Effects of Covariance Eigenstructure Estimation Error on Portfolio Risk Measurement

Stjepan Begušić (stjepan.begusic@fer.hr)

August 2, 2023

Laboratory for Financial and Risk Analytics (lafra.fer.hr)
Faculty of Electrical Engineering and Computing
University of Zagreb, Croatia

Laboratory for Financial
and Risk Analytics

Introduction

Covariance estimation and applications in finance

Portfolio optimization:

- For unconstrained portfolios, positive definiteness is crucial[1]
- Mitigate the ”error maximization” property: bias-variance tradeoff [2]

Portfolio risk prediction[3]:

- Positive definiteness not crucial but still nice to have
- Minimize estimation noise - get closer to the ”ground truth”
- Would be nice to have a single risk model for different portfolios

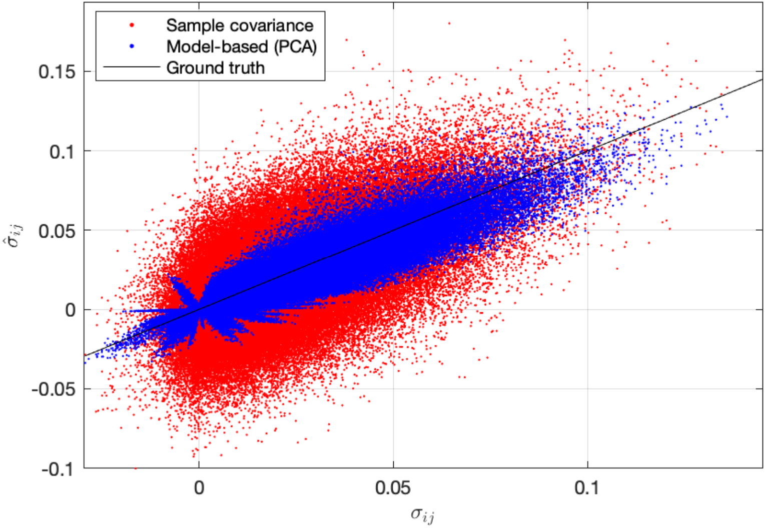
---

[1] Jagannathan, R., & Ma, T. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. Journal of Finance, 58(4), 1651–1683.

[2] Deshmukh, S., & Dubey, A. (2020). Improved Covariance Matrix Estimation with an Application in Portfolio Optimization. IEEE Signal Processing Letters, 27, 985–989.

[3] Menchero, J., & Li, P. (2020). Correlation Shrinkage: Implications for Risk Forecasting. Journal of Investment Management, 18(3).

Covariance estimation noise – example



How to measure the quality of (portfolio) variance estimation?

Primary interest – quality of the portfolio variance estimate:

$$\hat{\sigma}^2 = w^\intercal \hat{\Sigma} w$$

Variance forecast ratio:

$$VFR = \frac{\hat{\sigma}^2}{\sigma^2} = \frac{w^\intercal \hat{\Sigma} w}{w^\intercal \Sigma w} \tag{1}$$

- Ideally VFR is close to 1
- Overestimation of risk leads to $VFR > 1$, underestimation to $VFR < 1$

How to measure the quality of (portfolio) variance estimation?

Primary interest – quality of the portfolio variance estimate:

$$\hat{\sigma}^2 = w^\intercal \hat{\Sigma} w$$

Variance forecast ratio:

$$VFR = \frac{\hat{\sigma}^2}{\sigma^2} = \frac{w^\intercal \hat{\Sigma} w}{w^\intercal \Sigma w} \tag{1}$$

- Ideally VFR is close to 1
- Overestimation of risk leads to $VFR > 1$, underestimation to $VFR < 1$

In the example from the previous slide:

	$\|\|\hat{\Sigma} - \Sigma\|\|_F$	VFR
Sample estimate	12.60	0.91
PCA estimate	5.07	0.91

Problem

(How) do certain covariance estimators improve portfolio risk forecasting?

What is the effect of eigenvalue or eigenvector shrinkage on risk forecasting performance?

