Quantitative Portfolio Optimization

Team NAPPERS

Summary

Goal: Develop an end-to-end portfolio optimization pipeline that constructs monthly, long-only equity portfolios on an S&P 500–like universe and evaluates them against simple benchmarks using strictly out-of-sample performance.

Current Project Scope:

- Focuses on monthly, long-only, fully-invested equity portfolios; excludes leverage, shorting, options, and intraday trading.
- Restricts evaluation to historical backtests on the `sp500_monthly` sample
- Includes infrastructure, testing, and documentation as primary goals so that future extensions (e.g. new signals, risk models, or universes) can plug into the same framework.

Objective

$$\max_{w_t} \ \mu_t^\top w_t - \gamma\, w_t^\top \Sigma_t w_t - \lambda \|w_t\|_1 - \kappa \|w_t - w_{t-1}\|_1, \quad \text{s.t.} \ \sum_i w_{t,i} = 1, \ w_{t,i} \geq 0,$$

This objective balances four components:

Expected return: rewards allocations toward assets with higher estimated mean returns (estimated from rolling historical data).

Risk penalty: penalizes portfolio variance, with γ\gammaγ controlling risk aversion.

Weight regularization: discourages extreme or highly concentrated allocations by penalizing large absolute weights.

Turnover penalty: penalizes deviations from last month’s portfolio, serving as a proxy for transaction costs and encouraging stability over time.

Data

Sourced from Wharton Research Data Services (WRDS)

S&P 500 monthly data (2000 - present)

Useful fields include: date, ticker, price, return, shares outstanding, SIC codes

<u>Pros:</u>

- Dataset is cleaned and reputable
- Long historical coverage (25+ years of monthly data
- Breadth of data (returns are pre-calculated, outstanding shares and volume included)

Cons:

- Limited in scope beyond straightforward applications of portfolio optimization
- Huge dataset - computationally expensive, could lead to overfitting
- Access limitation - potentially leads to complications with theoretical distribution/real-world implementation of model
-

Moving forward:

- As project advances, we may need to acquire supplementary data - finer or coarser grain date-wise depending on eta
- In order to differentiate from previous project, may need to search for more niche dataset to narrow our application (ESG, alternate optimization strategy such as a two layer model or higher moments)

Preliminary Results

Optimized Portfolio

- Mean monthly return: 0.0119
- Annualized volatility: 0.1685
- Annualized Sharpe: 0.8451
- Max drawdown: -0.4620
- Cumulative return: 17.5990

Equal-Weight Benchmark

- Mean monthly return: 0.0119
- Annualized volatility: 0.1688
- Annualized Sharpe: 0.8459
- Max drawdown: -0.4622
- Cumulative return: 17.7497

Key result

Optimized portfolio $\approx$ equal-weight across return, risk, and Sharpe.

Preliminary Results

Distance from equal-weight

$$\mathbb{E}_t\left[\|w_t - w_t^{\mathrm{EW}}\|_1\right] \approx 0.0072$$

Turnover

$$\mathbb{E}_t\left[\|w_t - w_{t-1}\|_1\right] \approx 0.00023$$

Implications

- Weights remain **extremely close** to equal-weight at every date.
- Turnover is **near zero**: the strategy behaves like buy-and-hold with tiny nudges.
- With rolling-mean expected returns and sample covariance plus a turnover penalty, there is little incentive to move away from equal-weight.

Preliminary Results

Why does the optimizer collapse to equal-weight?

- Rolling mean estimates of expected returns are **noisy** (weak signal).
- Sample covariance can be **unstable** in high dimensions.
- The turnover penalty discourages reallocations unless the signal is strong.

Economic intuition

When signals are weak, diversification dominates, and equal-weight is hard to beat. With weak signals and a turnover penalty, the optimizer learns that equal-weight is already near-optimal.

Exploration: Skew Tensor

Let $X \in \mathbb{R}^n$ be a random vector with mean μ.
Define the third-order central moment tensor (coskewness tensor):

$$\mathcal{S}_{ijk} = \mathbb{E}\left[(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)\right]$$

$$\mathcal{S}_{ijk} = \mathcal{S}_{jik} = \mathcal{S}_{ikj} = \mathcal{S}_{kij} = \mathcal{S}_{kji} = \mathcal{S}_{jik}.$$

