Optimizing Investment Portfolios by NSGA-II

613k0007c Yu Pin Yi，613k0008c Wang Sin Hong

This experiment refers to the cited study and simplifies its methodology for practical implementation.

[2304.06675] A Learnheuristic Approach to A Constrained Multi-Objective Portfolio Optimisation Problem

I. Research Background and Objectives

A. Background:

1. Portfolio optimization is a core problem in the field of finance, aiming to maximize investment returns while minimizing risks.

B. Challenges:

1. Real-world non-convex problems require effective algorithms for solutions.
2. Multi-objective optimization involves considering multiple objectives within a portfolio, such as returns, risks, and liquidity.

C. Objectives:

1. Combine machine learning and multi-objective genetic algorithms (e.g., NSGA-II) to address multi-objective portfolio optimization problems.
2. Use surrogate models to accelerate algorithm convergence and improve solution quality.
3. Compare the optimization effectiveness of NSGA-II with the traditional Markowitz model.

II. Glossary of Terms

A. Markowitz Efficient Portfolio

1. Definition: The Markowitz Efficient Portfolio refers to the asset allocation that achieves the maximum return for a given level of risk or minimizes risk for a given level of return.
2. Application: It is used to guide how to allocate assets in a portfolio to achieve the optimal risk-return trade-off.

B. Expected Return

1. Definition: The expected return is the average return of an asset in the future, calculated based on past performance or market predictions. For each asset, the expected return is estimated using a weighted average of its historical returns.
2. Formula: Where E[Ri] is the expected return of asset i, and wis the weight of that asset in the portfolio.

$$E[R] = \sum_{i=1}^{n} wi\, E[Ri]$$

C. Risk

1. Definition: Risk is typically measured by volatility or standard deviation, representing the uncertainty or variability in an asset's returns. Higher risk indicates a broader range of potential returns.

D. Covariance Matrix

1. Definition: The covariance matrix is a data structure used to measure the return relationships between assets. Each element Cov(Ri,Rj) represents the covariance between asset i and asset j. Positive covariance indicates that the assets tend to move in the same direction, while negative covariance indicates they move in opposite directions. A covariance of zero indicates no relationship.
2. Application: In multi-asset portfolios, the covariance matrix helps calculate overall portfolio risk and is used in risk management and asset

E. Portfolio Volatility

1. Formula: Where wi is the weight of asset iii in the portfolio, σi is the volatility (standard deviation) of asset i, and Cov(Ri,Rj) is the covariance between assets i and j.

$$\sigma_\mathrm{p} = \sqrt{\sum_{i=1}^{N} wi^2 \sigma \mathrm{i}^2 + 2\sum_{i \neq j} wiwj\mathrm{Cov}(Ri, Rj)}$$

F. Sharpe Ratio

1. Definition: The Sharpe Ratio measures the ratio of portfolio return to risk and is used to compare the risk-adjusted returns of different assets or portfolios. A higher Sharpe Ratio indicates higher returns per unit of risk.
2. Formula: Where E[Rp] is the expected return of the portfolio, Rf is the risk-free return, and σp is the portfolio's standard deviation.

$$S = \frac{E[R_p] - Rf}{\sigma_\mathrm{p}}$$

G. Efficient Frontier

1. Definition: The efficient frontier is a curve consisting of optimal portfolios that provide the maximum return for a given level of risk or the minimum risk for a given level of return. These portfolios are the most efficient, meaning it is impossible to achieve higher returns without increasing risk or reduce risk without sacrificing returns.

III. Core Techniques and Methods

A. Genetic Algorithms

1. NSGA-II:A classic algorithm based on non-dominated sorting, specifically designed for solving multi-objective optimization problems. It generates a Pareto front, offering a set of solutions that represent the best trade-offs between conflicting objectives.

B. Non-Genetic Algorithm:

1. Efficient Frontier (Markowitz):Primarily used for solving single-objective optimization problems, focusing on optimal portfolio allocation based on Markowitz's mean-variance theory. It identifies portfolios that maximize return for a given risk or minimize risk for a given return.

