MATHEMATICAL ANALYSIS OF COVARIANCE MATRICES IN INDUSTRY BETAS AND VOLATILITES

Presented by Ethan Cox | Mathematics and Computer Science

Iowa State's 19th Annual Symposium on Undergraduate Research and Creative Expression

OVERVIEW

- Abstract
- Introduction
- Problem
- Objectives
- Methodology
- Result
- Conclusion
- Reference

ABSTRACT

This project investigates the mathematical structure of covariance matrices derived from industry betas and volatilities. Using data from 49 U.S. industry portfolios, we compute monthly CAPM betas and volatilities from daily returns, and apply Principal Component Analysis (PCA) for dimensionality reduction. We explore both the estimation and structure of high-dimensional covariance matrices and analyze the explanatory power of their principal components. Results uncover dominant risk factors and the interdependencies between industries, offering insights into systematic risk and practical tools for financial modeling, particularly in high-dimensional settings.

INTRODUCTION

- Covariance matrices are central in portfolio theory and financial risk management.
- Industry betas and volatilities are key indicators of market sensitivity and total risk.
- In high-dimensional settings (many industries/assets), estimation becomes a challenge due to noise and instability.
- My study applied PCA (Principal Component Analysis) and shrinkage techniques to reveal structure and reduce dimensionality, improving reliability and interpretability.

PROBLEM

- How can we mathematically characterize and improve the estimation of covariance matrices derived from industry betas and volatilites?

- In high-dimensional financial data:
    - Covariance estimates are noisy.
    - Traditional methods may fail to capture structure or produce unstable results
- There is a need to reduce noise and extract meaningful patterns from these matrices.

OBJECTIVE

1. Estimate monthly CAPM betas and volatilities for 49 U.S. industry portfolios over a long historical period (1926-2024).
2. Apply Principal Component Analysis to:
   - Understand common drivers of risk.
   - Reduce dimensionality.
   - Quantify variance explained by principal components.
3. Construct correlation matrices of principal components to reveal structural links between beta and volatility factors.
4. Experiment with shrinkage estimators to improve the stability and predictive quality of beta estimates.

METHODOLOGY

Data Collection

- Daily returns for 49 U.S. industry portfolios (Kenneth French data library).
- Daily market returns (Fama-French factors).
- Project completed in Python.

Beta Estimation

- Compute monthly CAPM betas using rolling regressions on daily data.

Volatility Estimation

- Compute monthly volatilities as the standard deviation of daily returns x 100.

METHODOLOGY

PCA Analysis

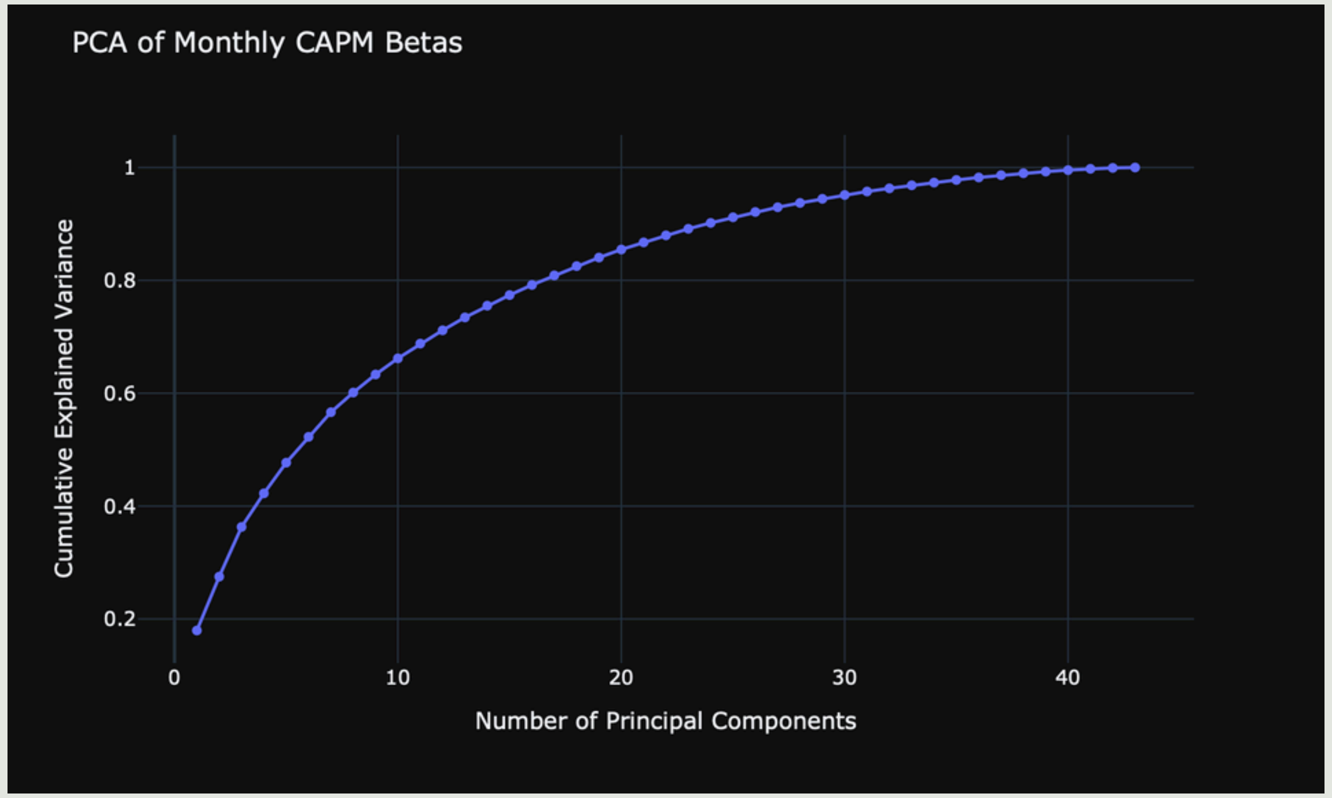
- Apply PCA to the beta matrix and the volatility matrix.
- Extract eigenvalues and principal components.
- Analyze explained variance and factor loadings.

