

Large–Deviation–Based Portfolio Optimization

Ryan Dai

Department of Quantitative Finance

2025.06.09

Content

1. Introduction
2. Market Assumption and Portfolio Setup
3. Cramér's Theorem and the Large-Deviation Objective
4. Empirical Results
5. Future Research direction
6. Reference

Introduction: Portfolio optimization

- Mean-Variance portfolio optimization: Find a weight to
 - Maximize expected return
 - Minimize variance
- Large Deviation method:
 - Find a weight to maximize the probability that average log-return of a portfolio exceeds a target c .

Market Assumption and Portfolio Setup

Definition 2.1 (Asset Prices and Returns). Let $\mathbf{S}_t = (S_{1,t}, \dots, S_{n,t})$ denote the vector of asset prices at day $t \in \{0, 1, \dots, T\}$. The one-day *simple return* vector is

$$\mathbf{r}_t = (r_{1,t}, \dots, r_{n,t}), \quad r_{i,t} := \frac{S_{i,t+1}}{S_{i,t}} - 1.$$

We focus on log-returns, which are additive over time and better suited for long-horizon analysis.

Definition 2.2 (Portfolio Weights and Log-Returns). A static allocation is represented by a weight vector

$$\mathbf{w} = (w_1, \dots, w_n) \in \mathcal{W} \subseteq \mathbb{R}^n, \quad \mathcal{W} := \left\{ \mathbf{w} \mid \sum_i w_i = 1, 0 \leq w_i \leq 0.1 \right\}.$$

The portfolio value at time t , given weights \mathbf{w} , is

$$P_t(\mathbf{w}) := \mathbf{w}^\top \mathbf{S}_t,$$

and the one-day *log-return* of the portfolio is

$$L_t(\mathbf{w}) := \log \left(\frac{P_{t+1}(\mathbf{w})}{P_t(\mathbf{w})} \right) = \log (1 + \mathbf{w}^\top \mathbf{r}_t).$$

Market Assumption and Portfolio Setup

Definition 2.3 (Cumulative and Average Log-Return). Over the full horizon $[0, T]$, the portfolio's cumulative and average log-returns under weight \boldsymbol{w} are

$$\log \left(\frac{P_T(\boldsymbol{w})}{P_0(\boldsymbol{w})} \right) = \sum_{t=0}^{T-1} L_t(\boldsymbol{w}) \quad (\text{cumulative log-return}),$$

$$\bar{L}_T(\boldsymbol{w}) := \frac{1}{T} \sum_{t=0}^{T-1} L_t(\boldsymbol{w}) \quad (\text{average log-return}).$$

Theorem 2.1 (Log-return distribution under multivariate GBM). Assume the asset price vector $\boldsymbol{S}(t) = (S_1(t), \dots, S_n(t))^{\top}$ follows a multivariate geometric Brownian motion

$$d\boldsymbol{S}(t) = \text{diag}(\boldsymbol{S}(t)) \boldsymbol{\mu} dt + \text{diag}(\boldsymbol{S}(t)) \boldsymbol{\sigma} d\boldsymbol{W}(t),$$

Cramér's Theorem (theoretical)

Theorem 3.1 (Cramér's Theorem). *Let $\{X_t\}_{t \geq 0}$ be i.i.d. real-valued random variables with logarithmic moment-generating function*

$$\Lambda(\lambda) := \log \mathbb{E}[e^{\lambda X_0}] < \infty \quad \text{for all } \lambda \in \mathbb{R}.$$

Then for any $c > \mathbb{E}[X_0]$, the large-deviation upper bound holds:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log \Pr \left(\frac{1}{T} \sum_{t=1}^T X_t \geq c \right) = -I(c),$$

where the rate function $I(c)$ is given by the Legendre transform

$$I(c) = \sup_{\lambda \in \mathbb{R}} \{ \lambda c - \Lambda(\lambda) \}.$$

Cramér's Theorem (Estimate)

Remark 3.1 (Cramér-type Probability Estimate). In practice, we apply Cramér's theorem to approximate rare-event probabilities of the form $\Pr(\bar{L}_T(\mathbf{w}) \geq c)$, where $\{L_t(\mathbf{w})\}$ are portfolio log-returns. Assuming the returns are i.i.d. and have a well-defined log-moment generating function, the approximation is

$$\Pr(\bar{L}_T(\mathbf{w}) \geq c) \approx \exp(-T I(c; \mathbf{w})), \quad T \gg 1,$$

with the rate function defined as

$$I(c; \mathbf{w}) := \sup_{\lambda \in \mathbb{R}} \left\{ \lambda c - \log \mathbb{E} \left[e^{\lambda L_1(\mathbf{w})} \right] \right\}.$$

Objective function

Lemma 3.1 (MGF of a Normal Random Variable). *Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then its moment-generating function is*

$$M_X(\lambda) = \mathbb{E}[e^{\lambda X}] = \exp\left(\lambda\mu + \frac{1}{2}\lambda^2\sigma^2\right), \quad \lambda \in \mathbb{R}.$$

