Optimizing Investment Portfolios by NSGA-II

(NSGA-II 在最佳化投資組合應用)

613k0007c 余品誼，613k0008c 王心泓

本實驗引用：對此論文實驗稍作簡化

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

1. 研究背景與目標

1.1 背景：

1. 投資組合優化是財務領域中的核心問題，目標是同時**最大化投資收益**與**最小化風險**。

1.2 挑戰：

1. 現實中的非凸（non-convex）問題需要有效的算法解決。
2. 多目標優化需要考慮投資組合中的多個目標（收益、風險、流動性）。

1.3 目標：

1. 結合機器學習與多目標基因演算法（如 NSGA-II），解決多目標投資組合優化問題。
2. 使用代理模型加速算法收斂並提升結果品質。
3. 比較不同 NSGA-II 與傳統 Markowitz 模型的優化效果

2. 名詞解釋：

2.1 Markowitz 最優組合（Markowitz Efficient Portfolio）

1. **定義**：Markowitz 最優組合是指在給定風險的情況下能夠實現最大回報的資產配置，或者在給定回報的情況下最小化風險的資產配置。
2. 應用：用於指導如何分配投資組合中的資產，以達到最佳的風險回報比。

2.2 預期回報（Expected Return）

1. **定義**：預期回報是根據過去的表現或市場預測來計算的資產未來的平均回報。對於每個資產，預期回報是根據其歷史回報的加權平均來估算的。
2. **公式**：E[Ri] 是第 i 支資產的預期回報，wi 是該資產在投資組合中的權

重

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

2.3 風險（Risk）

1. **定義**：風險通常用波動率或標準差來度量，指的是資產回報的不確定性或變異度。風險越高，表示回報的變動範圍越大。
2. **公式**：投資組合的風險（方差）由資產的**協方差矩陣**計算，並且受到資產之間相關性的影響。

2.4 協方差矩陣（Covariance Matrix）

1. **定義**：協方差矩陣是用來衡量資產之間回報關聯的數據結構。每個元素 Cov(Ri,Rj) 表示資產 i 和資產 j 之間的協方差。如果資產的回報高度相關，則協方差為正；若負相關，則協方差為負；無相關則協方差為零。
2. **應用**：在多資產組合中，協方差矩陣幫助計算整體組合的風險，並且用于風險管理和資產配置。

2.5 投資組合的波動率（Volatility）

1. **公式**：wi 是資產 i 在投資組合中的權重、σi 是資產 iii 的波動率（標準差）、Cov(Ri,Rj)是資產 i 和資產 j 之間的協方差。

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

2.6 夏普比率（Sharpe Ratio）

1. **定義**：夏普比率是衡量投資組合回報與風險的比率，用於比較不同資產或投資組合的風險調整後的回報。夏普比率越高，表示每單位風險所獲得的回報越高。
2. **公式**：E[Rp] 是投資組合的預期回報，Rf 是無風險資產的回報，σp 是投資組合的標準差。

$$S = \frac{E[R_p] - Rf}{\sigma_p}$$

2.7 有效前沿（Efficient Frontier）

1. **定義：**是一條由最優投資組合組成的曲線，這些投資組合在特定風險水平下提供最大回報，或在特定回報水平下提供最小風險。這些投資組合是最有效的，意味著無法在不增加風險的情況下獲得更高回報，或者在不降低回報的情況下減少風險。

3. 核心技術與方法

3.1 演算法

基因演算法 Genetic Algorithms, GAs）解決多目標優化問題：

1. **NSGA-II**：基於非支配排序的經典算法，用於多目標優化。

非基因演算法：

1. **EfficientFrontier(Markowitz)**：主要用於解決單目標優化問題，專注於基於 Markowitz 理論的最優投資組合配置。

特徵	**EfficientFrontier**(Markowitz)	**NSGA-II** (多目標進化演算法)
目標	單目標優化（回報最大化或風險最小化）	多目標優化（同時優化回報和風險）
優化方式	基於 **均值-方差理論**，使用數學優化方法	基於 **進化算法**，使用選擇、交叉、變異等操作
計算方法	**二次規劃** 或 **線性規劃**，數學優化方法	**遺傳算法**，探索多目標解空間
結果	單一最優解	Pareto 前沿解（多組最優解）
適用問題	投資組合優化，已知回報和風險之間的折衷	當需要處理多個目標，且這些目標之間存在折衷時
計算複雜度	計算較快，適用於較簡單的問題	計算較慢，需要更多的計算資源和時間
使用場景	回報和風險的單一目標優化	同時考慮多個目標，適用於多目標優化的情況

