

The Dispersion Bias

Correcting a large source of error in minimum variance portfolios

Lisa Goldberg Alex Papanicolaou Alex Shkolnik

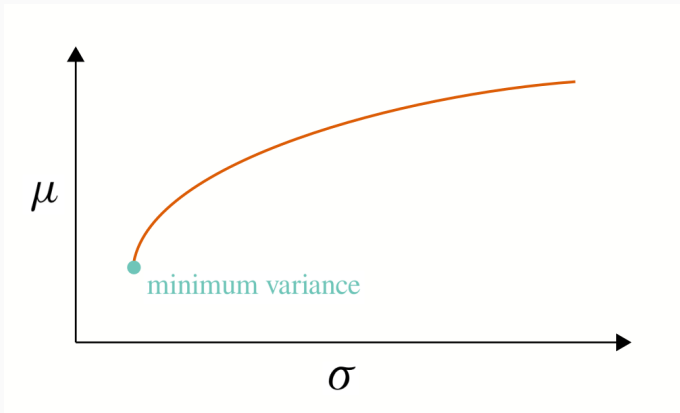
3 November 2017

Swissquote Conference 2017 on FinTech

Optimized portfolios and the impact of estimation error

Optimized portfolios

Since (Markowitz 1952), quantitative investors have constructed portfolios with mean-variance optimization.



The impact of estimation error

In practice, optimization relies on an estimate of the mean and covariance matrix ($\hat{\Sigma}$ estimates Σ).

Estimation error leads to two types of errors:

- **You get the wrong portfolio:** Estimation error distorts portfolio weights so optimized portfolios are never optimal.
- **And it's probably riskier than you think it is:** A risk-minimizing optimization tends to materially underforecast portfolio risk.

We measure both errors in simulation.

Measuring the impact of estimation error in simulation

(Squared) tracking error of an optimized portfolio \hat{w} measures its distance from the optimal portfolio w_* :

$$\mathcal{T}_{\hat{w}}^2 = (\hat{w} - w_*)^\top \Sigma (\hat{w} - w_*)$$

Tracking error is the width of the distribution of return differences between w and \hat{w} . Ideally, tracking error should be as close to 0.

Variance forecast ratio measures the error in the risk forecast as:

$$\mathcal{R}_{\hat{w}} = \frac{\hat{w}^\top \hat{\Sigma} \hat{w}}{\hat{w}^\top \Sigma \hat{w}}$$

Ideally, the variance forecast ratio should be as close to 1.

In simulation,

- generate returns, estimate $\hat{\Sigma}$, compute \hat{w} ;
- compute w_* using Σ (accessible in simulation);
- measure the errors.

$$\mathcal{J}_{\hat{w}}^2 = (\hat{w} - w_*)^\top \Sigma (\hat{w} - w_*)$$

$$\mathcal{R}_{\hat{w}} = \frac{\hat{w}^\top \hat{\Sigma} \hat{w}}{\hat{w}^\top \Sigma \hat{w}}$$

Minimum variance

Why minimum variance?

Theory

Error amplification: Highly sensitive to estimation error.

Error isolation: Impervious to errors in expected return.

Insight into a general problem: Informs our understanding of how estimation error distorts portfolios and points to a remedy.

Practice

Large investments: For example, the Shares Edge MSCI Min Vol USA ETF had net assets of roughly \$14 billion on Sept. 8, 2017.

The true **minimum variance portfolio** w_* is the solution to:

$$\begin{aligned} \min_{x \in \mathbb{R}^N} \quad & x^\top \Sigma x \\ & x^\top \mathbf{1}_N = \mathbf{1}. \end{aligned}$$

In practice, we construct an **estimated minimum variance portfolio**, \hat{w} , that solves the same problem with $\hat{\Sigma}$ replacing Σ .

Some past approaches

Plug-in estimates $\hat{\Sigma}$:

- sample covariance matrix
 - *underforecasts risk by factor $(1 - N/T)_+$*
- covariance regularization
- low-dimensional approximation (factor model)
- bayes/shrinkages estimates (structured model)

Also, bootstrap resampling and stochastic optimization.

Factor models form our starting point.

Factor models

Factor models and equity markets

Beginning with the development of the Capital Asset Pricing Model (CAPM) in (Treynor 1962) and (Sharpe 1964), **factor models** have been central to the analysis of equity markets.

In a **fundamental model**, human analysts identify factors. Fundamental models have been widely used by equity portfolio managers since (Rosenberg 1984) and (Rosenberg 1985).

