Portfolio Selection Revisited
in memory of Harry Markowitz

Lisa Goldberg(lrg@berkeley.edu)
in collaboration with Haim Bar, Hubeyb Gurdogan, Alec Kercheval and Alex Shkolnik

Blessings of Dimensionality Workshop
17 July, 2024

You are a junior research assistant to a trader, who asks you to quickly estimate the risk of a minimum variance portfolio in the S&P 500 with a year's worth of daily data

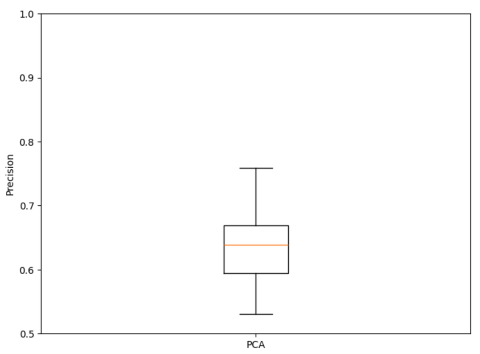
$$\min_{w \in \mathbb{R}^p} w^\top \Sigma w$$
$$w^\top \mathbf{1} = 1$$

$$w_* = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}}$$

$$V^2(w_*) = \frac{1}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}}$$

How good is your estimate?

The estimate is imprecise

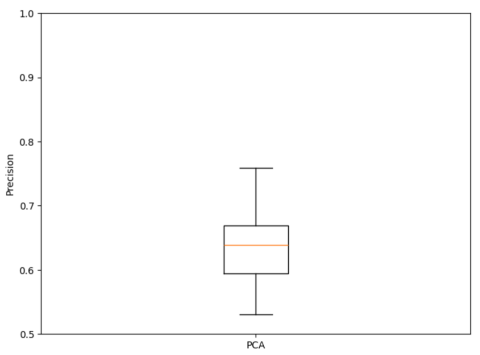


Precision

$$\mathcal{P} = \frac{\text{estimated volatility}}{\text{true volatility}}$$

Data are generated from a one factor model, $r = \beta f + \epsilon$. Factor returns f are drawn independently from $N(0, 0.16^2)$ and specific returns are drawn independently from $N(0, 0.50)$. While values of β are parameters to be estimated, they are drawn from $N(1, 0.5)$. Boxplot generated with 200 simulated paths. Analysis and graphics by Rahul Vinoth.

Can we rescue the situation with fancy factor models, random matrix theory and convex optimization?



Data are generated from a one factor model, $r = \beta f + \epsilon$. Factor returns f are drawn independently from $N(0, 0.16^2)$ and specific returns are drawn independently from $N(0, 0.50)$. While values of β are parameters to be estimated, they are drawn from $N(1, 0.5)$. Boxplot generated with 200 simulated paths. Analysis and graphics by Rahul Vinoth.

The Bianchi experiment

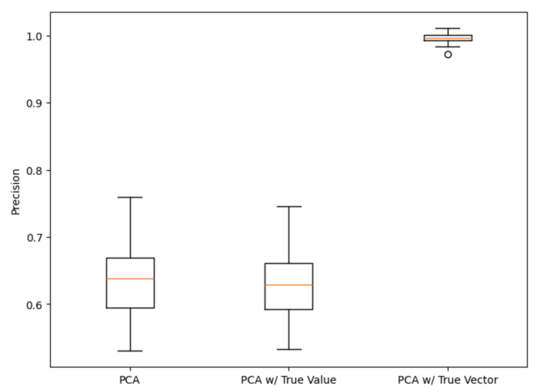
In simulation:

Replace the estimated eigenvalue with the truth, leaving the corrupted eigenvector in place

Switch the roles of eigenvalue and eigenvector in the previous step

Data are generated from a one factor model, $r = \beta f + \epsilon$. Factor returns f are drawn independently from $N(0, 0.16^2)$ and specific returns are drawn independently from $N(0, 0.50)$. While values of β are parameters to be estimated, they are drawn from $N(1, 0.5)$. Boxplot generated with 200 simulated paths. Analysis and graphics by Rahul Vinoth.