Experiment:

- High dimension low sample size ($p > n$) assumed
- A single-factor model is assumed, with several model-based estimators (PCA, S-POET, JSE) considered
- Non-optimized portfolios (equal weighted and market-cap weighted), optimized portfolios (minimum variance)

Model

Model

- Random vector of observable variables $X = [X_1, X_2, ..., X_p]^\intercal$
- Unobservable common factor $F \in \mathbb{R}$, with variance σ^2
- Unobservable factor loadings $\beta = [\beta_1, \ldots, \beta_p]^\intercal$
- Unobservable specific factors $\varepsilon = [\varepsilon_1, \varepsilon_2, ..., \varepsilon_p]^\intercal$, with covariance $\delta^2 I$

Model:

$$X = \beta F + \varepsilon.$$

Covariance under the model:

$$\Sigma = \sigma^2 \beta\beta^\intercal + \delta^2 I = \eta^2 bb^\intercal + \delta^2 I,$$

where $b = \beta/|\beta|$ is a unit vector, and $\eta^2 = \sigma^2|\beta|^2$.

Estimators

- Data sample $X \in \mathbb{R}^{p \times n}$, with $p > n$
- Sample covariance S
  - Largest eigenvalue $\hat{\lambda}^2$ and corresponding eigenvector $\hat{h}$
  - Mean of the remaining nonzero eigenvalues:

$$\hat{l}^2 = \frac{\mathrm{tr}(\hat{\Sigma}_{sample}) - \hat{\lambda}^2}{n - 1}$$

- Consider the class of estimators:

$$\hat{\Sigma} = \hat{\eta}^2 \hat{b} \hat{b}^\intercal + \hat{\delta}^2 I$$

VFR under the model

Portfolio variance and VFR:

$$\sigma_p^2 = w^\intercal \Sigma w = \eta^2 (w^\intercal h)^2 + \delta^2 w^\intercal w$$

$$VFR = \frac{\hat{\eta}^2 (w^\intercal \hat{h})^2 + \hat{\delta}^2 w^\intercal w}{\eta^2 (w^\intercal h)^2 + \delta^2 w^\intercal w}$$

- For large p, the idiosyncratic component is diversified away and only the common component matters.

Portfolios

Portfolios

Equal-weighted portfolio:

$$w_{EW} = [1/p, 1/p, ..., 1/p]^\intercal$$

Special case - VFR of the EW portfolio:

$$\sigma_{EW}^2 = w_{EW}^\intercal \Sigma w_{EW} = \eta^2 (w_{EW}^\intercal h)^2 + \delta^2 w_{EW}^\intercal w_{EW} = \eta^2 \mu_h^2 + \delta^2 / p$$

$$VFR_{EW} = \frac{\hat{\eta}^2 \mu_{\hat{h}}^2 + \hat{\delta}^2 / p}{\eta^2 \mu_h^2 + \delta^2 / p}$$

- For large p, only the eigenvalue and eigenvector mean matter (but not its direction)!

Portfolios

Capitalization-weighted portfolio:

$$w_{MC} = \frac{\tilde{w}}{\sum \tilde{w}}, \quad \tilde{w}_i = \frac{1}{i}, \quad i = 1, ..., p$$

Minimum variance portfolio (using a covariance estimate $\hat{\Sigma}_{est.}$):

$$w_{est.} = \frac{\hat{\Sigma}_{est.}^{-1} \mathbf{1}}{\mathbf{1}^\intercal \hat{\Sigma}_{est.}^{-1} \mathbf{1}}$$

Estimators

Estimators

PCA estimator:

- $\hat{\eta}^2_{PCA} = \hat{\lambda}^2 - \frac{n-1}{p-1}\hat{l}^2$
- $\hat{b}_{PCA} = \hat{h}$
- $\hat{d}^2_{PCA} = \frac{n-1}{p-1}\hat{l}^2$

[4] Wang, W., & Fan, J. (2017). Asymptotics of empirical eigenstructure for high dimensional spiked covariance. Annals of Statistics, 45(3), 1342–1374.

