Analysis of Return-Covariance Eigenvectors and Associated Minimum-Variance Portfolios

Alex Bernstein *(joint with Alex Shkolnik)*
abernstein@ucsb.edu

September 13, 2024

Department of Statistics & Applied Probability
University of California, Santa Barbara

Introduction

Minimum Variance Portfolios with a Long-Only Constraint

Let $\mathrm{e} = \begin{pmatrix} 1, & 1, & \dots, & 1 \end{pmatrix} \in \mathbb{R}^p$. We are interested in solutions of:

$$\begin{aligned} \min_{x \in \mathbb{R}^p} \quad & x^\top \mathbf{\Sigma} x \\ \text{s.t.} \quad & x^\top \mathrm{e} = 1 \\ & x \geq 0 \end{aligned}$$

where

$$\mathbf{\Sigma} = \sigma^2 \beta \beta^\top + \mathbf{\Delta}$$

and $\sigma^2 \in \mathbb{R}_+$, $\beta \in \mathbb{R}^p$, and $\mathbf{\Delta} = \textbf{diag}(\delta_1^2, \delta_2^2, \dots, \delta_p^2)$ is diagonal in $\mathbb{R}^{p \times p}$ with $\delta_i^2 > 0$ for $i \in \{1, \dots, p\}$.

The $x \geq 0$ condition is the Long-Only (LO) constraint. If it is excluded, we say that the portfolio is Long-Short (LS).

Long-Only Portfolio Solution

Theorem (Solution to the Minimum Variance Portfolio)

Define

$$\psi(t) = \frac{\sum_{t\beta_i<1} \beta_i/\delta_i^2}{1/\sigma^2 + \sum_{t\beta_i<1} \beta_i^2/\delta_i^2}$$

The minimum variance portfolio x can be found by computing the following:

$$\theta = \psi(\theta)$$

$$w_i = \frac{\max(1-\theta\beta_i, 0)}{\delta_i^2}, \quad x_i = \frac{w_i}{\sum_{j=1}^p w_j}$$

θ can be found via fixed-point iterations (i.e. computing $\theta_{k+1} = \psi(\theta_k)$ until $\theta_k = \theta_{k+1}$)

But How Do We Estimate Parameters?

Our solution x is clearly a function of three parameters: σ^2, β and $\boldsymbol{\Delta}$.

These are unknown parameters, and must be estimated

Principal Component Analysis Estimation

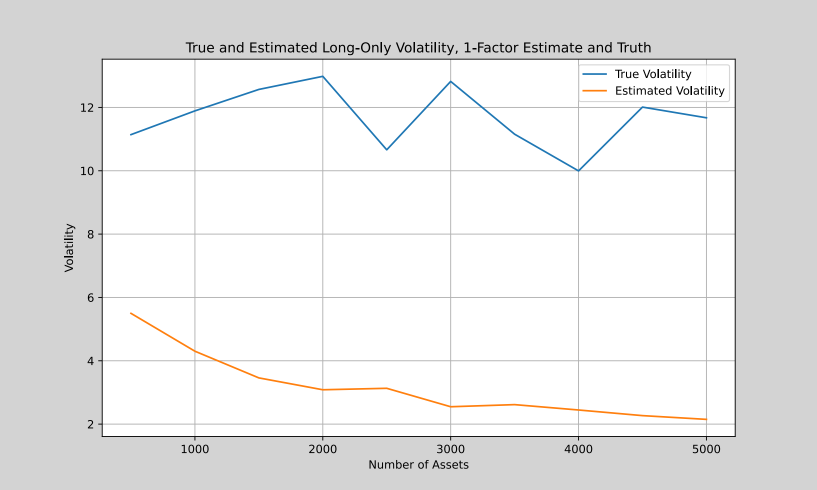
Model Parameters for Estimated 1-Factor Model

Define:

- $Y \in \mathbb{R}^{p \times n}$ be a matrix of security returns for p securities
- $\boldsymbol{\Sigma}$ is the True Covariance of all p securities
- $S = \frac{YY^\top}{n}$ is the $p \times p$ Sample Covariance matrix
- By spectral decomposition, we may write $S = \hat{\sigma}^2 \hat{\beta} \hat{\beta}^\top + G$ for some matrix G
- $\hat{\boldsymbol{\Sigma}} = \hat{\sigma}^2 \hat{\beta} \hat{\beta}^\top + \textbf{diag}(G)$. We define $\hat{\boldsymbol{\Delta}} = \textbf{diag}(G)$ as the diagonal matrix containing estimates for idiosyncratic variance

How does the portfolio $x(\hat{\sigma}^2, \hat{\beta}, \hat{\boldsymbol{\Delta}})$ do?