C. Heuristic Algorithms

1. **Surrogate Models:**
    a. Reduce the number of expensive objective function evaluations.
    b. Gaussian Process Regression (GPR)**:** Commonly used to construct smooth surrogate models. GPR provides uncertainty estimates during the optimization process, making it highly effective in multi-objective optimization tasks.

D. Methodology

1. **Data Processing and Prediction:**
    a. Utilize stock data from **five companies** ("AAPL", "MSFT", "GOOGL", "AMZN", "TSLA") for the period **2013-01-01 to 2024-01-01** for testing.
    b. Split the data into **70% for optimization** and **30% for backtesting** to evaluate the performance of the optimization algorithms.

```
# **將數據分為 70% 用於優化，30% 用於回測**
train_size = int(len(prices) * 0.7)
train_prices = prices.iloc[:train_size]  # 用於優化的訓練數據
test_prices = prices.iloc[train_size:]  # 用於回測的測試數據
```

    c. Calculate the expected return and covariance matrix of the asset

```
# **步驟 2: 計算資產的預期收益與協方差矩陣**
returns = expected_returns.mean_historical_return(train_prices)
cov_matrix = risk_models.sample_cov(train_prices)
```

2. Objective Function and Parameter Design**:**
    a. **Objective:** Maximize the Sharpe ratio while minimizing portfolio volatility.

```
out["F"] = [portfolio_volatility, -sharpe_ratio]
```

3. **Genetic Algorithm Execution:**
    a. Randomly initialize a population of 200 individuals and perform selection, crossover, and mutation based on fitness values.
    b. Run for 100 generations and calculate the hypervolume and runtime for each algorithm.

```
# **步驟 5: 啟動優化問題並使用代理模型**
problem = PortfolioOptimizationProblem(returns=returns.values, cov_matrix=cov_matrix.values, surrogate_model=surrogate_model)

# **步驟 6: 定義 NSGA-II 多目標優化演算法**
algorithm = NSGA2(pop_size=200)
```

```python
# **步驟 7: 執行優化**
print("正在執行多目標優化...")
res = minimize(problem,
               algorithm,
               termination=('n_gen', 100),
               seed=42,
               save_history=True,
               verbose=True)
print("優化完成。")
```

4. **Surrogate Model Assistance:**

   a. Integrate surrogate models into the algorithm and adjust surrogate parameters to improve efficiency.

5. **Backtesting:**

   a. Perform simulated backtesting using the asset weights generated by the algorithms.

   b. Compare the optimization results of different algorithms.

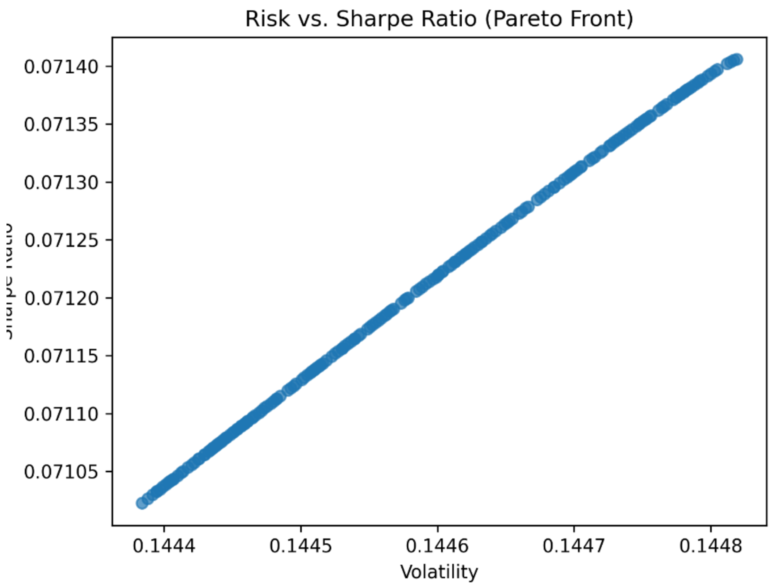
```python
# 計算每期的日回報
returns = prices.pct_change().dropna()  # pct_change() 計算百分比變化，dropna() 去除缺失值

# 投資組合的日回報：權重加權後的每日回報
portfolio_returns = (returns * weights).sum(axis=1)

# 計算累積回報（指數回報）
cumulative_returns = (1 + portfolio_returns).cumprod()
```