Estimators

PCA estimator:

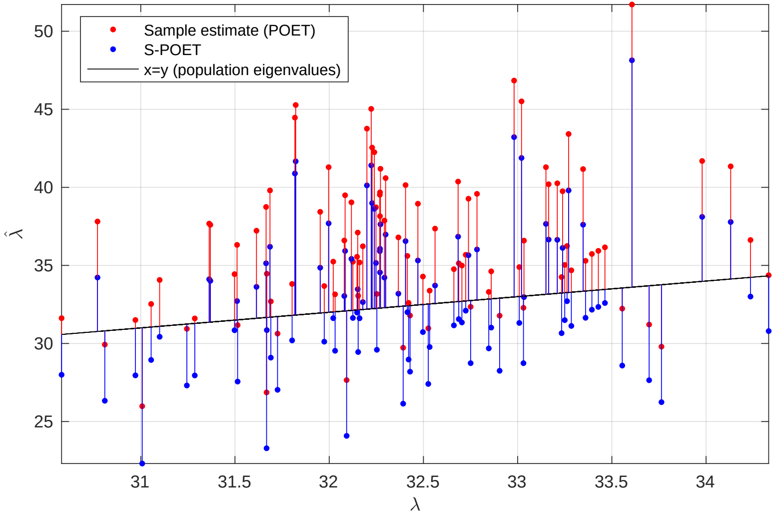
- $\hat{\eta}^2_{PCA} = \hat{\lambda}^2 - \frac{n-1}{p-1}\hat{l}^2$
- $\hat{b}_{PCA} = \hat{h}$
- $\hat{d}^2_{PCA} = \frac{n-1}{p-1}\hat{l}^2$

S-POET estimator[4]:

- Eigenvalue shrinkage constant: $c_{SPOET} = \frac{\mathrm{tr}(\hat{\Sigma}_{sample}) - \hat{\lambda}^2}{p-1-p/n}$
- $\hat{\eta}^2_{SPOET} = \hat{\lambda}^2 - \frac{p}{n}c_{SPOET}$
- $\hat{b}_{SPOET} = \hat{h}$
- $\hat{d}^2_{SPOET} = \frac{p-1}{p}c_{SPOET}$

[4] Wang, W., & Fan, J. (2017). Asymptotics of empirical eigenstructure for high dimensional spiked covariance. Annals of Statistics, 45(3), 1342–1374.

Eigenvalue shirnkage



JSE estimator[5]:

- Eigenvector shrinkage:

$$\hat{h}_{JSE} = \frac{\hat{\mu}_h \mathbf{1} + c_{JSE}(\hat{h} - \hat{\mu}_h \mathbf{1})}{|\hat{\mu}_h \mathbf{1} + c_{JSE}(\hat{h} - \hat{\mu}_h \mathbf{1})|},$$

  with $\hat{\mu}_h = 1/p \sum \hat{h}_i$ and $c_{JSE} = \hat{l}^2/p$
- $\hat{\eta}^2_{JSE} = \hat{\eta}^2_{SPOET}$
- $\hat{b}_{JSE} = \hat{h}_{JSE}$
- $\hat{d}^2_{JSE} = \hat{d}^2_{SPOET}$

[5] Goldberg, L. R., & Kercheval, A. N. (2023). James–Stein for the leading eigenvector. Proceedings of the National Academy of Sciences of the United States of America, 120(2).

Eigenvector shrinkage

Dispersion bias[6]:

