Large–Deviation–Based Portfolio Optimization

Ryan Dai
Department of Quantitative Finance
2025.06.09

Content

Introduction: Portfolio optimization

- Mean-Variance portfilio 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.

Portfolio Setup

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

$$\boldsymbol{r}_t = (r_{1,t}, \ldots, r_{n,t}), \qquad r_{i,t} := \frac{S_{i,t+1}}{S_{i,t}} - 1.$$

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

$$\boldsymbol{w} = (w_1, \ldots, w_n) \in \mathcal{W} \subseteq \mathbb{R}^n, \qquad \mathcal{W} := \Big\{\boldsymbol{w} \mid \sum_i w_i = 1,\ 0 \le w_i \le 0.1\Big\}.$$

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

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

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

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

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

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

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

Market Assumption

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

$$\mathrm{d}\boldsymbol{S}(t) = \operatorname{diag}\big(\boldsymbol{S}(t)\big)\, \boldsymbol{\mu}\, \mathrm{d}t + \operatorname{diag}\big(\boldsymbol{S}(t)\big)\, \boldsymbol{\sigma}\, \mathrm{d}\boldsymbol{W}(t),$$

where $\boldsymbol{\mu} \in \mathbb{R}^n$ *is the drift vector,* $\boldsymbol{\sigma} \in \mathbb{R}^{n \times n}$ *is the volatility matrix, and* $\boldsymbol{W}(t)$ *is an n-dimensional standard Brownian motion.*

For any fixed weight vector $\boldsymbol{w}$*, let the portfolio value be* $P(t) := \boldsymbol{w}^\top \boldsymbol{S}(t)$*. The one-day log-return*

$$L_t(\boldsymbol{w}) := \log\Big(\tfrac{P_{t+1}(\boldsymbol{w})}{P_t(\boldsymbol{w})}\Big)$$

is i.i.d. with

$$L_t(\boldsymbol{w}) \sim \mathcal{N}(\mu_L, \sigma_L^2), \qquad \mu_L := \boldsymbol{w}^\top \boldsymbol{\mu} - \tfrac{1}{2}\boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w},\ \sigma_L^2 := \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w},$$

where $\boldsymbol{\Sigma} := \boldsymbol{\sigma}\boldsymbol{\sigma}^\top$ *is the log-return covariance matrix.*

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*

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

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

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

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 $\Pr(\overline{L}_T(\boldsymbol{w}) \geq c)$, where $\{L_t(\boldsymbol{w})\}$ are portfolio log-returns. Assuming the returns are i.i.d. and have a well-defined log-moment generating function, the approximation is

$$\Pr\big(\overline{L}_T(\boldsymbol{w}) \geq c\big) \approx \exp\big(-T\, I(c; \boldsymbol{w})\big), \qquad T \gg 1,$$

with the rate function defined as

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

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*

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

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

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

Substituting

$$\mu_L = \boldsymbol{w}^\top \boldsymbol{\mu} - \tfrac{1}{2}\boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}, \qquad \sigma_L^2 = \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w},$$

the large-deviation objective becomes

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

Find the optimum via FOC (first order condition)

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

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

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

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

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

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

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

Then

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

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

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

Closed-form optimal weights. Substituting $\lambda^\star$ yields the closed-form solution

$$\boxed{\boldsymbol{w}^\star = \sqrt{\frac{2c}{\boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}}\,\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}. \tag{1}$$

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 $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ and daily performance target $c = \log(1.5)/252$,

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

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

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

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

to obtain $\boldsymbol{w}_{\mathrm{MV}}$.

(iii) Equal-weight portfolio. The naïve benchmark is $\boldsymbol{w}_{\mathrm{EQ}} = \left(1/n, \ldots, 1/n\right)^\top$.

Performance Metrics

- Annual return: annully Comounding
- 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)

Metric \ Window	10	11	12	13	14	15	16	17	18	19	20
Annualised Return											
LDP	0.0977	0.0966	0.0905	0.0943	**0.1027**	0.1015	**0.1047**	**0.1083**	**0.1056**	0.1007	0.1006
MV	0.0756	0.0840	0.0918	0.0869	0.0940	0.0873	0.0894	0.0889	0.0799	0.0656	0.0685
EQ	0.1026	0.1026	0.1026	0.1026	0.1026	0.1026	0.1026	0.1026	0.1026	0.1026	0.1026
Annualised Volatility											
LDP	**0.1821**	**0.1814**	**0.1804**	**0.1811**	**0.1794**	**0.1790**	**0.1807**	**0.1766**	**0.1780**	**0.1779**	**0.1802**
MV	0.2101	0.2065	0.2041	0.2055	0.2058	0.2040	0.2118	0.2101	0.2096	0.2073	0.2074
EQ	0.1952	0.1952	0.1952	0.1952	0.1952	0.1952	0.1952	0.1952	0.1952	0.1952	0.1952
Sharpe Ratio											
LDP	**0.5368**	**0.5327**	0.5016	0.5210	**0.5727**	**0.5670**	**0.5795**	**0.6134**	**0.5931**	**0.5661**	**0.5579**
MV	0.3600	0.4069	0.4496	0.4227	0.4567	0.4281	0.4219	0.4233	0.3813	0.3164	0.3301
EQ	0.5254	0.5254	0.5254	0.5254	0.5254	0.5254	0.5254	0.5254	0.5254	0.5254	0.5254
Max Drawdown											
LDP	**-0.4544**	**-0.4368**	**-0.4657**	**-0.4670**	**-0.4757**	**-0.4890**	**-0.5107**	**-0.5079**	**-0.4708**	**-0.4568**	**-0.4685**
MV	-0.5738	-0.4782	-0.5518	-0.5581	-0.5666	-0.5866	-0.6265	-0.6312	-0.6168	-0.5838	-0.5681
EQ	-0.5692	-0.5692	-0.5692	-0.5692	-0.5692	-0.5692	-0.5692	-0.5692	-0.5692	-0.5692	-0.5692
Average Turnover											
LDP	0.6327	0.6104	0.5995	0.6067	0.5896	0.5603	0.5760	0.5484	0.5145	0.4934	0.4562
MV	0.6470	0.6662	0.6346	0.6383	0.5564	0.5542	0.5360	0.5080	0.5036	0.5391	0.4882
EQ	0.1544	0.1544	0.1544	0.1544	0.1544	0.1544	0.1544	0.1544	0.1544	0.1544	0.1544

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