3.2 啟發式算法：

- **代理模型（Surrogate Models）：**
  - 減少昂貴的目標函數評估次數。
  - 高斯過程回歸（Gaussian Process Regression, GPR）：通常用來構建平滑的代理模型，它在優化過程中能夠提供不確定性估計，這使得它在多目標優化中表現良好。

3.3 方法流程

1. **數據處理與預測：**
   - 使用五檔日期從 2013-01-01 到 2024-01-01 股票(["AAPL", "MSFT", "GOOGL", "AMZN", "TSLA"]數據做測試
   - 把資料分成 7:3，70%做優化，30%做回測

```python
# **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
```

   - 計算資產的預期收益與協方差矩陣

```python
# **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. **目標函數與參數設計：**
   - 目標：最大化夏普比率，同時最小化波動性。

```python
out["F"] = [portfolio_volatility, -sharpe_ratio]  # 夏普比率作為最小化目標
```

3. **基因演算法運行：**
   - 隨機初始化種群 200，根據適應度值進行選擇、交叉與變異。
   - 運行 100 代後，計算各算法的超體積與運行時間。

```python
# **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. **代理模型輔助**：

   - 在演算法中嵌入代理模型，通過調整代理參數提升效率。

5. **回測（Backtesting）**：

   - 使用算法生成的資產權重進行模擬回測，比較不同演算法的優化效果。

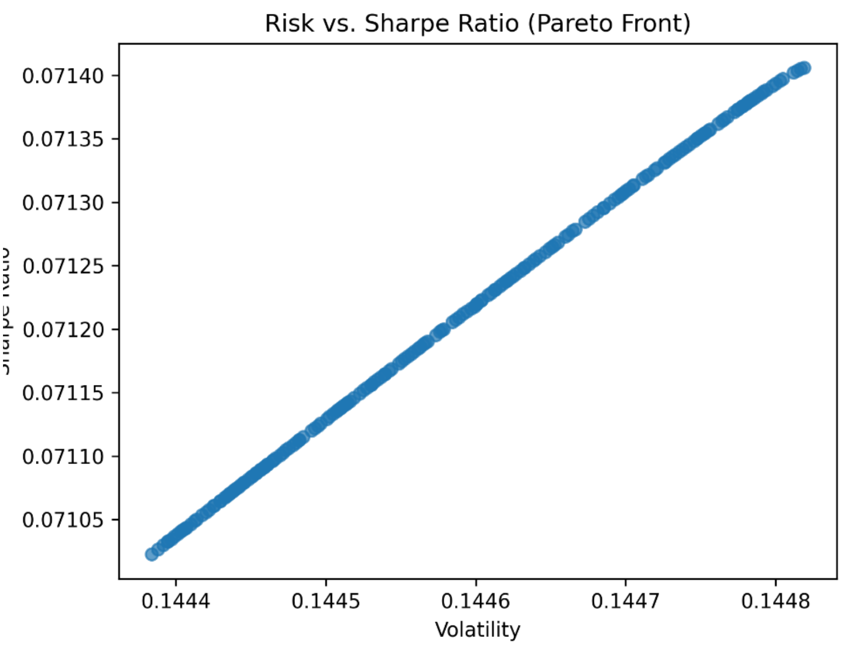
```python
# 計算每期的日回報
returns = prices.pct_change().dropna()  # pct_change() 計算百分比變化，dropna() 去除缺失值

# 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()
```