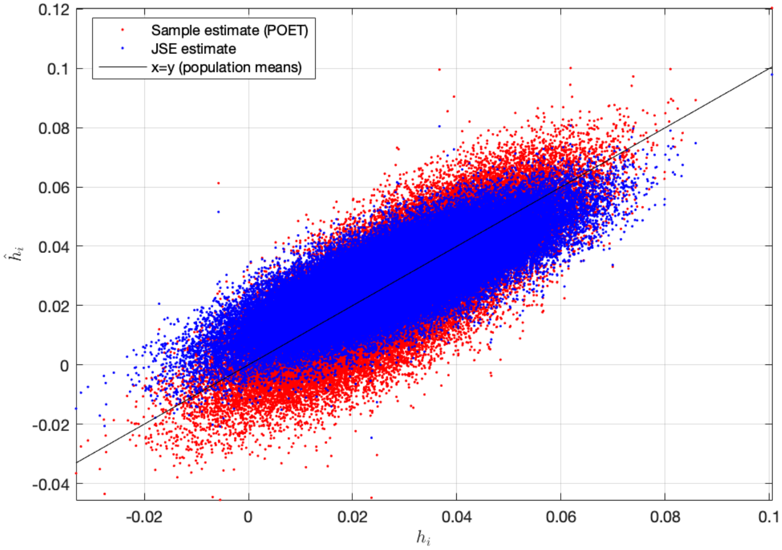
- Sample eigenvectors have lower mean $\hat{\mu}_h < \mu_h$ and excess dispersion:

$$d^2(\hat{h}) = \frac{1}{p}\sum_{i=1}^{p}(\frac{h_i}{\hat{\mu}_h} - 1)^2 = \frac{1 - p\hat{\mu}_h^2}{p\hat{\mu}_h^2} > \frac{1 - p\mu_h^2}{p\mu_h^2}$$

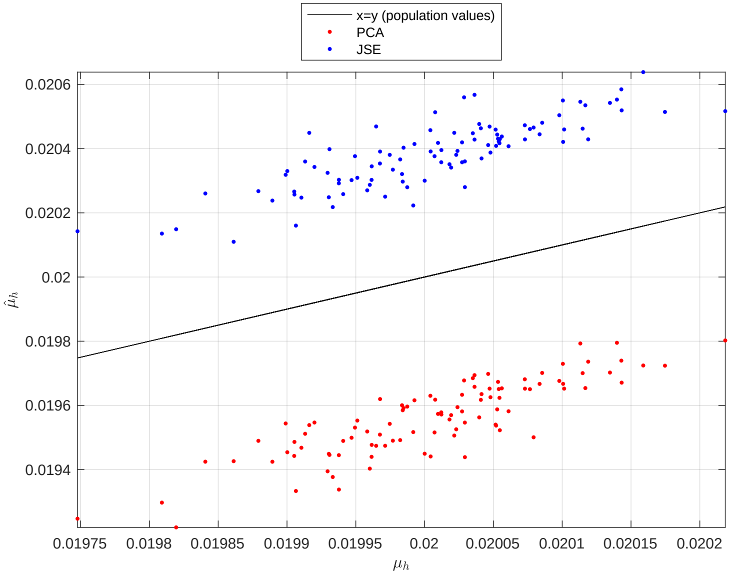
- The excess dispersion induces an optimization bias in minimum variance portfolios - eliminated by the JSE estimator

[6] Goldberg, L. R., Papanicolaou, A., & Shkolnik, A. (2022). The Dispersion Bias. SIAM Journal on Financial Mathematics, 13(2), 521–550.

Eigenvector shrinkage



Eigenvector shrinkage



Asymptotically, the amount of excess dispersion is known[7]:

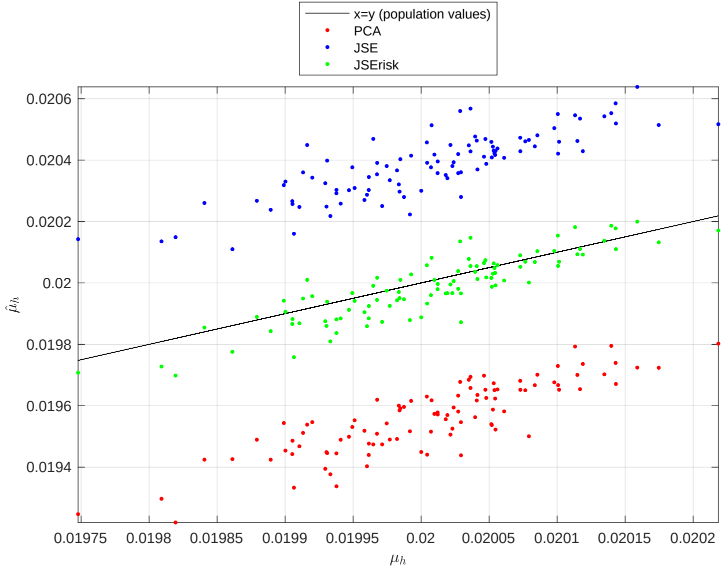
$$d^2_\infty(\hat{h}) = d^2_\infty(h) + \frac{1-\psi^2}{p\hat{\mu}_h^2}, \quad \text{with } \psi^2 = \frac{\hat{\lambda}^2 - \hat{l}^2}{\hat{\lambda}^2} \tag{2}$$

