Portfolio Volatility Estimation: PCA and James-Stein Approaches

Zhuoli Jin

Department of Statistics and Applied Probability

September 20, 2024

Overview

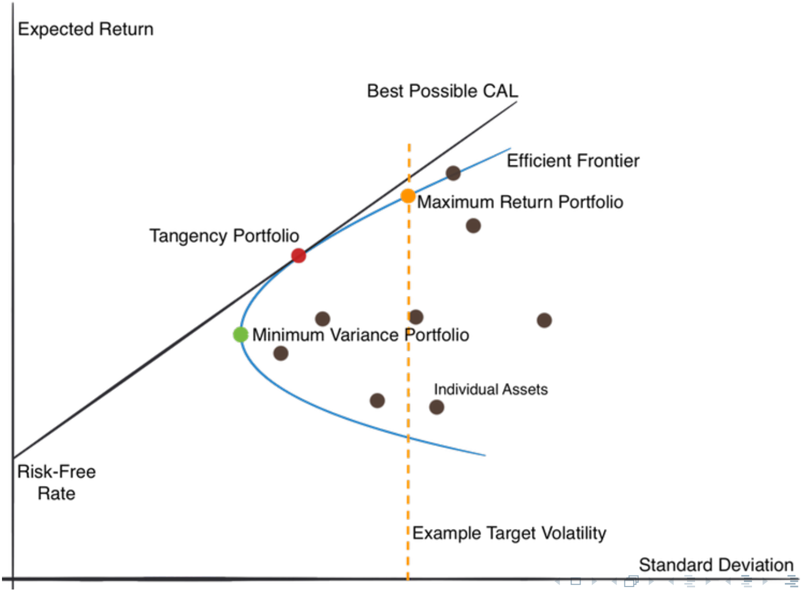
The Markowitz mean-variance portfolio is obtained from solving the following:

$$\begin{aligned} \min_{\omega} \quad & w^\top \Sigma w \\ \text{subject to} \quad & \mu^\top w \geq \alpha \\ & e^\top w = 1, \end{aligned} \tag{1}$$

where e is a p-vector of ones, α is a return target. μ is (true) means of asset returns and Σ is (true) covariance of asset returns.

Goal: determine $w = (w_1, w_2, \cdots, w_p)^\top$.

Efficient Frontier



Solution to Markowitz Problem

Use the Lagrange multipliers, the solution to Eq 1 is:

$$w = \gamma_e \Sigma^{-1} e + \gamma_\mu \Sigma^{-1} \mu. \tag{2}$$

- Case 1: $\frac{\mu^\top \Sigma^{-1} e}{e^\top \Sigma^{-1} e} \geq \alpha$:
  Pick the minimum variance solution $w = \frac{\Sigma^{-1} e}{e^\top \Sigma^{-1} e}$.
- Case 2: $\frac{\mu^\top \Sigma^{-1} e}{e^\top \Sigma^{-1} e} < \alpha$:
  Pick the tangency portfolio $w = \frac{1}{2} \Sigma^{-1} (\lambda_1 \mu + \lambda_2 e)$ with
  $A = e^\top \Sigma^{-1} e, B = e^\top \Sigma^{-1} \mu, C = \mu^\top \Sigma^{-1} \mu$.

Quick Discussion

Denote $p \times n$ matrix R as asset return.

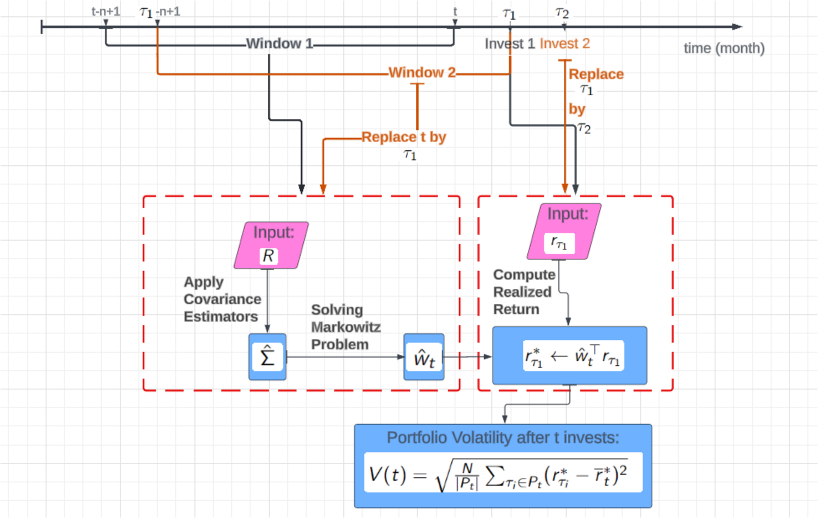
- μ: vector of estimated returns.
  Use m to denote the estimate.
  Common estimates:
    - Sample mean: $m = (\overline{R}_1, \overline{R}_2, \cdots, \overline{R}_p)^\top$ with $\overline{R}_i = \frac{1}{n}\sum_{j=1}^{n} R_{i,j}$
    - James-Stein return constraint.
- Σ: covariance matrix of the asset returns.
  Use $\hat{\Sigma}$ to denote the estimate.
  Sample covariance: S.
  Common estimates:
    - PCA estimator
    - James-Stein shrinkage estimator
    - Ledoit-Wolf shrinkage estimator
    - Other estimators (to be listed later).