4. 實驗結果與評估

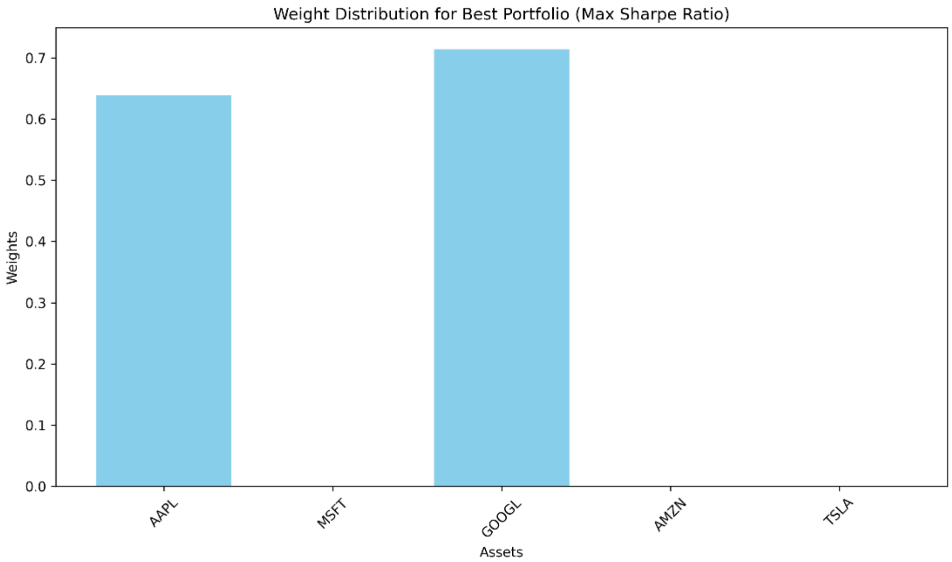
4.1 有效前沿（Efficient Frontier）



1. 風險與回報的折衷：這張圖反映了在給定風險水平下，如何選擇能夠提供最佳回報的投資組合。隨著風險增加，回報也可能增加，但投資者需要找到最適合自己風險偏好的點。
2. 最優解選擇：這條曲線代表了 Pareto 前沿。在這條曲線上的解是最優的，代表了在不同風險下，能夠實現的最大夏普比率。這些解是風險與回報之間的最佳折衷。
3. 夏普比率最大化：對於大多數風險中立或風險厭惡的投資者，他們會選擇 Pareto 前沿上的夏普比率最高的組合，這代表了最佳的風險調整回報。