A different shrinkage constant to account for the asymptotic excess dispersion:

$$c_{JSErisk} = \frac{\psi - p\hat{\mu}_h^2}{1 - p\hat{\mu}_h^2} \tag{3}$$

[7] Goldberg, L. R., Papanicolaou, A., & Shkolnik, A. (2022). The Dispersion Bias. SIAM Journal on Financial Mathematics, 13(2), 521–550.

Eigenvector shrinkage



Experimental setup

Experimental setup

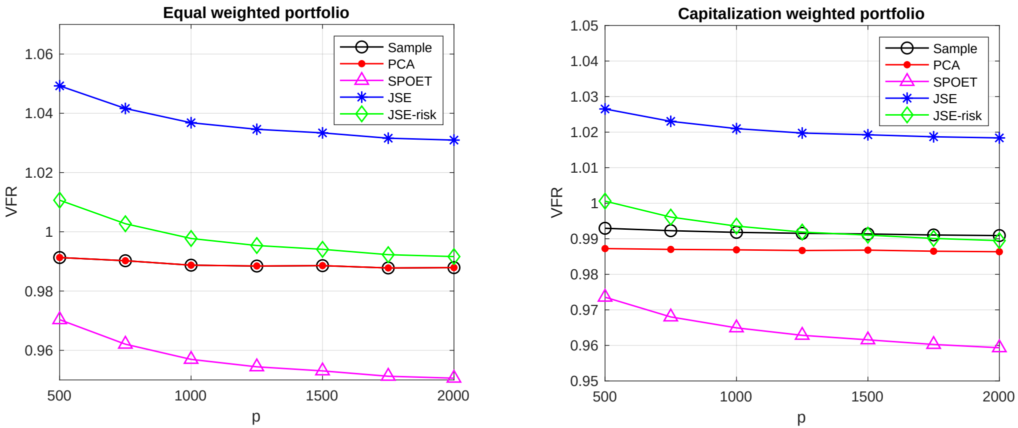
Factor exposures β_i	$\mathcal{N}(1, 0.25)$
Factor returns F	$\mathcal{N}(0, 0.16^2)$
Idiosyncratic returns ε_i	$T(5)$, mean 0 and variance 0.60^2
Number of variables p	$[500, 750, ..., 2000]$
Sample size n	250
Number of simulations N_s	100

For each simulation:

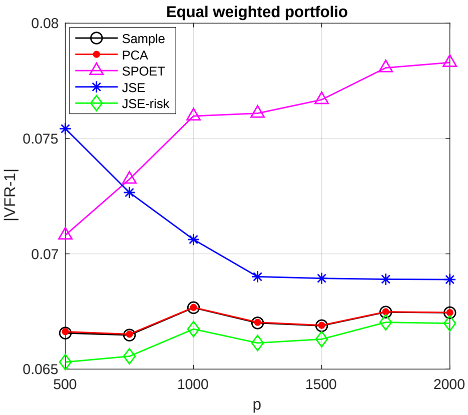
- Generate the model (factor exposures)
- Simulate the data matrix X with maximum p
- Slice the data to obtain the sample for specific p
- Obtain portfolios and variance forecasts for considered estimators