PCA For Long-Only Portfolios



Parameterized James-Stein Eigenvector Estimation

Shrinkage Methodology

Principal Components are estimated vectors, and can thus be improved via "shrinkage"

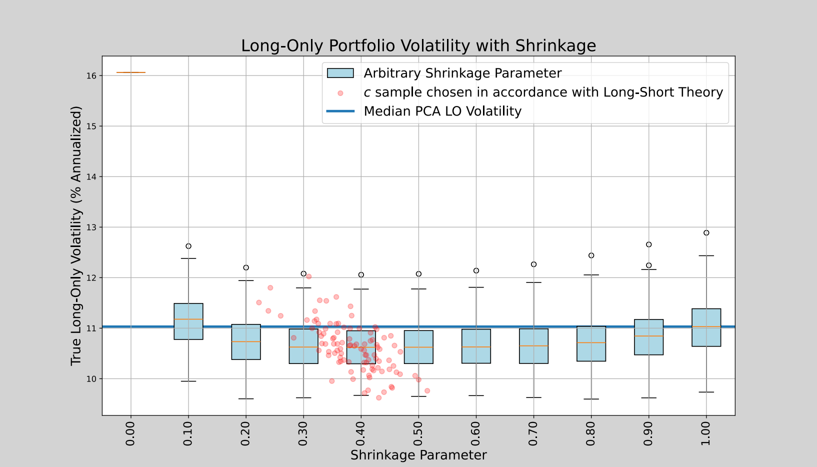
James-Stein Eigenvector Shrinkage

Let $\hat{\beta}$ be the first principal component and $q = \frac{\mathrm{e}}{|e|}$ (i.e. q is the vector identical values of length 1). Define:

(1) $$\hat{\eta} = \hat{\eta}(c) = c\hat{\beta} + (1-c)\langle\hat{\beta}, q\rangle q, \quad 0 \le c \le 1$$

The choice of c is a parameter. A theoretically optimal value for c for Long-Short portfolios is available in the literature. How does it perform in the Long-Only setting?

PCA For Long-Only Portfolios



Choice of Shrinkage Parameter

- It appears that the Long-Short JS parameter does somewhat well in this situation, but we have no theoretical justification in the Long-Only case
- How do we choose an optimal shrinkage parameter in the Long-Only Case?
- Introduce a new idea: portfolio sensitivity

Portfolio Sensitivity Hypothesis

Guiding Intuition

- Our goal is to find a portfolio that performs well even when our estimates (for example, of β) are wrong
- Portfolios that have weights that are not sensitive to changes in parameters should be more robust
- Ideally, portfolio sensitivity (which we will define) captures sensitivity to estimation error
- We will test this hypothesis on the James-Stein Shrinkage methodology

Portfolio Sensitivity I

Theorem (β Portfolio Derivative)
We can differentiate the (long-only) minimum-variance portfolio x with respect to the model parameters β_j as follows:

(2)

$$\frac{\partial x_i}{\partial \beta_j} = \sum_{k=1}^{p} \frac{\partial x_i}{\partial w_k}\frac{\partial w_k}{\partial \beta_j}; \quad \frac{\partial x_i}{\partial w_k} = \frac{\mathbf{1}_{\{i=k\}} \sum_{\ell=1}^{p} w_\ell - w_i \mathbf{1}_{\{\theta\beta_i<1\}}}{\sum_{(\ell=1}^{p} w_\ell)^2}$$

(3)

$$\frac{\partial w_k}{\partial \beta_j} = \frac{-\mathbf{1}_{\{\theta\beta_k<1\}}}{\delta_k^2}\left(\beta_k \frac{\partial \theta}{\partial \beta_j} + \theta \mathbf{1}_{\{j=k\}}\right)$$

(4)

$$\frac{\partial \theta}{\partial \beta_j} = \left(\frac{1}{\delta_j^2}\right)\left(\frac{1 - 2\theta(\beta_j)\beta_j}{1/\sigma^2 + \sum_{\theta\beta_k<1} \beta_k^2/\delta_k^2}\right)\mathbf{1}_{\{\theta\beta_i<1\}}$$

Similar theorems exist for σ^2 and δ_i^2

Portfolio Sensitivity II

We have a notion of sensitivity (i.e. a derivative) with respect to an individual β_i element.

We usually care about the statistical aspects of β, however. We therefore want a notion of sensitivity with respect to sample statistics g of β; with $g : \mathbb{R}^p \mapsto \mathbb{R}$ Define:

$$\frac{\partial x_i(\beta)}{\partial g(\beta)}(v) = \frac{\langle \nabla x_i(\beta), v \rangle}{\langle \nabla g(\beta), v \rangle}$$

a natural choice is $v = \nabla g(\beta)$ as this maximizes the change in $g(\beta)$

Choice of Statistic: Dispersion

We are focusing on a specific statistic of β, namely the *Dispersion*. Define:

- Let $\bar{x} = \frac{\sum_{i=1}^{p} \beta_i}{p}$ be the Sample Mean of β
- Let $\mathrm{Var}(\beta) = \frac{\sum_{i=1}^{p} (\beta_i - \bar{\beta})^2}{p}$ be the Variance of β