IV. Results and Evaluation

A. Efficient Frontier

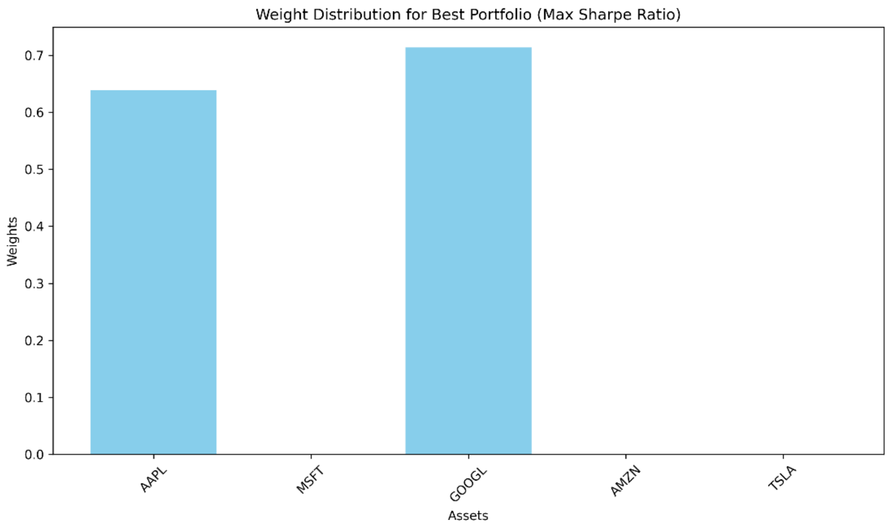


1. **Trade-off Between Risk and Return:**

   This chart illustrates how to select the optimal portfolio that provides the best return for a given risk level. As risk increases, return may also increase, but investors must identify the point that aligns with their risk preference.

2. **Optimal Solution Selection:**
   The curve represents the Pareto frontier, where solutions are optimal. These points signify the maximum Sharpe ratio achievable under different levels of risk, providing the best trade-offs between risk and return.
3. **Maximizing Sharpe Ratio:**
   For most risk-neutral or risk-averse investors, the portfolio with the highest Sharpe ratio on the Pareto frontier is the ideal choice, as it represents the best risk-adjusted return.

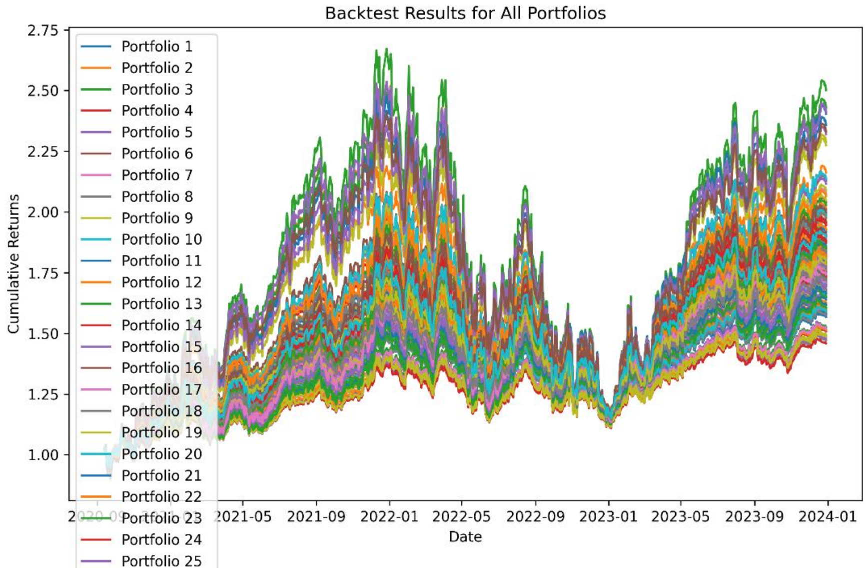
B. Optimized Portfolio Using NSGA-II with the Highest Sharpe Ratio