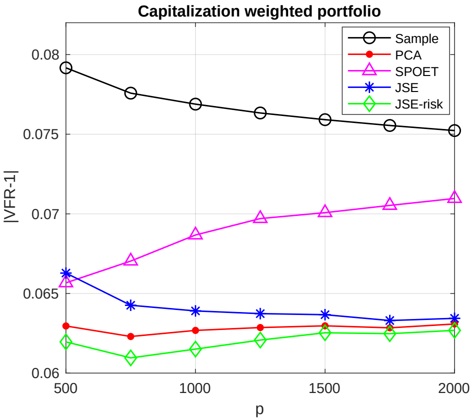
Portfolio variance forecasts

Bias in variance forecasting (non-optimized portfolios)

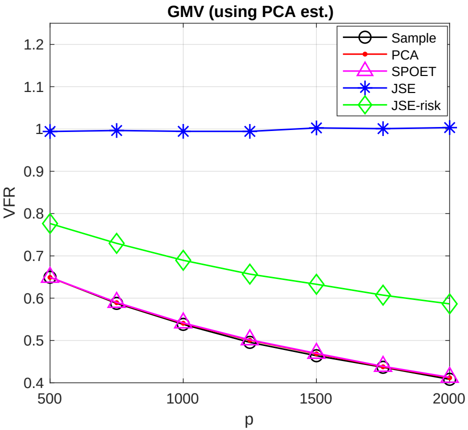


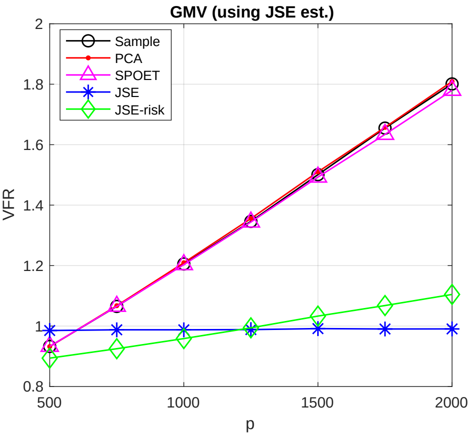
Relative error in variance forecasting (non-optimized portfolios)



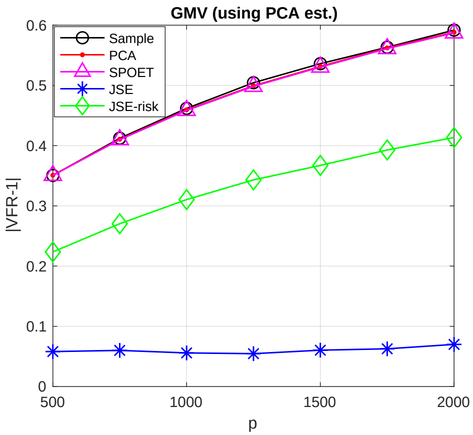


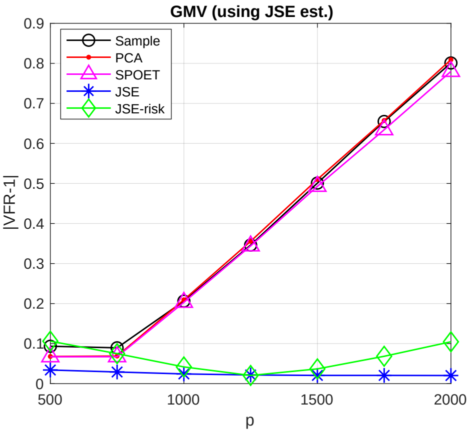
Bias in variance forecasting (minimum variance portfolios)





Relative error in variance forecasting (minimum variance portfolios)





Concluding remarks

Concluding remarks

- In HL, model-based covariance estimators may help overcome the bias-variance tradeoff - reduce estimation noise
- Correcting for eigenvalue bias leaves eigenvector bias exposed - corrected by the JSE
- Non-optimized portfolio VFR improved by adjusting the JSE shrinkage constant
- Minimum variance portfolio VFR is always prone to optimization bias
- How to formulate a target function for covariance estimates w.r.t. portfolio risk forecasting?
- Can a single risk model outperform for various portfolios, or do we have to accept a trade-off?