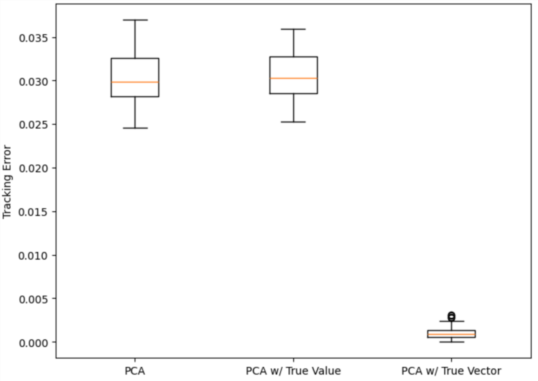
It appears to be errors in eigenvectors that lower precision



As we improve precision by correcting eigenvectors, we also move the optimized portfolio closer to the optimum

Tracking error, the workhorse of financial services, is the distance between two portfolios

$$TE^2 = (w_* - \mathbf{w})^\top \mathbf{\Sigma} (w_* - \mathbf{w})$$



Data are generated from a one factor model, $r = \beta f + \epsilon$. Factor returns f are drawn independently from $N(0, 0.16^2)$ and specific returns are drawn independently from $N(0, 0.50^2)$. While values of β are parameters to be estimated, they are drawn from $N(1, 0.5)$. Analysis and graphics by Rahul Vinoth.

How do we correct eigenvectors when we don't know ground truth?

We turned to the practitioner literature, notably, a 2011 paper by Roger Clarke, Harindra deSilva and Steven Thorley, which explains how to write the weights of a minimum variance portfolio in terms of the parameters of a 1-factor model:

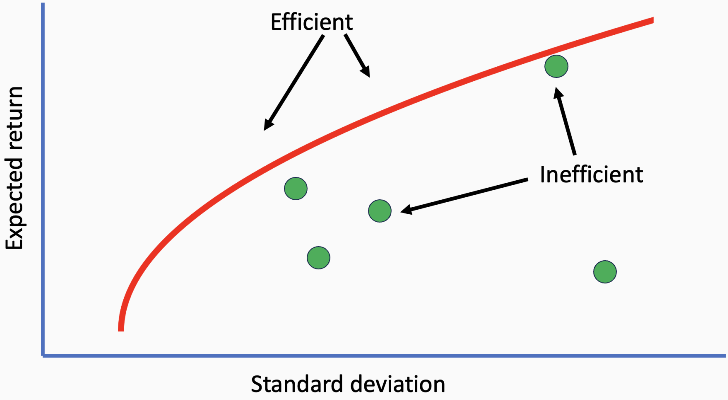
$$w_* = \frac{\sigma_{MV}^2}{\delta^2}\left(\mathbf{1} - \frac{1}{\beta_{LS}}\beta\right)$$

we applied the Woodbury formula and took limits

In the beginning....

In 1952, Harry Markowitz launched modern finance by framing it as a tradeoff between expected return and variance



Markowitz did not originally draw the efficient frontier the way we do today

Markowitz's mean-variance optimization is SOP in financial services today

An optimized portfolio is a solution to a quadratic optimization with linear constraints. More generally, consider a $p \times k$ matrix of constraint gradients C and k-vector of constraint targets T:

$$\min_{w \in \mathbb{R}^p} w^\top \Sigma w$$
$$w^\top C = T$$

The prototypical optimized portfolio specified by:

$$C = \begin{pmatrix} 1 & m_1 \\ \vdots & \vdots \\ 1 & m_p \end{pmatrix} \qquad T = \begin{pmatrix} 1 \\ m \end{pmatrix}$$

where the m_i's are security expected returns and m is a target portfolio expected return.