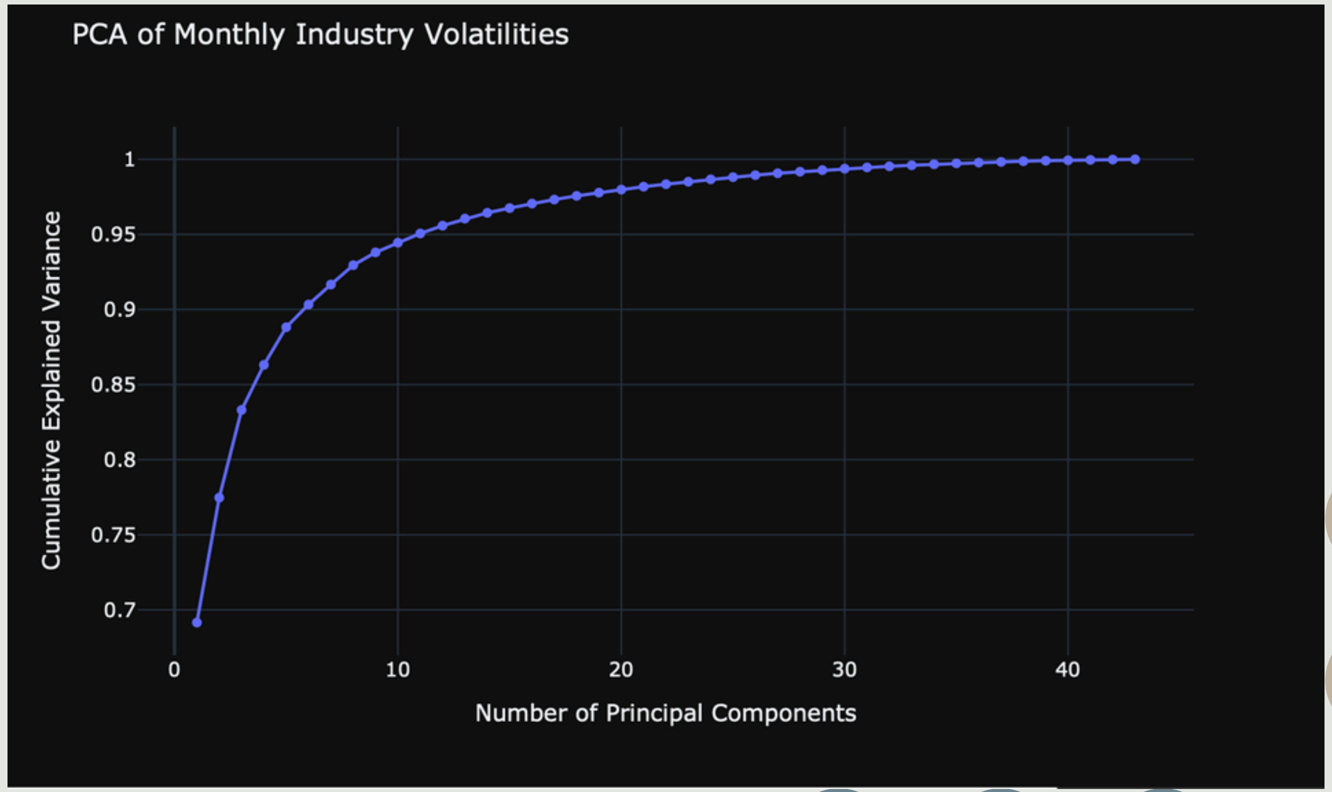
Shrinkage Estimator

- Use rolling windows and PCA to form historical trend-based shrinkage estimates of betas.

RESULT

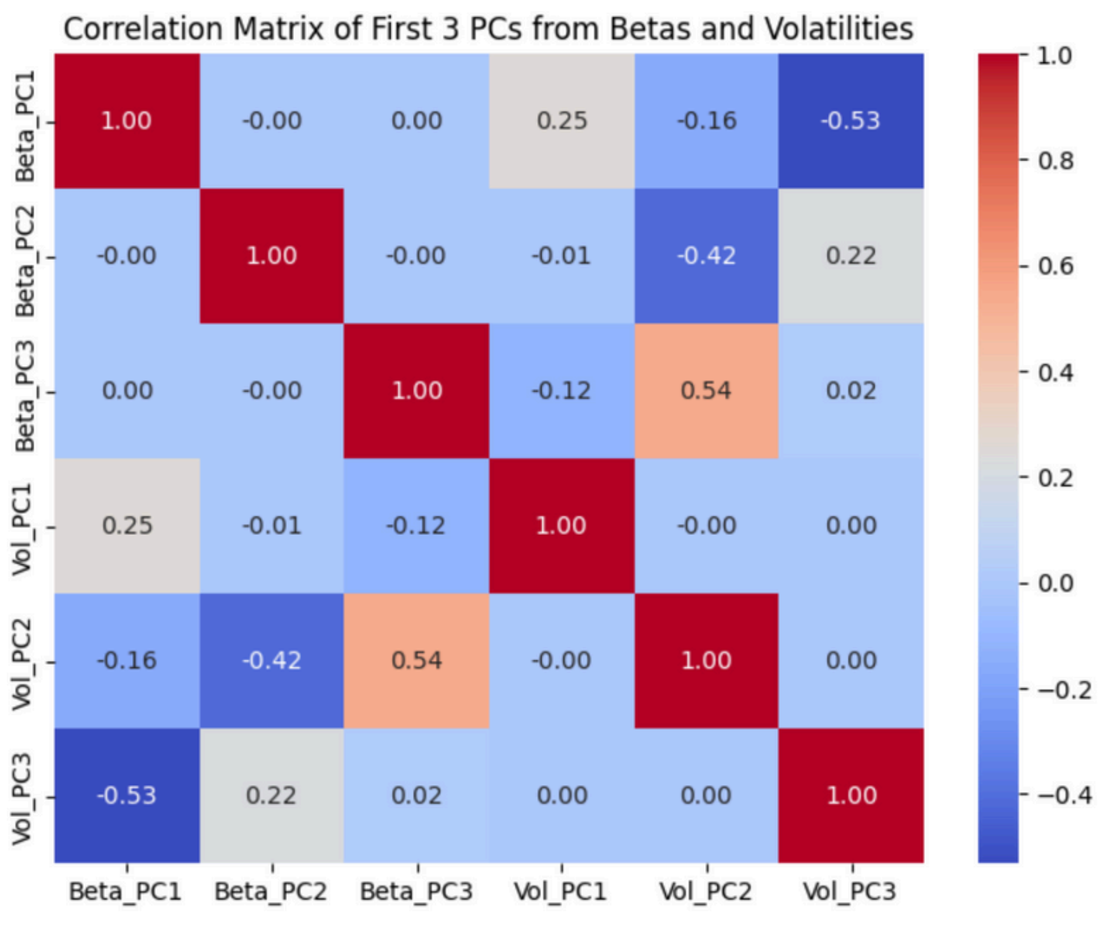
- PCA on Indsutry Betas:
  - The first 2-3 principal components explain a significant portion (~70%) of the total variation.
  - Reveals strong factor structure: one or two common risk sources dominate.
- PCA on Volatilities:
  - A similar structure was observed, with the first few PCs capturing systemic volatility trends.
  - However, these volatility PCs revealed an even stronger factor structure.