Portfolio	Volatility	Sharpe Ratio	AAPL	MSFT	GOOGL	AMZN	TSLA
Portfolio 2	0.144819	0.071405981	0.638786	3.90E-05	0.713327495	2.14E-06	4.44E-07

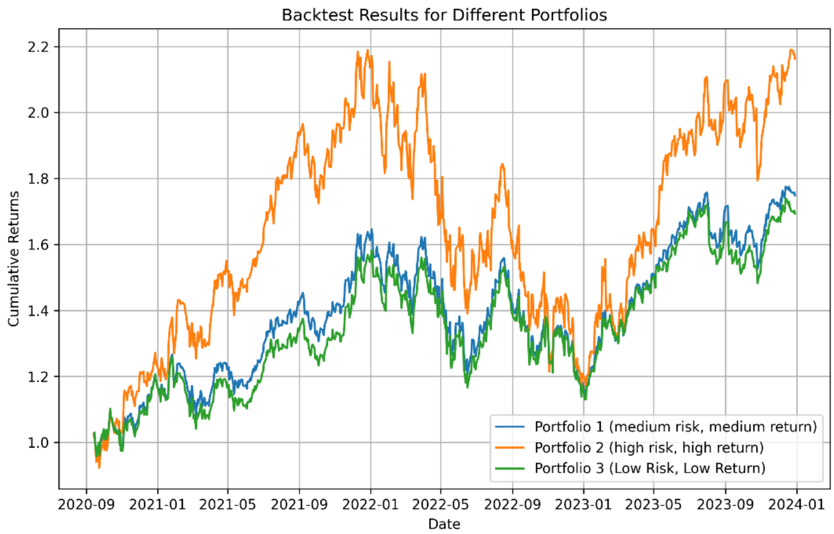
1. The portfolio with the highest Sharpe ratio represents the best risk-adjusted return, indicating that it achieves the highest profitability per unit of risk.

C. Backtesting All Portfolios Optimized Using NSGA-II



1. Using NSGA-II to optimize portfolios generates multiple optimal solutions representing different risk-return trade-offs. These solutions allow investors to select the best portfolio based on their specific risk tolerance and return expectations.

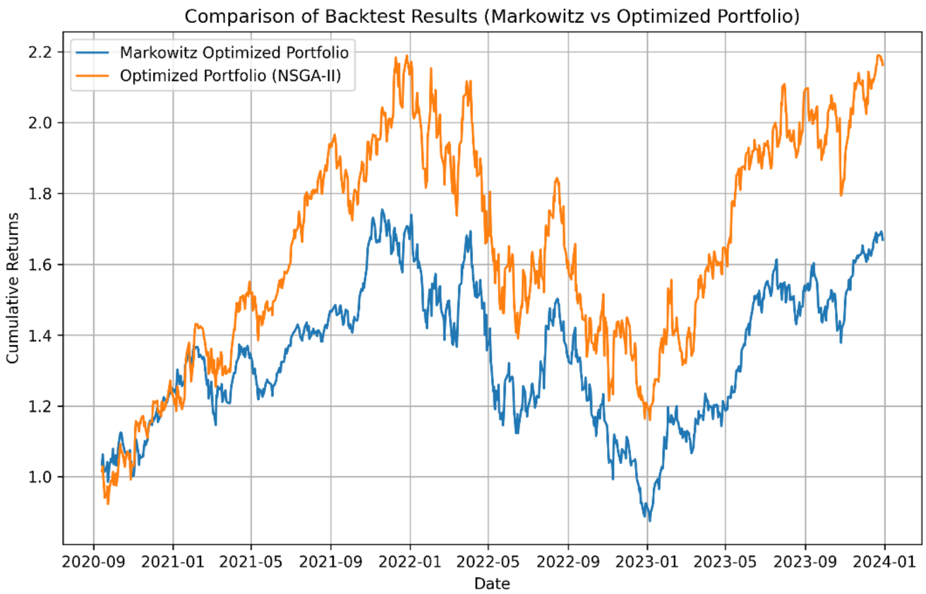
D. Backtesting Results for Three Strategies (High Return, Low Risk, Medium Risk and Return)