In a **statistical model** (such as PCA, factor analysis, etc), machines identify factors. An enormous academic literature on PCA models has descended from (Ross 1976).

PCA is the focus of our analysis.

A one-factor model

The return generating process is specified by

$$R = \phi\beta + \epsilon$$

where ϕ is the return to a market factor, β is the N -vector of factor exposures, δ^2 is the N -vector of diversifiable specific returns.

When the ϕ and ϵ are uncorrelated, the security covariance matrix can be expressed as

$$\Sigma = \sigma^2\beta\beta^\top + \Delta,$$

where σ^2 is the variance of the factor and the diagonal entries of Δ are specific variances, δ^2 .

In practice, we have only estimates: $\hat{\sigma}^2$, $\hat{\beta}$ and $\hat{\delta}^2$.

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

We measure the errors in estimated parameters, of course.

But our focus is how errors in parameter estimates affect portfolio metrics: **tracking error** and **variance forecast ratio**.

PCA model estimate (sample covariance \mathbf{S}):

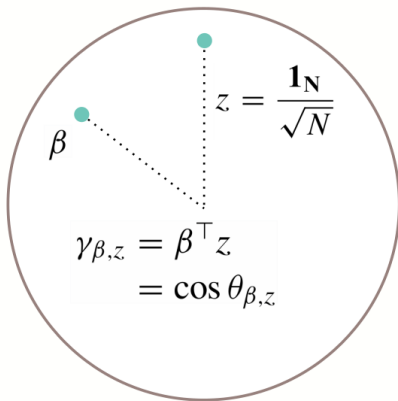
- $\hat{\beta}$ – first eigenvector of \mathbf{S} ,
- $\hat{\sigma}^2$ – largest eigenvalue of \mathbf{S} ,
- $\hat{\delta}^2$ – OLS regression of returns on the estimated factor.

PCA approximates true factors well for N large and Σ .

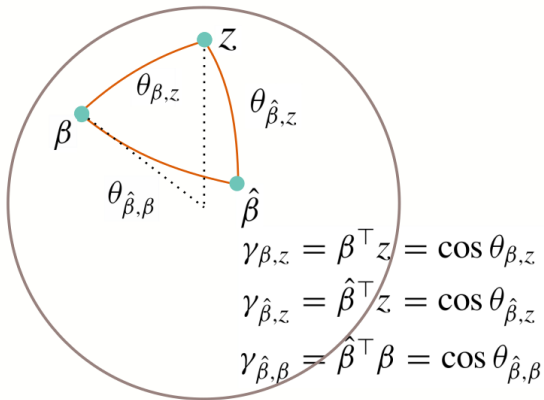
Sample eigenvectors behave differently for N large, T fixed.