4.2 使用 NSGA-II 優化投

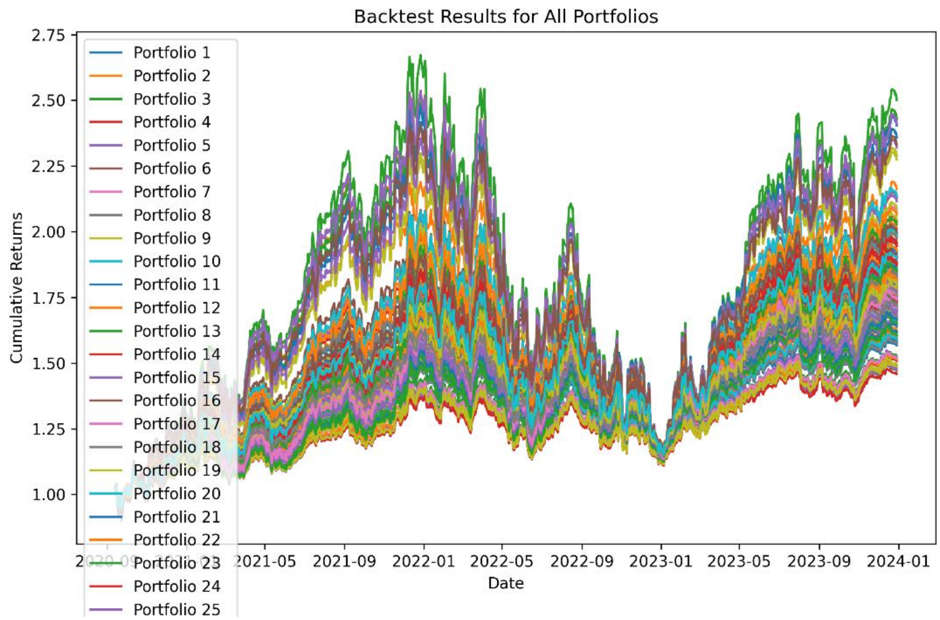
資組合後取夏普比率最高的投資組合



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

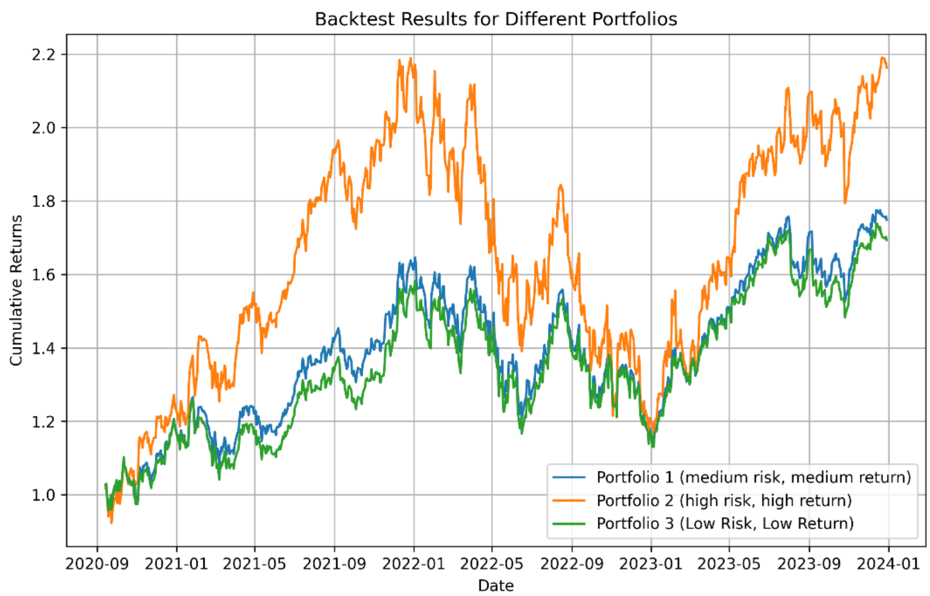
蒐集 Sharpe 比率最高的投資組合，代表每單位風險獲利越大。

4.3 使用 NSGA-II 優化投資組合後回測所有組合



使用 NSGA-II 優化投資組合可以得到多個風險與報酬比例的最佳解。

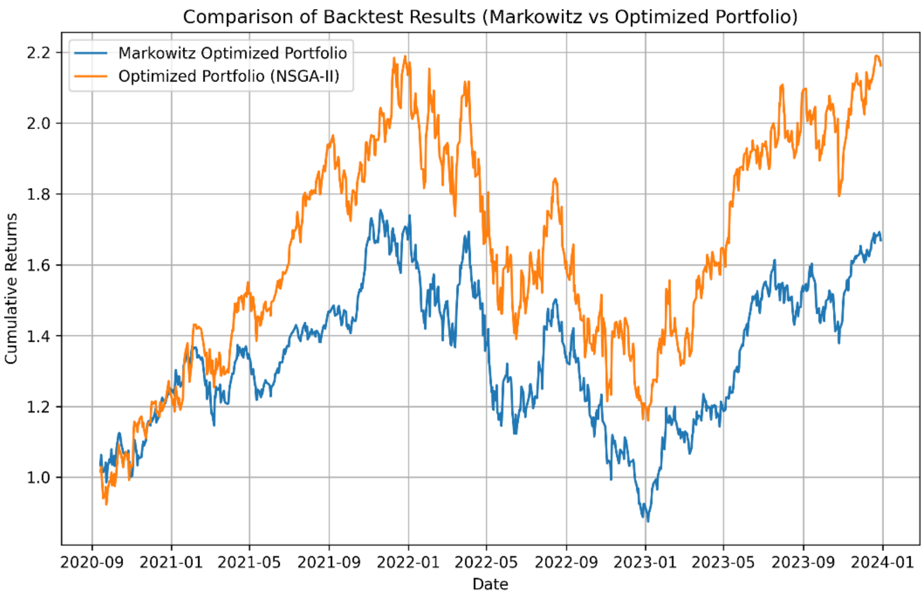
4.4 使用 NSGA-II 優化投資組合後制定策略三種策略的回測成果(高回報，低風險，中度風險中度回報)