1. **Relationship Between Risk and Return:**
   Backtesting results show that the relationship between risk and return is not always straightforward. High-risk portfolios do not always yield the highest returns, while low-risk portfolios can sometimes provide better returns. This highlights the complexity of the market and challenges the traditional linear risk-return relationship.
2. **Stable Returns Through NSGA-II Optimization:**

Backtesting demonstrates that portfolios optimized using NSGA-II achieve positive returns across varying levels of risk during a ten-month backtesting period. This shows that NSGA-II effectively identifies stable solutions under diverse risk preferences. Despite fluctuations in returns, overall investment outcomes remain positive.

E. Comparison of NSGA-II and Markowitz Theory in Maximizing Sharpe Ratio



1. **NSGA-II Outperforms in Backtesting Results:**
   Portfolios optimized with NSGA-II for maximum Sharpe ratio consistently deliver more stable and superior returns. This shows that NSGA-II provides a better balance between risk and return, offering a flexible and efficient multi-objective optimization approach.
2. **Limitations of the Markowitz Method:**
   The Markowitz theory focuses on selecting the optimal asset allocation that maximizes return for a given risk level or minimizes risk for a given return. However, it lacks flexibility for adjusting the return-risk ratio based on individual investor preferences. This makes it less suitable for investors who want to tailor their portfolios according to their unique risk tolerance.
3. **Modern Algorithmic Needs:**
   As financial markets grow increasingly complex, modern investors need more than just a single "optimal" solution. Multi-objective optimization algorithms like NSGA-II or risk-adjusted strategies provide this flexibility, enabling investors to adjust their risk levels for a given return, better aligning with contemporary investment demands.

V. Conclusion

A. Outstanding Performance of NSGA-II in Maximizing Sharpe Ratio

Backtesting results show that the NSGA-II optimization method excels in multi-objective portfolio optimization. It not only maximizes the Sharpe ratio but also identifies the most suitable investment strategies tailored to varying risk preferences. Compared to traditional methods, NSGA-II delivers more stable and flexible returns.

B. Limitations of the Markowitz Method

While Markowitz theory effectively identifies single optimal asset allocations (maximizing returns and minimizing risk), it is limited in addressing the diverse risk tolerance of investors. This restricts its application for investors who want to adjust their return-risk ratios according to individual preferences.

C. Market Demands and the Advantages of Modern Algorithms

With increasing complexity in financial markets, investors' needs have evolved beyond single-objective optimization. Multi-objective optimization algorithms like NSGA-II offer flexibility in balancing risk and return based on investor preferences. These algorithms provide diverse solutions, aligning with the demands of modern investors.

D. Complexity of the Risk-Return Relationship

Backtesting data highlights that the relationship between risk and return is not linear. High risk does not always lead to high returns, and low risk can sometimes result in favorable returns. This further underscores the complexity of financial markets, necessitating multi-objective optimization algorithms to find balanced solutions across diverse risk preferences.

E. Stability and Practicality of NSGA-II Optimization

Over a ten-month backtesting period, portfolios optimized using NSGA-II consistently achieved positive returns, regardless of high or low-risk levels. The algorithm demonstrated stability across varying market conditions, proving its significance and practical value in modern financial investment.

VI. Future Directions

Future research can explore dynamic portfolio adjustment by integrating real-time market data with reinforcement learning and multi-objective optimization. This could enable the creation of adaptive investment strategies to respond to market volatility. Additionally, incorporating deep learning as surrogate models may enhance computational efficiency. Expanding the scope to include multi-asset classes such as bonds, commodities, and cryptocurrencies can validate performance in real-time trading. Moreover, introducing advanced risk metrics like CVaR (Conditional Value at Risk) or maximum drawdown optimization can address investors' demand for robustness in extreme market conditions, further advancing the practical applications of portfolio optimization techniques.