There is a closed-form expression for an optimized portfolio, but estimation error is everywhere

The solution is given by:

$$w_* = \gamma_{\mathbf{1}} \Sigma^{-1} \mathbf{1} + \gamma_m \Sigma^{-1} m$$

a combination of characteristic (one-constraint) portfolios, weighted by their shadow prices (Lagrangians)

characteristics are $\mathbf{1}$ and m

everything not in bold (everything except $\mathbf{1}$) is corrupted by estimation error

$$a = \mathbf{1}^\top \Sigma^{-1} \mathbf{1}, b = \mathbf{1}^\top \Sigma^{-1} m, d = m^\top \Sigma^{-1} m, \gamma_{\mathbf{1}} = \frac{d - bT}{ad - b^2}, \gamma_m = \frac{aT - b}{ad - b^2}.$$

The Markowitz enigma, factor models and random matrix theory

Evidently realizing that classical statistics would not provide what he needed, Markowitz considered alternative ways to estimate optimization inputs in 1952

Perhaps there are ways, by combining statistical techniques and the judgment of experts, to form reasonable probability beliefs μ_{ij}, σ_{ij} One suggestion as to tentative μ_{ij}, σ_{ij} is to use the observed μ_{ij}, σ_{ij} for some period of the past. I believe that better methods, which take into account more information, can be found. I believe that what is needed is essentially a "probabilistic" reformulation of security analysis. I will not pursue this subject here, for this is "another story." It is a story of which I have read only the first page of the first chapter.

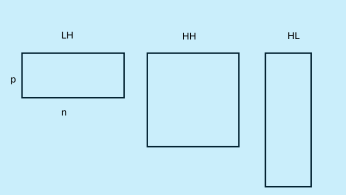
Harry Markowitz (1952)

Portfolios optimized with estimates are not efficient—In 1989 Richard Michaud wrote that mean-variance optimizers are "estimation error maximizers"

The precision of

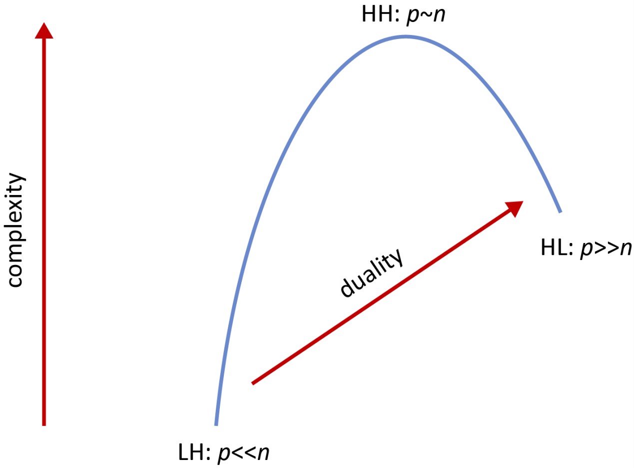
$$w_* = \gamma_{\mathbf{1}} \Sigma^{-1} \mathbf{1} + \gamma_m \Sigma^{-1} m$$

depends on the covariance matrix Σ and means m, which are estimated from data



Optimization selectively amplifies errors in an estimated covariance matrix and the problem is severe when the number of securities p is large relative to the observations n

We work in the HL regime of random matrix theory, where we benefit from duality and concentration of measure, a blessing of dimensionality



In HL, a sample covariance matrix S cannot be used in high-dimensional optimization, but still contains useful information

$$S = \begin{pmatrix} & \vdots & \\ \dots & \sigma_{ij} \dots & \\ & \vdots & \end{pmatrix}$$

$$\begin{aligned} \sigma_{ij} &= \sigma_i \sigma_j \rho_{ij} \\ &= \mathrm{vol} \cdot \mathrm{vol} \cdot \mathrm{corr} \end{aligned}$$

Spectral decomposition

$$S = \lambda^2 h h^\top + \dots.$$

$$\ell^2 = \frac{Tr(S) - \lambda^2}{n-1}$$

Factor models reduce dimension and are widely used in finance

Correlations across many variables are driven by a relatively small number of factors.

1904: Charles Spearman's two-factor theory of intelligence
1963: Bill Sharpe's one-factor market model
1974: Barr Rosenberg's Barra
1976: Stephen Ross's arbitrage pricing theory

Suppose returns to p securities follow a one-factor model

$$r = \beta f + \epsilon$$

Then:

$$\Sigma = \sigma^2 \beta\beta^\top + \delta^2 I = \eta^2 b b^\top + \delta^2 I$$

We need to estimate two positive scalars, η^2 and δ^2, and a unit p-vector b.

