{Y})}{N-1}$$

- X_i and Y_i are the returns of asset X and asset Y at time i, respectively,
- X bar, Y bar are the average returns of asset X and asset Y over N observations, and
- *N* is the total number of observations.

```
###### Calculate the covariance matrix of daily returns
covariance_matrix = daily_returns.cov()
```

Figure -3

Where the covariance matrix C is:

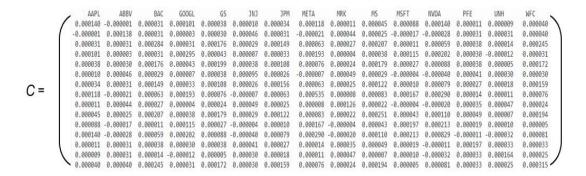


Figure -3 (covariances matrix of portfolio daily returns)

Model Optimization and Diversification: The optimization problem aimed to maximize the expected return, subject to several constraints: the sum of asset weights must equal 1, no individual weight can exceed 20%, and all weights must be non-negative.

Furthermore, the Conditional Value at Risk (CVaR) was constrained to remain below a threshold, ensuring that the portfolio adheres to a specified level of risk tolerance. The optimization was solved using the ECOS solver, a numerical optimization package suited for convex optimization problems. In this model, first to, initialize random seed and define the number of assets in the portfolio, set the maximum allowable weight for any single asset, and determine the risk tolerance level for the CVaR calculation. The mean returns and the covariance matrix are converted from their respective data structures into NumPy arrays, which are efficient for numerical computations (Figure-4).

```
# Initialization
np_random.seed (0)
num_assets = 15
max_weight = 0.20  # Maximum weight of 20% for any asset
alpha = 0.05  # Risk tolerance level

mean_returns = np.array( mean_return_annualized.values)
cov_matrix = np.array([covariance_matrix.values])
```

Figure-4

After initializing the variables, define a decision variable weight using the convex (cvxpy) library, which will be an array representing the *proportion of the portfolio to allocate to each asset*. This will be the main impact on output of the optimization.

```
# Define decision variables
weights = cp.Variable(num_assets)
```

Figure-5

In this portfolio optimization model, calculating the expected portfolio return by weighting the mean returns of each asset with the proportion of the portfolio's capital allocated to them. The portfolio's standard deviation, a proxy for risk, is computed using a quadratic form, which considers both the weights and the inter-asset covariance.

To manage extreme risk, Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) as optimization variables. VaR represents a threshold below which we expect the portfolio's losses to stay with a high confidence level, while CVaR provides the expected average loss in the worst-case scenarios exceeding VaR, thus capturing tail-end risk. These metrics are essential in ensuring our portfolio meets the defined risk tolerance level.

```
# Define the portfolio standard deviation (for risk)
portfolio_std_dev = cp.quad_form(weights, cov_matrix)
VaR = cp.Variable() # Create a variable for the VaR
portfolio_losses = -expected_return # Calculate the portfolio losses (negative returns)
cvar = cp.Variable() #average losses beyond the VaR
```

Figure-6

In the portfolio optimization model, a set of constraints to achieve a desirable result between risk and return. The total asset weights are constrained to one sum, which ensures the full investment of the portfolio's budget. Additionally, each asset's weight must be non-negative, no short selling, and no individual asset's weight can exceed a set maximum, preventing over-concentration and promoting diversification within the portfolio.

Also set constraints Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) constraints for risk management. The VaR constraint is formulated the assumption of normally distributed returns, where the portfolio's standard deviation, a measure of risk, cannot exceed the VaR threshold. This provides as a limit on the level of risk we are willing to accept.

Furthermore, the CVaR constraint seeks to ensure that the expected losses beyond the VaR do not exceed an acceptable level, as defined by confidence interval (alpha). This constraint limits the average of our potential extreme losses and keep portfolio risk in check. The mention of CVaR being less than or equal to zero is corrected; CVaR, in this context, is an average value and is restricted by the constraint to remain within our risk tolerance.

```
# Constraint for the VaR (assuming normal distribution of returns)
VaR_constraint = portfolio_std_dev <= VaR

# Constraint for the CVaR: the average value of the losses that exceed the VaR
cvar_constraint = cp_sum(cp_pos(portfolio_losses - VaR)) / (num_assets * alpha) <= cvar

# Define optimization constraints, including the CVaR constraint
constraints = [
    cp_sum(weights) == 1, # Budget constraint
    weights >= 0, # Non-negativity constraint
    weights <= max_weight, # Diversification constraint
    VaR_constraint, # VaR should not exceed the standard deviation by too much
    cvar_constraint # CVaR constraint</pre>
```