Factor models, minimum variance and estimation error

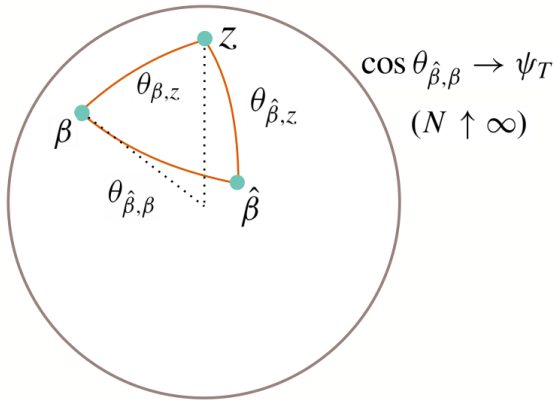
The dispersionless vector



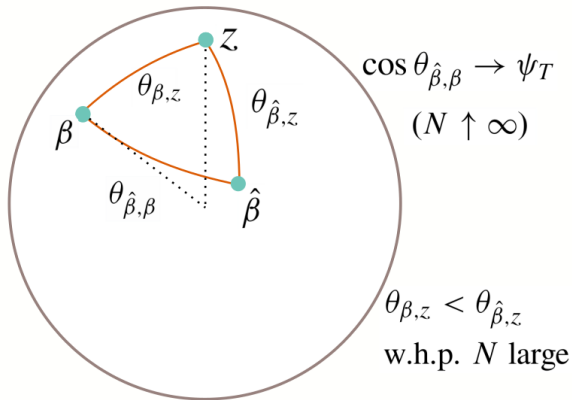
Geometry of the problem



Sample eigenvector behavior



The dispersion bias



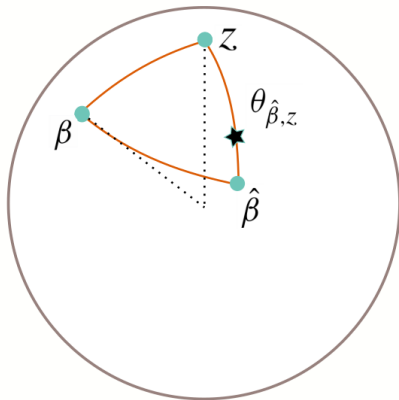
For fixed T and large N , let $r = \frac{\gamma_{\beta,z}}{\gamma_{\hat{\beta},z}}$.

$$\mathcal{T}_{\hat{w}}^2 \asymp \frac{\sigma_N^2}{N} (r - \gamma_{\beta,\hat{\beta}})^2$$
$$\mathcal{R}_{\hat{w}} \asymp \frac{\hat{\delta}^2}{\sigma_N^2 (r - \gamma_{\beta,\hat{\beta}})^2 + \delta^2}$$

PCA estimator has $\mathcal{R}_{\hat{w}} \rightarrow 0$ and $\mathcal{T}_{\hat{w}}^2$ positive as N becomes large.

Decreasing $r - \gamma_{\beta,\hat{\beta}}$ lowers tracking error and raises variance forecast ratio, both desirable. Equivalent to minimizing $\theta_{\beta,\hat{\beta}}$.

Moving in the right direction



Correcting biases

Dispersion bias correction for a standard one-factor model

Almost surely, some shrinkage of $\hat{\beta}$ toward z along the geodesic on the sphere connecting the two points lowers tracking error and raises variance forecast ratio of a minimum variance portfolio.

The oracle estimate of β is given by:

$$\hat{\beta}^* \propto \hat{\beta} + \rho^* z$$

where

$$\rho^* = \frac{\gamma_{\beta,z} - \gamma_{\beta,\hat{\beta}}\gamma_{\hat{\beta},z}}{\gamma_{\beta,\hat{\beta}} - \gamma_{\beta,z}\gamma_{\hat{\beta},z}}.$$

The oracle zeroes out $r - \gamma_{\beta,\hat{\beta}}$, and it has a useful limit as $N \rightarrow \infty$.

Eigenvalue bias correction for a standard one-factor model

For a large N minimum variance portfolio, (bias in) the largest eigenvalue **does not affect tracking error and variance forecast ratio...**

... as predicted by the **Bianchi experiment**, which shows that replacing an estimated eigenvector with a true eigenvector improves portfolio metrics even when the estimated eigenvalue is not corrected.

But eigenvalue correction is important for other portfolios, so we do it.

$$\hat{\sigma}_{\rho^*}^2 = \left(\frac{\gamma_{\hat{\beta},z}}{\gamma_{\hat{\beta}^*,z}} \right)^2 \hat{\sigma}^2$$

Impact of bias correction theorem for a standard one-factor model

For fixed T with $N \rightarrow \infty$

Model	Tracking error	Variance forecast ratio
PCA	bounded away from 0	$\rightarrow 0$
Oracle	$\rightarrow 0$	$\rightarrow 1$
Target	0	1

A bonafide shrinkage estimator

Our estimate begins with the asymptotic (large N) formula for the oracle estimator

$$\rho^* = \frac{\gamma_{\beta,z}}{1 - \gamma_{\beta,z}^2} (\Psi_T - \Psi_T^{-1}),$$

where Ψ_T is a positive random variable that is expressed in terms of χ_T and asymptotic estimates for eigenvalues.

We rely on (Yata & Aoshima 2012) for the latter.

The emphasis of the fixed T large N regime dates back to (Connor & Korajczyk 1986) and (Connor & Korajczyk 1988).