Testing Framework I

r_t	Return to p assets on time (month) t
R_t	$R_t = (r_{t-n+1}, \cdots, r_t)$ a $p \times n$ matrix of returns at time t with lookback window of size n
Y_t	R_t centered, $Y_t = R_t - [\overline{r}_t, \cdots, \overline{r}_t]$ with $\overline{r_t}$ be the p-vector average of the n columns of R_t. Needed to compute sample covariance matrix
$\hat{w}_t$	Estimated portfolio weights at time t
N	The number of times t is incremented until a year goes by, e.g. r_t in units of monthly return implies $N = 12$
P_t	Set of times counting back from t

Table: Common Notations in Testing Framework

Testing Framework II



PCA I

1. Eigen decomposition on centered sample covariance matrix S.
2. Write sample covariance matrix as sum of two parts: the first part consists of K leading eigenpairs, the second part is the reminder.

$$S = H_{p \times k} H_{p \times k}^{\top} + G$$

3. The specific risk estimate Δ:

$$\Delta = diag(G)$$

Check Appendix 16 for details.

JSE I

1. Estimate return constraint vector μ using James-Stein shrinkage.
2. Apply shrinkage to the eigenvectors.

$$H_{JS} = HC + M(I - C),$$

where $M_{p \times k}$ is shrinkage target based on constraints μ and e, and $C_{k \times k}$ is the shrinkage matrix.

Check Appendix 18 for details.

Numerical Model

$r = \mu + BX + \epsilon$

- Number of factors: 7.
- $\mu \in \mathbb{R}^p$: Expected security return with mean 6.28.
- $\alpha = 8$: the portfolio return shall exceed 8.
- $B_{p\times 7}$: Factor loading (e.g: market factor, size factors, sector-specific factor, etc.) Check Appendix 13 for details.
- $x \in \mathbb{R}^7$: Factor return, Gaussian distributed
- $\epsilon \in \mathbb{R}^p$: Idiosyncratic risk, Gaussian distributed

We collect observations of r in $\mathbb{R}^p$, forming a data matrix for analysis.

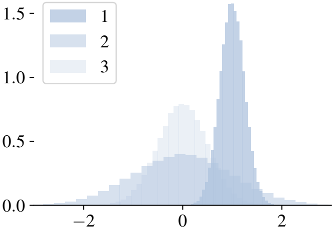
Numerical Results

p	PCA $V(t)$	James-Stein $V(t)$
500	7.82	6.74
1000	7.08	6.36
2000	7.18	5.07
3000	5.96	3.74
100000	5.61	0.84

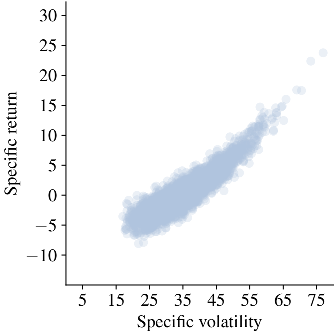
Table: Portfolio volatility statistics ($n = 125$, $\mu = 8$, 100 investment dates)

Appendix

Factor Loadings I

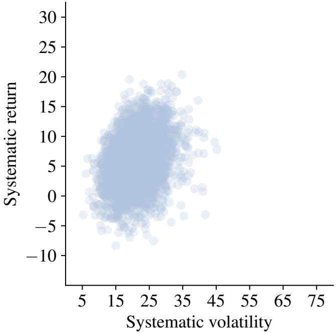


(a) Load Hist

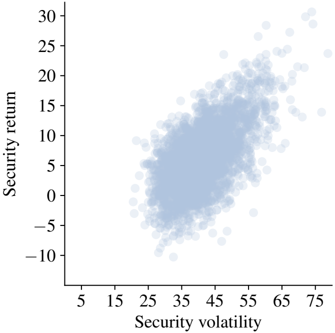


(b) Svol_Sret

Factor Loadings II

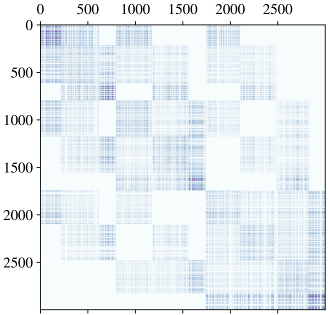


(c) Bvol_Bret



(d) vols_alph

Factor Loadings III



(e) Block Post

PCA I

PCA Assumption

- $p >> n$
- Low Effective Rank ($K << p$)
- Sparsity or Low-Rank Structure (the covariance matrix can be well approximated by a matrix with fewer non-zero eigenvalues).
- Noiseless Data or Low Noise