Under the Gaussian log-return assumption $L_1(\mathbf{w}) \sim \mathcal{N}(\mu_L, \sigma_L^2)$, Lemma 3.1 implies that

$$I(c; \mathbf{w}) = \sup_{\lambda \in \mathbb{R}} \left\{ \lambda c - \lambda \mu_L - \frac{1}{2} \lambda^2 \sigma_L^2 \right\} = \frac{(c - \mu_L)^2}{2\sigma_L^2}.$$

Substituting

$$\mu_L = \mathbf{w}^\top \boldsymbol{\mu} - \frac{1}{2} \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}, \quad \sigma_L^2 = \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w},$$

the large-deviation objective becomes

$$f(\mathbf{w}) := \frac{\left(c - \mathbf{w}^\top \boldsymbol{\mu} + \frac{1}{2} \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}\right)^2}{2 \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}}, \quad \mathbf{w} \in \mathcal{W}.$$

Find the optimum via FOC (first order condition)

Step 1: First-order condition (FOC). Imposing $\nabla_w f(w) = 0$ and applying the quotient rule gives

$$-\mu + \Sigma w = \frac{c - w^\top \mu + \frac{1}{2} w^\top \Sigma w}{w^\top \Sigma w} \Sigma w.$$

Let $\gamma := (c - w^\top \mu + \frac{1}{2} w^\top \Sigma w) / (w^\top \Sigma w)$; then $(1 - \gamma) \Sigma w = \mu$, so

$$w = \frac{1}{1 - \gamma} \Sigma^{-1} \mu.$$

Consequently, the optimal weight vector always lies on the $\Sigma^{-1} \mu$ ray; its exact scale is determined by γ (fixed in the next step).

Step 2: Solve for the scale. Write $w = \lambda \Sigma^{-1} \mu$, $\lambda > 0$ and substitute into $f(w)$. Let

$$d := \mu^\top \Sigma^{-1} \mu > 0.$$

Then

$$f(\lambda) = \frac{[c + d(-\lambda + \frac{1}{2} \lambda^2)]^2}{2\lambda^2 d}, \quad \lambda > 0.$$

Differentiating and setting $f'(\lambda) = 0$ gives the unique minimiser

$$\lambda^* = \sqrt{\frac{2c}{d}}.$$

Closed-form optimal weights. Substituting λ^* yields the closed-form solution

$$w^* = \sqrt{\frac{2c}{\mu^\top \Sigma^{-1} \mu}} \Sigma^{-1} \mu. \quad (1)$$

Data set (45 US stocks)

GICS Sector	Tickers
Communication Services	DIS
Consumer Discretionary	MCD, F, TGT
Consumer Staples	KO, PG, WMT, PEP, MO, SYY, CL
Energy	XOM, CVX
Financials	TRV, AXP, BK, USB, WFC, AIG, BAC, C, MMC
Health Care	JNJ, MRK, PFE
Industrials	GE, MMM, BA, CAT, HON, DE, EMR, GD, LMT, UNP
Information Technology	IBM, TXN, HPQ, XRX, AVT
Materials	DD, ECL, AA
Utilities	XEL, ED

Time horizon: 1980 - 2024

Training window: 10, 11, 12, ..., 20

Portfolio Construction

(i) **Large-deviation portfolio w^\star .** Using the closed-form solution (1) with the empirical $(\hat{\mu}, \hat{\Sigma})$ and daily performance target $c = \log(1.5)/252$,

$$\tilde{w} = \sqrt{\frac{2c}{\hat{\mu}^\top \hat{\Sigma}^{-1} \hat{\mu}}} \hat{\Sigma}^{-1} \hat{\mu}.$$

Because $\tilde{w} \notin \mathcal{W}$, we project it onto \mathcal{W} by solving $\min_{w \in \mathcal{W}} \|w - \tilde{w}\|_2^2$, yielding the implementable w_{LDP} .

(ii) **Mean-variance (MV) portfolio.** We solve the classical Markowitz problem

$$\max_{w \in \mathcal{W}} (\hat{\mu}^\top w - \tfrac{1}{2} w^\top \hat{\Sigma} w),$$

to obtain w_{MV} .

(iii) **Equal-weight portfolio.** The naïve benchmark is $w_{\text{EQ}} = (1/n, \dots, 1/n)^\top$.

Performance Metrics

- Annual return: annually Compounding
- Volatility: annualized sample variance
- Sharpe ratio: assuming risk-free rate is 0
- Max Drawdown: the largest peak-to-trough decline
- Turnover: L1 distance between portfolio weights at each rebalance point (annually)

Result (table)

[illegible]

Conclusion

1. The Large-Deviation-based Portfolio (LDP) optimization method may effectively reduce **volatility** and **maximum drawdown** compared to traditional approaches.
2. **A training window of 16–18 years** appears to offer a good trade-off between estimation reliability and adaptability, producing more stable out-of-sample performance.

Research direction

- (i) Extend beyond Gaussian returns.
- (ii) Incorporate transaction costs into the objective or constraints.
- (iii) Develop adaptive rules for calibrating the return target c .

Github

https://github.com/ryandaiqf/LDP_Portfolio/tree/main