Only r is observed. $r \in \mathbb{R}^p$, $f \in \mathbb{R}$ and $\epsilon \in \mathbb{R}^p$ are random (but not necessarily Gaussian) and $\beta \in \mathbb{R}^p$ is an unknown parameter. $\sigma^2 = \mathrm{var}(f)$, $\delta^2 = \mathrm{var}(\epsilon)$, $\eta^2 = \sigma^2 |\beta|^2$ and $\mathrm{corr}(f, \epsilon) = 0$. Returns of public equities tend to be positively correlated, so the leading eigenvalue of a sample security return covariance matrix tends to grow approximately linearly with p and the entries of the leading eigenvector tend to have a common sign.

Spectral analysis of the sample covariance matrix suggests a factor-based estimate of Σ

For a unit p vector , set:

$$\Sigma^{v} = (\lambda^2 - \ell^2) v v^\top + \frac{n}{p}\ell^2 I,$$

The estimate Σ^v varies with the unit vector v, and a standard choice is the leading sample eigenvector h of the sample covariance matrix S.

Since $\lim_{p\to\infty}(\lambda^2 - \ell^2)/p = \lim_{p\to\infty} \eta^2/p$ is finite and $\lim_{p\to\infty} n\ell^2/p = \delta^2$, we keep the estimates of η^2 and δ^2 fixed. For any v, $\mathrm{Trace}(\Sigma^v) = \mathrm{Trace}(S)$ is an unbiased estimate of $\mathrm{Trace}(\Sigma)$.

An outsider's perspective

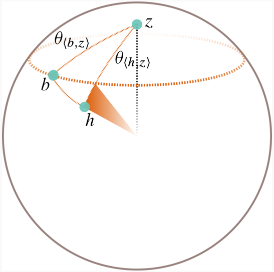
Concentration of measure implies that when $p >> n$, the sample leading eigenvector h is bound to be further away from an anchor point z than the population leading eigenvector b

In a suitable spiked factor model, almost surely as $p \to \infty$,

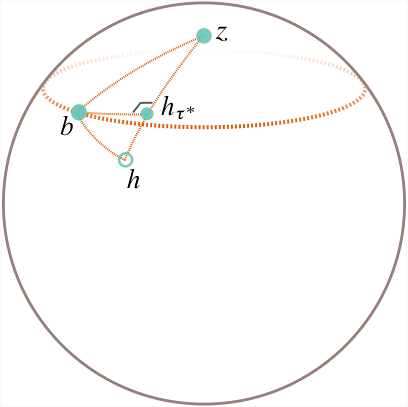
$$\langle h, z \rangle = \langle h, b \rangle \langle b, z \rangle$$

with

$$\langle h, b \rangle = \frac{\lambda^2 - \ell^2}{\lambda^2}$$

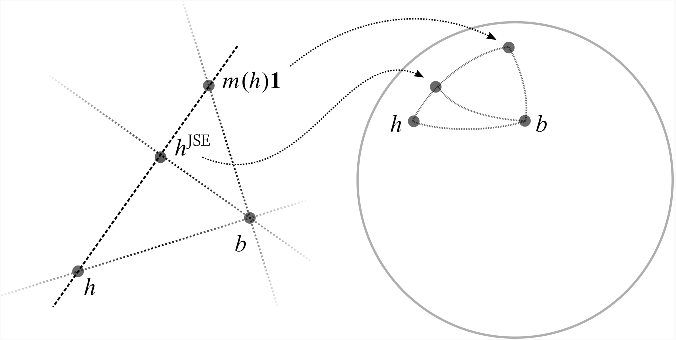


Applying the right amount of shrinkage of h toward the anchor point z yields a better estimate of b



source: Goldberg, Papanicalaou & Shkolnik (2022)