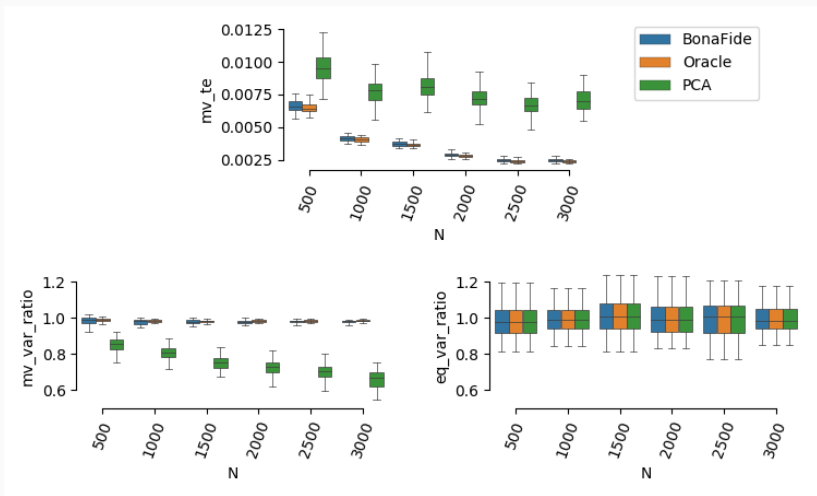
Our proofs rely heavily on (Shen, Shen, Zhu & Marron 2016) and (Wang & Fan 2017).

Numerical results

Calibrating the one-factor model

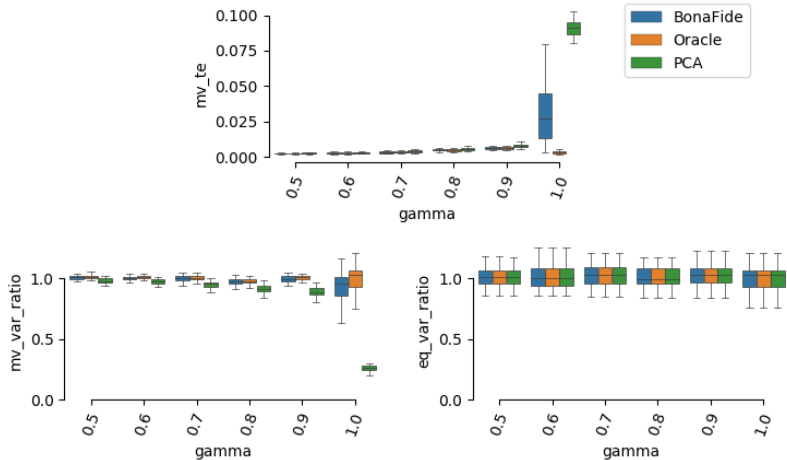
Parameter	Value	Comment
β	normalized so $ \beta _2 = 1$	factor exposure
$\gamma_{\beta,z}$	0.5–1.0	controls dominant factor dispersion
σ^2	dominant eigenvalue of Σ	annualized factor volatility of 16%
δ^2	specific variances	annualized specific volatilities drawn from [10%, 64%]

Numerical results from a one-factor model, $\gamma_\beta = 0.90$



Simulation based on 50 samples

Numerical results from a one-factor model, $N = 500$



Simulation based on 50 samples

**Beta shrinkage has been used by
practitioners since the 1970s**

The innovators

In the 1970s, Oldrich Vasicek and Marshall Blume observed **excess dispersion** in betas estimated from time series regressions, and they proposed adjustments.

(**Vasicek 1973**) shrinks estimated betas toward their cross-sectional mean using a Bayesian formula.

(**Blume 1975**) uses the empirically observed average shrinkage of betas on individual stocks in the current period relative to a previous period to adjust forecast betas for the next period.

An ultra-simplified version of the Blume adjustment is on the exam taken by aspiring Chartered Financial Analysts (CFA)s.

Quizlet

CFA Level 2 Equity

Multifactor Models		
Build-up Model	$r = R_F + \text{equity risk premium} + \text{size premium} + \text{specific-company premium}$	☆
Blume Adjustment	$\text{adjusted beta} = (2/3 \text{ raw beta} + 1/3)$	☆
WACC	$\text{MV debt} / \text{Total Assets} \cdot (1-T) r_d + \text{MV equity} / \text{Total Assets} \cdot r_e$	☆

Chat about Blume on AnalystForum.com



chowder (/profile/chowder)

Jun 7th, 2009 2:35pm

205 AF Points

(/profile/chowder)

wtf is Blume method?



bpdulog (/profile/bpdulog)

Jun 7th, 2009 2:36pm

CFA Charterholder

3,696 AF Points

(/profile/bpdulog)

I don't know but this is what I did:

$2/3^*$ provided beta + .33

I think I read that in the Secret Sauce



topher (/profile/topher)

Jun 7th, 2009 2:36pm

1,649 AF Points

(/profile/topher)


I didn't know what Blume was either, but they asked for adjusted beta which is:

$(2/3)\text{Beta} + 1/3*1 = 1.27$

 (/comment/91005043#comment-91005043)

Covariance Shrinkage methods

A flat market mode acknowledged as a target by Bouchaud, Bun, & Potters.