1. 回測結果顯示風險與回報的關係：這張圖強調了風險與回報並不總是簡單對應的，這表明即便是高風險的投資組合，有時回報未必會非常高，同樣低風險的投資組合也可能提供較好的回報。這反映了市場中的複雜性，並挑戰了傳統的風險-回報線性關係。

2. NSGA-II 優化的穩定回報：使用 NSGA-II 進行優化，結果表明即使是不同風險程度的投資組合，在十個月的回測期內，大多數情況下仍能實現正回報，證明這種多目標優化算法能夠有效地找到在多樣化風險偏好下的穩定解，即使回報波動較大，最終的投資收益也是積極的。

4.4 NSGA-II 與 Markowitz theory 演算法最大化夏普比率回測比較



1. 最大化夏普比率的 NSGA-II 投資組合回測結果較好：這表明，從回測結果來看，NSGA-II 優化方法（透過最大化夏普比率）通常能提供更穩定且優秀的回報，滿足投資者對於風險與回報之間平衡的需求，是一種更為靈活且高效的多目標優化方法。
2. Markowitz 方法的局限性：Markowitz 理論強調的是**最優資產配置**，即選擇能夠提供最大化回報且風險最小的投資組合。然而，它並沒有提供針對不同投資者風險偏好的靈活調整方式。因此，對於那些希望根據個人風險承受能力來調整回報和風險比例的投資者，這種方法顯得不夠靈活。
3. 現代需求的算法：隨著金融市場的複雜性提升，現代投資者的需求不僅是尋找「最優解」，還需要能夠根據自身的風險偏好來選擇合適的投資策略。多目標優化算法（例如 NSGA-II）或**基於風險調整的策略**可以提供這種靈活性，使得投資者能夠在給定回報的情況下，調整承擔的風險水平。

5. 結論

5.1 最大化夏普比率的 NSGA-II 投資組合表現突出

回測結果顯示，使用 NSGA-II 優化方法進行多目標投資組合配置，不僅能實現夏普比率的最大化，還能根據不同風險偏好找到最適合的投資策略。相較於傳統方法，NSGA-II 能夠提供更穩定且靈活的回報。

5.2 Markowitz 方法的局限性

雖然 Markowitz 理論能有效找到單一最優資產配置（最大化回報且風險最小），但其局限在於無法滿足投資者對個人風險承受能力的差異化需求。因此，對於希望根據自身偏好調整回報與風險比例的投資者來說，其應用範圍受到限制。

5.3 市場需求與現代算法的優勢

金融市場日益複雜，投資者需求也不再侷限於單一目標的優化。NSGA-II 等多目標優化算法能夠根據投資者的偏好，靈活調整風險與回報，提供更多元的解決方案，符合現代投資者的需求。

5.4 風險與回報的關係複雜性

回測數據強調，風險與回報並非簡單的線性關係。高風險並不總是意味著高回報，低風險也可能帶來良好的收益。這再次證明，市場的複雜性需要多目標優化算法來應對，以在多樣化的風險偏好中找到平衡。

5.5 NSGA-II 優化的穩定性與實用性

十個月的回測結果顯示，不論是高風險還是低風險的投資組合，NSGA-II 算法都能實現正回報，且在不同市場情況下表現出穩定性。這證明了多目標優化算法在現代金融投資中的重要性和實用價值。

6. 延伸

未來可以探索如何動態調整投資組合，結合市場即時數據應用強化學習與多目標優化，設計出能自適應市場波動的投資策略。此外，引入深度學習作為代理模型，提升計算效率，並將研究範圍擴展至多資產類別（如債券、商品、加密貨幣等），驗證其在實時交易中的表現。同時，可以引入更複雜的風險指標（如 CVaR 或最大回撤）進行優化，滿足投資者在極端市場中的穩健性需求，進一步推動投資組合優化技術的實際應用。