The shrinkage is given by a James-Stein type formula



$$h_C = \underset{C}{proj}(h),\, c^{JSE} = \frac{\ell^2/\lambda^2}{1-|h_C|^2},\, h^{\mathrm{JSE}} \propto (1-c^{JSE})h + c^{JSE} h_C$$

$$C = k\text{−dimensional subpace of}\, R^p,\, \angle(b, C) < \pi/2$$

source: Goldberg & Kercheval (2023), drawing by Alex Shkolnik

Portfolio selection revisited

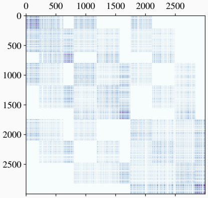
Investors use risk-adjusted excess return, or Sharpe ratio, to distinguish portfolios

$$SR = \frac{r - r_f}{\sigma}$$

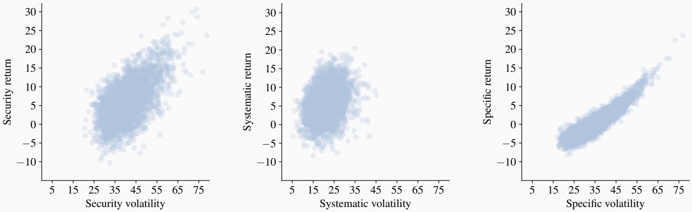
Some investors select the portfolio with highest Sharpe ratio (or information ratio) in a backtest

We simulate data with a seven-factor return generating process

factor	count
market	1
styles	2
industries	4

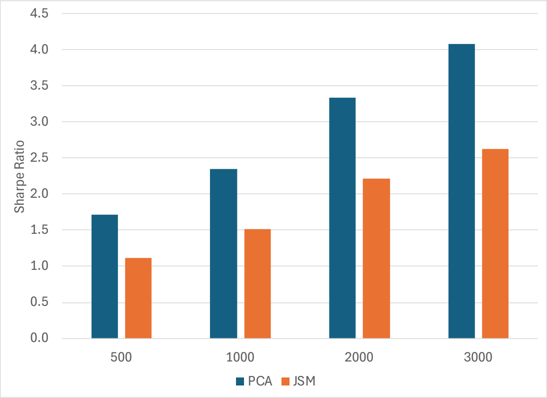


Industry membership visualization



Complete calibration details are in Appendix A of Shkolnik, Kercheval, Gurdogan, Goldberg & Bar (2024).

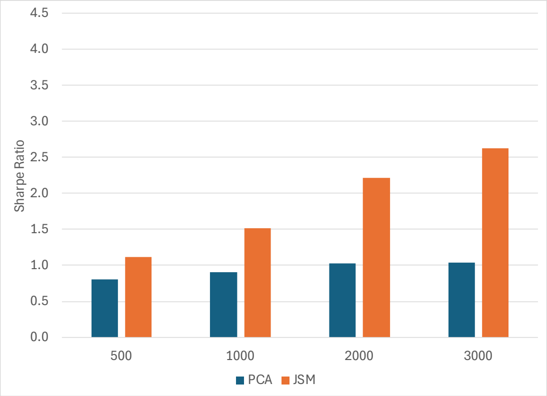
On the basis of estimated Sharpe Ratios, the investor chooses PCA



Estimated

Data are generated from the seven-factor model in Shkolnik et al. (2024). Values are averages over 400 simulated

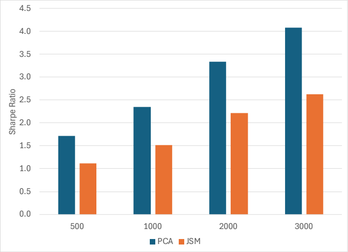
When, in fact, JSM has higher true Sharpe Ratios



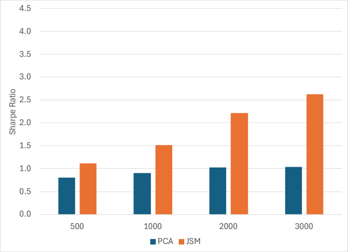
True

Data are generated from the seven-factor model in Shkolnik et al. (2024). Vales are averages over 400 simulated paths.

An investor can see estimates but not ground truth



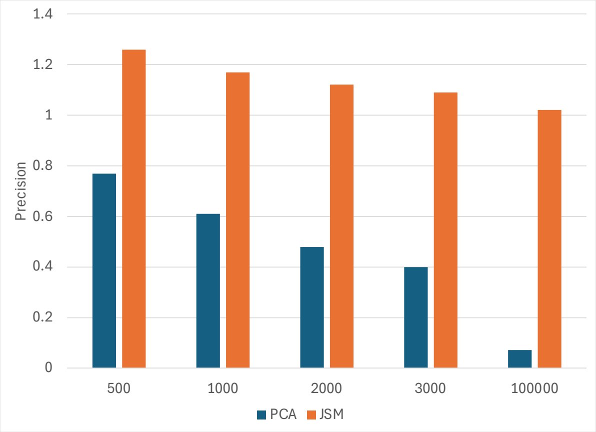
Estimated



True

Data are generated from the seven-factor model in **?**. Values are averages over 400 simulated paths.