Case	Indices	Representative Entry	Number of Unique Terms	Growth
Single	$i = j = k$	$\mathcal{S}_{iii} = \mathbb{E}[(X_i - \mu_i)^3]$	n	$\mathcal{O}(n)$
Double	$i = j \neq k$	$\mathcal{S}_{iik} = \mathbb{E}[(X_i - \mu_i)^2(X_k - \mu_k)]$	$n(n-1)$	$\mathcal{O}(n^2)$
Triple	$i \neq j \neq k$	$\mathcal{S}_{ijk} = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)]$	$\binom{n}{3} = \frac{n(n-1)(n-2)}{6}$	$\mathcal{O}(n^3)$
Total	—	—	$\frac{n(n+1)(n+2)}{6}$	$\mathcal{O}(n^3)$

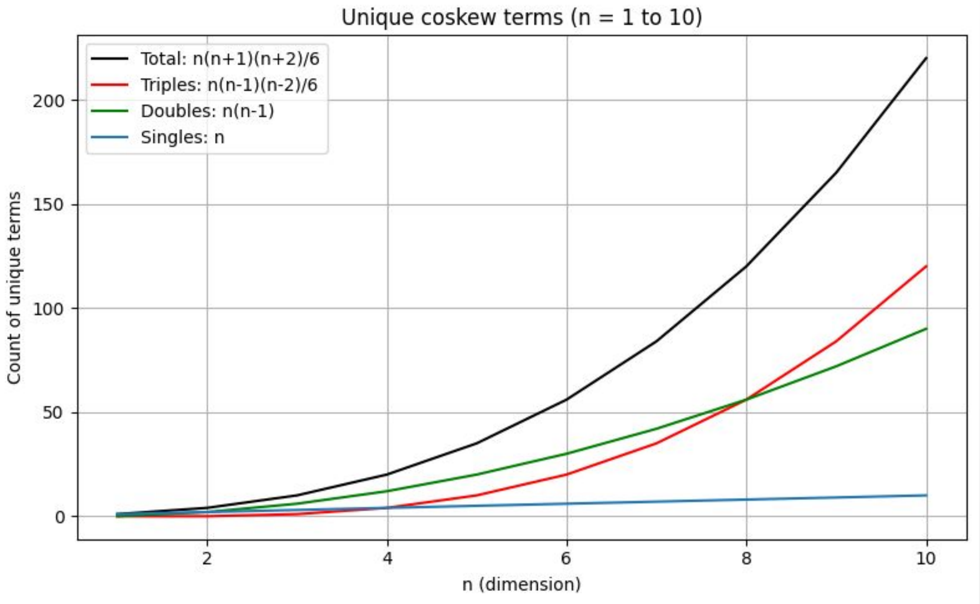
Exploration: Skew Tensor 2

As dimensionality increases, the structure of coskewness becomes overwhelmingly dominated by

triple-variable interactions.

This has implications for:

- Storage complexity,
- Estimation stability,
- Regularization needs,
- Dimensionality reduction strategies.



Exploration: Skew Tensor 3

For portfolio optimization we want to maximize the portfolio skew at the cost of portfolio variance, where portfolio skew is defined by the contraction of the tensor over the weight indices.

$$\boxed{\text{Skew}_p = \sum_{i,j,k} w_i w_j w_k \mathcal{S}_{ijk}}$$

$$\text{Skew}_p = \mathcal{S}[w, w, w].$$

$$\text{Skew}_p = w^\top \mathcal{S}(w, w),$$

Formulation	**Objective**	**First-Order Condition (FOC)**
Skew–Variance Tradeoff	$\max_w \mathcal{S}[w,w,w] - \lambda w^\top \Sigma w \quad \text{s.t. } \mathbf{1}^\top w = 1$	$3\,\mathcal{S}(w,w) - 2\lambda\Sigma w - \gamma\mathbf{1} = 0$
Unconstrained Skew–Variance	$\max_w \mathcal{S}[w,w,w] - \lambda w^\top \Sigma w$	$3\,\mathcal{S}(w,w) = 2\lambda\Sigma w$
Scale-Invariant (Skew Ratio)	$\max_w \dfrac{\mathcal{S}[w,w,w]}{(w^\top \Sigma w)^{3/2}}$	$\mathcal{S}(w,w) = \theta\,\Sigma w$
Mean–Variance–Skew	$\max_w \mu^\top w - \dfrac{\lambda}{2} w^\top \Sigma w + \eta\mathcal{S}[w,w,w] \quad \text{s.t. } \mathbf{1}^\top w = 1$	$\mu - \lambda\Sigma w + 3\eta\,\mathcal{S}(w,w) - \gamma\mathbf{1} = 0$