Rotational invariance hypothesis (RIH)

- In the absence of any *cogent prior* on the eigenvectors of \mathbf{C} , one can assume that the estimator $\hat{\mathbf{C}}(\mathbf{E})$ is a member of a *Rotationally Invariant Ensemble* (RIH)
- Remark: certainly not true in Finance, at least for the largest eigenvalue (“*market mode*”) because

$$\mathbf{V}_1 \approx (1, 1, \dots, 1) / \sqrt{N}$$

but it is OK for the bulk.

- “Cleaning” \mathbf{E} within RIH: *keep the eigenvectors, play with eigenvalues* to find:

$$\hat{\mathbf{C}}(\mathbf{E}) = \mathbf{U} \hat{\Lambda} \mathbf{U}^\dagger$$

Marc Potters (Capital Fund Management)

4 / 22

Honey, I shrunk the sample covariance matrix

Seminal work (Ledoit & Wolf 2004) introduced a linear shrinkage correction to the sample covariance utilizing a constant correlation target

$$C = \alpha E + (1 - \alpha)[(1 - \rho)I + \rho ee^T].$$

Our vector z represents a constant correlation mode so there are parallels.

We also take aim at a specific misbehaving artifact in factor models and minimum variance portfolios and leverage a different asymptotic theory.

Summary and ongoing research

We identified a **large, damaging dispersion bias** in PCA-estimated factor models.

We determined an **oracle correction** that elevates portfolio construction and risk forecasting for a minimum variance portfolios for fixed T as $N \rightarrow \infty$.

And we have developed a **bona fide (data-driven) correction** that has proven effective in empirically-calibrated simulation.

Our results can be viewed as **an extension and formalization** of ideas that have been known by practitioners since the 1970s.

Extension to multi-factor models.

Application to sparse low rank factor extractions.

Investigation of a wider class of portfolios.

Empirical studies.

Thank you



Photograph by Jim Block

- Blume, Marshall E. (1975), 'Betas and their regression tendencies', *The Journal of Finance* **30**(3), 785–795.
- Connor, Gregory & Robert A. Korajczyk (1986), 'Performance measurement with the arbitrage pricing theory: A new framework for analysis', *Journal of financial economics* **15**, 373–394.
- Connor, Gregory & Robert A. Korajczyk (1988), 'Risk and return in equilibrium apt: Application of a new test methodology', *Journal of financial economics* **21**, 255–289.
- Ledoit, Oliver & Michael Wolf (2004), 'Honey, i shrunk the sample covariance matrix', *The Journal of Portfolio Management* **30**, 110–119.

- Markowitz, Harry (1952), 'Portfolio selection', *The Journal of Finance* **7**(1), 77–91.
- Rosenberg, Barr (1984), 'Prediction of common stock investment risk', *The Journal of Portfolio Management* **11**(1), 44–53.
- Rosenberg, Barr (1985), 'Prediction of common stock betas', *The Journal of Portfolio Management* **11**(2), 5–14.
- Ross, Stephen A (1976), 'The arbitrage theory of capital asset pricing', *Journal of economic theory* **13**(3), 341–360.
- Sharpe, William F (1964), 'Capital asset prices: A theory of market equilibrium under conditions of risk', *The Journal of Finance* **19**(3), 425–442.

- Shen, Dan, Haipeng Shen, Hongtu Zhu & Steve Marron (2016), 'The statistics and mathematics of high dimensional low sample size asymptotics', *Statistica Sinica* **26**(4), 1747–1770.
- Treynor, Jack L (1962), Toward a theory of market value of risky assets. Presented to the MIT Finance Faculty Seminar.
- Vasicek, Oldrich A. (1973), 'A note on using cross-sectional information in bayesian estimation of security betas', *The Journal of Finance* **28**(5), 1233–1239.
- Wang, Weichen & Jianqing Fan (2017), 'Asymptotics of empirical eigenstructure for high dimensional spiked covariance', *The Annals of Statistics* **45**(3), 1342–1374.

Yata, Kazuyoshi & Makoto Aoshima (2012), 'Effective pca for high-dimension, low-sample-size data with noise reduction via geometric representations', *Journal of multivariate analysis* **105**(1), 193–215.