It is low precision of volatility estimates that severely distort Sharpe Ratios for PCA



Data are generated from the seven-factor model in Shkolnik et al. (2024). Values are averages over 400 simulated paths.

Final thoughts

Practical problems can generate new theory

The challenge of creating an estimated security return covariance matrix suitable for optimization posed by Harry Markowitz more than 70 years ago has been approached in countless ways.

Using concentration of measure arguments, which we access via an outsider's perspective, our work on this issue has generated:

- Data-driven methods for:
  - correcting sample eigenvectors of in the HL regime
  - mitigating the transmission of estimation error in a PCA-based covariance matrix via quadratic optimization
  - increasing the precision of optimized portfolios
- Some clarity on asymptotic regimes of random matrix theory

Harry Markowitz inspires us to look at stubborn problems with fresh eyes

At Markowitz's thesis defense, committee member Milton Friedman told him:

> *Harry, I read your dissertation. I don't see any problems with the math, but this is not a dissertation in economics. We can't give you a Ph.D. in economics for a dissertation that isn't about economics.*

Today, mathematical finance is central to economics.

Collaborators

Jeongyoun Ahn
Haim Bar
Jacob Biernat
Stjepan Begušić
Hubeyb Gurdogan
Sungkyu Jung
Alec Kercheval
Jacob Lan
Youhong Lee
Alex Papanicolaou
Harrison Selwitz
Alex Shkolnik
Simge Ulucam
Rahul Vinoth

References

Clarke, R., deSilva, H. & Thorley, S. (2011), 'Minimum-variance portfolio composition', *The Journal of Portfolio Management* **37**(2), 31–45.

Goldberg, L. R., Gurdogan, H. & Kercheval, A. (2024), Portfolio optimization via strategy-specific eigenvector shrinkage.

Goldberg, L. R. & Kercheval, A. N. (2023), 'James Stein for the leading eigenvector', *Proceedings of the National Academy of Sciences* **120**.

Goldberg, L. R., Papanicalaou, A. & Shkolnik, A. (2022), 'The dispersion bias', *SIAM Journal on Financial Mathematics* **13**(2), 521–550.

Goldberg, L. R., Papanicolaou, A., Shkolnik, A. & Ulucam, S. (2020), 'Better betas', *The Journal of Portfolio Management* **47**(1), 119–136.

Gurdogan, H. & Kercheval, A. (2022), 'Multi anchor point shrinkage for the sample covariance matrix', *SIAM Journal on Financial Mathematics* **13**(3), 1112–1143.

Gurdogan, H. & Shkolnik, A. (2024), The quadratic optimization bias of large covariance matrices.

Markowitz, H. (1952), 'Portfolio selection', *Journal of Finance* **7**(1), 77–91.

Michaud, R. (45), 'The Markowitz optimization enigma: Is 'optimized' optimal?', *Financial Analysts Journal* **1**(31–42).

Rosenberg, B. & McKibben, W. (1973), 'The prediction of systematic and specific risk in common stocks', *Journal of Financial and Quantitative Analysis* **8**(2), 317–333.

Ross, S. (1976), 'The arbitrage theory of capital asset pricing', *Journal of Economic Thoery* **13**, 341–360.

Sharpe, W. F. (1963), 'A simplified model for portfolio analysis', *Management Science* **9**(2), 277–293.

Shkolnik, A. (2022), 'James-Stein estimation of the first principal component', *Stat* **11**(1).

Shkolnik, A., Kercheval, A., Gurdogan, H., Goldberg, L. R. & Bar, H. (2024), Portfolio selection revisited.

Spearman, C. (1904), 'General intelligence, objectively determined and measured', *American Journal of Psychology* **15**(201–293).