PCA II

PCA Recipe

1. Compute Sample Covariance Matrix: the sample covariance matrix S is computed as $S = \frac{1}{n} YY^\top$.
2. Perform Spectral Decomposition: decompose S into its eigenvalues and eigenvectors
$$S = V \Lambda V^T,$$
where $\Lambda = diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and $V = (v_1, v_2, \cdots, v_n)$ with each v_i the corresponding eigenvector of λ_i
3. Select Principal Components: choose the top K PCs – so we have V_K and Λ_K. Now we have $S = HH^T + G$ with $HH^\top = V_K \Lambda_K V_K^\top$ and G being the residuals.
4. Construct PCA-Based Covariance Matrix:
$$\hat{\Sigma}_{PCA} = HH^\top + \rho(G), \tag{3}$$
with $\rho()$ being the regularization operator on G to make $\hat{\Sigma}$ of full rank, for example
$$\rho(G) = Diag(G). \tag{4}$$

James-Stein I

JSE Assumption

1. n is fixed, $p \to \infty$.
2. $0 < \liminf_{p\uparrow\infty} \inf_{|v|=1} \langle v, \Sigma_U v\rangle < \limsup_{p\uparrow\infty} \sup_{|v|=1} \langle v, \Sigma_U v\rangle < \infty$
3. $\lim_{p\uparrow\infty} (B^\top B)/p$ exists as an invertible $K \times K$ matrix with fixed $K \geq 1$.
4. $\lim_{p\uparrow\infty} (H^\top B)/H$ exists as an invertible $K \times K$ matrix with fixed $K \geq 1$ for H in Eq**??**.
5. For $\zeta \in \mathbb{R}^p$ non vanishing and residing entirely in $Col(H)$, denote $\zeta_B = H(H^\top H)^{-1} H^\top B$,

$$\limsup_{p\uparrow\infty} |\zeta_B|/|\zeta| < 1, \text{ and } \liminf_{p\uparrow\infty} |\zeta| \neq 0.$$

6. For any $g \in \mathbb{R}^n$, $J = I - \frac{gg^\top}{|g|^2}$, F, U satisfies the following:
1.F and U are latent
2. K is known and the number of nonzero eigenvalues of sample covariance matrix S is larger than K.
3.$F^\top JF$ is invertible almost surely.
4.$\lim_{p\uparrow\infty} U^\top U/p = \gamma^2 I$ almost surely for some $\gamma > 0$
5. $\limsup_{p\uparrow\infty} \left\| JU^{-1}B \right\| /p = 0$ almost surely for some matrix norm $\|\cdot\|$ on $\mathbb{R}^{n\times K}$
6. $\limsup_{p\uparrow\infty} |JU^\top z|/\sqrt{p} = 0$ almost surely for $z = \zeta/|\zeta|$

James-Stein II

James-Stein Return Constraint

1. The James-Stein recipe to improve the sample mean estimate $m = \bar{r}$ is given for any p-vector $M \neq m$ by

$$m_{JS} = cm + (1-c)M, \quad c = 1 - \frac{\nu^2}{(m-M)^\top (m-M)}, \tag{5}$$

where

$$\nu^2 = \frac{tr(G)}{n_+ - K}, \tag{6}$$

and n_+ is the number of nonzero eigenvalues of S in PCA recipe.

Step 2: The shrinkage target M maybe any p-vector, but commonly be the grand mean

$$M = \frac{\langle m, e\rangle}{\langle e, e\rangle} e = (\sum_{i=1}^{p} m_i / p)e. \tag{7}$$

James-Stein III

James-Stein Recipe

1. For any estimate $\rho(G)$ that is not a scalar matrix, we replace Y by $Y = \rho(G)^{-1/2}(R - [\bar{r}, \bar{r}, \cdots, \bar{r}])$, where $\rho(G)^{-1/2}$ is diagonal with $\rho(G)_{ii}^{-1/2} = 1/\sqrt{\rho(G)_{ii}}$.
2. Recompute H following $S = HH^\top + G$ but from the reweighted sample covariance matrix S using above updated Y.
3. Take m_{js} in Eq5 and assemble the matrix
$$A = \rho(G)^{-1/2}(m_{js}\ e). \tag{8}$$
4. Compute the pesudo-inverse $A^\dagger = (A^\top A)^{-1}A^\top$, and take ν^2 in Eq 7, we define the variables
$$M = AA^\dagger H \quad J = (H - M)^\top(H - M) \quad C = I - \nu^2 J^{-1}. \tag{9}$$
5. The James-Stein estimator for H is
$$H_{JS} = HC + M(I - C) \tag{10}$$
for C a $K \times K$ matrix and M a $p \times k$ matrix.
6. The James-Stein estimator for Σ is
$$\hat{\Sigma}_{JS} = \rho(G)^{1/2}(H_{JS}H_{JS}^\top + I)\rho(G)^{1/2}. \tag{11}$$

References I