The Dispersion of β is defined as:

$$\mathrm{disp}(\beta) = \frac{\sqrt{\mathrm{Var}(\beta)}}{\bar{\beta}} = \frac{\sqrt{\frac{1}{n} \sum_{\ell=1}^{n} (\beta_\ell - \bar{\beta})^2}}{\bar{\beta}} \tag{5}$$

Dispersion Sensitivity

Using the our definitions of sensitivity and dispersion, we can define the sensitivity of a single portfolio weight with respect to dispersion:

$$\frac{\partial x_i}{\partial \mathrm{disp}(\beta)} = \frac{\langle x_i(\beta), \nabla \, \mathrm{disp}(\beta) \rangle}{\langle \nabla \, \mathrm{disp}(\beta), \nabla \, \mathrm{disp}(\beta) \rangle}$$

In order to choose our shrinkage parameter, we need a way to compare sensitivities of $x(\beta_1)$ and $x(\beta_2)$ for the entire vector. To do so, define:

$$\varrho(\beta) = \left| \begin{pmatrix} \frac{\partial x_1}{\partial \mathrm{disp}(\beta)} & \frac{\partial x_2}{\partial \mathrm{disp}(\beta)} \dots & \frac{\partial x_p}{\partial \mathrm{disp}(\beta)} \end{pmatrix} \right|_2$$

as the sensitivity of $x(\beta)$ with respect to dispersion

Testing Dispersion Sensitivity Hypothesis

Long-Only Robustness

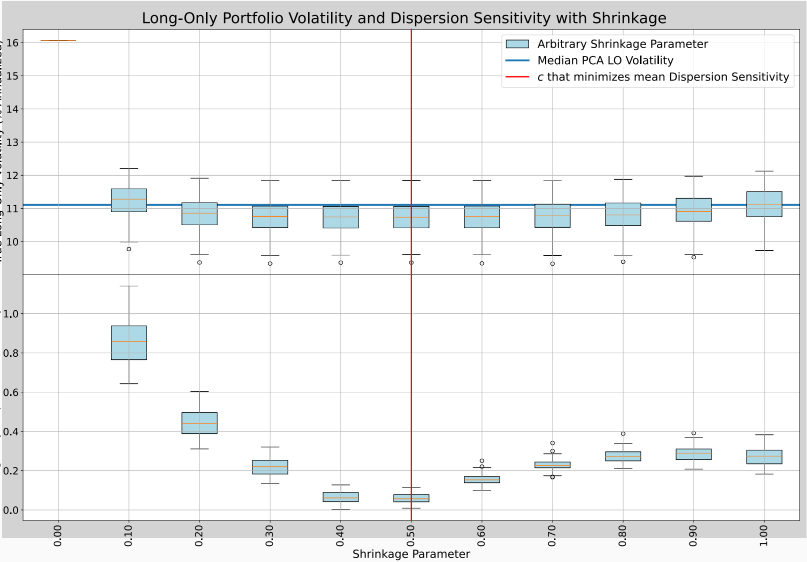
... Back to shrinkage

Objective: We want to find a better-performing portfolio via shrinkage on β

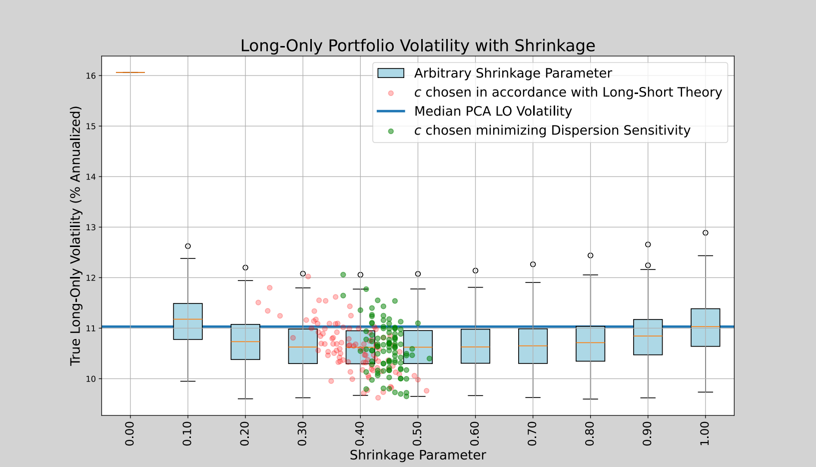
Idea: select c that minimizes $\varrho(\eta(c))$.

This can be done via a parameter search, as one may compute $\varrho(\eta(c))$ for know values of c.

Dispersion Sensitivity and Long-Only Volatility Comparison



Effect of Minimizing Dispersion Sensitivity



Mysterious Similarity to Long-Short James Stein

Optimal selection via dispersion sensitivity generally matches the performance (i.e. experimental LO Volatility) of the optimal portfolio found when Long-Short optimal James-Stein β shrinkage is used to correct β.

A priori- there is no reason for this match, as the corresponding Long-Short JS correction also has much higher dispersion sensitivities when compared to the ϱ-minimizing portfolio.