Figure-7

The objective function is to maximize the expected return of the portfolio. This is a straightforward goal in portfolio optimization to get the highest possible return for the risk taken.

```
objective = cp.Maximize(expected_return)
```

Figure-8

The optimization problem with the objective function and the constraints is then solved using the ECOS solver, which is suitable for convex optimization problems. Upon solving, the weights variable will be updated with the optimal asset weights that maximize the expected return while observing the defined constraints.

```
# Solve optimization problem using ECOS solver
problem = cp.Problem(objective, constraints)
problem.solve(solver=cp.ECOS)
```

Figure-9

Conclusion

After solving the optimization problem, the algorithm successfully determined the optimal asset to diversify (no individual weight can exceed 20%) and allocation in the portfolio. The results show the proportion of investments allocated among different assets to achieve maximum expected returns while respecting the identified risk tolerance. Additionally, the calculated CVaR provided valuable insight into the portfolio's potential downside risk beyond the specified confidence level.

Assets	Optimal Weight (annualized)	Mean Return (annualized)	Volatility (annualized)
AAPL	0.0000	27.79	18.81
META	0.0000	21.46	18.67
NVDA	0.0000	7.18	26.75
GOOGL	0.2000	57.68	27.26
MSFT	0.0000	18.08	22.38
JNJ	0.0000	9.81	15.51
PFE	0.2000	33.95	19.81
MRK	0.2000	191.38	36.72
ABBV	0.0000	25.60	17.81

UNH	-0.0000	-4.35	25.17
JPM	0.2000	68.62	22.29
BAC	0.2000	270.04	45.72
WFC	-0.0000	-27.17	22.29
GS	0.0000	11.66	20.32
MS	0.0000	25.39	28.17

Figure-10 (Asset diversification and allocation)

Figures 10 and 11 illustrate the final asset allocation. Certain assets such as GOOGL, PFE, MRK, JPM, and BAC are weighted at the maximum limit of 20%, indicating a strong model preference for these assets within the established risk parameters. This suggests that the model perceives these investments as significant contributors to portfolio performance while remaining within the acceptable volatility limits.

Other assets, such as AAPL, META, and NVDA, have been assigned zero weight, suggesting that within the context of this specific portfolio and its risk-return profile, they do not contribute optimally to the portfolio's objective under the model's constraints.

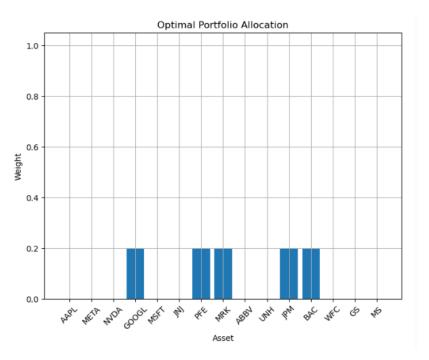


Figure-11(Asset diversification and allocation)

Limitations

- Normal Distribution: The normal distribution in market returns fails to capture the dynamic nature of actual returns, which may exhibit significant skewness that deviates from the symmetrical distribution model.
- Risk Aversion: The model presumes a static risk tolerance level, denoted as alpha, which does not adequately represent investors' evolving risk preferences, which may vary over time or in response to fluctuating market conditions.
- Numerical Solver: The application of different numerical solvers could result in disparate outcomes, a consequence of varying levels of solver reliability and the specific constraints of the Embedded Conic Solver (ECOS).
- **Single-Period Focus**: The current model framework overlooks the complexities of multiperiod investment strategies, including the necessity of periodic rebalancing, which is crucial for the sustained growth and risk management of long-term investment portfolios.

References

Olteanu, O., & Treanta, S. (2023). Symmetry in Mathematical Analysis and Functional Analysis (p. 190). MDPI-Multidisciplinary Digital Publishing Institute.

Olabode, A. S. (2020). Operations Research, Mathematics, Computer Science and Statistics: The Relationships. Data Science: Journal of Computing and Applied Informatics, 4(1), 59-70.

Fox, W. P., Burks, R., Fox, W. P., & Burks, R. (2019). Mathematical Modeling, Management Science, and Operations Research for Military Decision-Making. Applications of Operations Research and Management Science for Military Decision Making, 1-15.

Prékopa, A., & Lee, J. (2018). Risk tomography. European Journal of Operational Research, 265(1), 149-168. https://doi.org/10.1016/j.ejor.2017.07.055.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of Conditional Value-at-Risk. Journal of Risk, 2(3), 21-41.