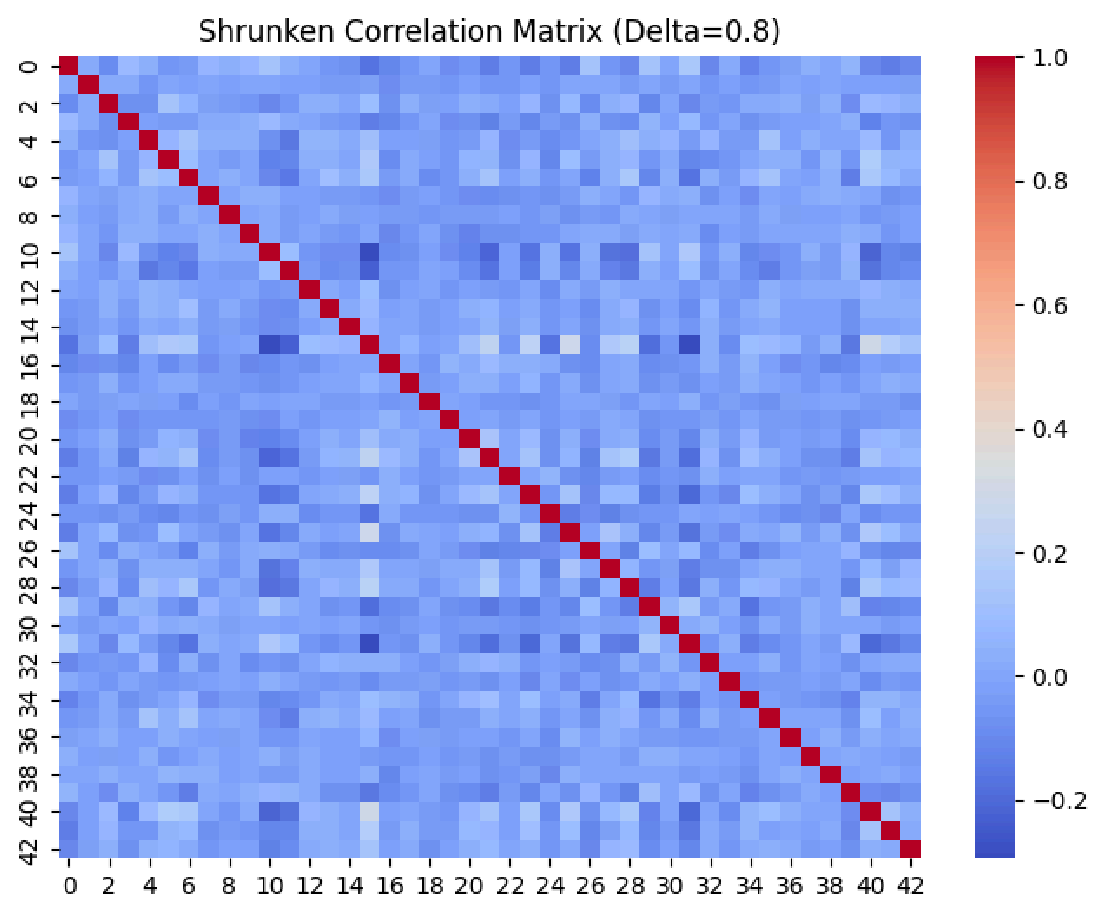




RESULT

- Correlation Matrix of First 3 PCs (Betas vs Volatilities):
  - Indicates interdependencies between systematic risk (beta) and total risk (volatility).
- Shrinkage Estimator Insight:
  - Combines current-month beta with PCA-derived historical trend, reducing estimation noise.
  - Provides more stable and reliable risk profiles across industries.





CONCLUSION

- Covariance matrices of industry betas and volatilities exhibit a strong low-rank structure.
- PCA is mostly effective in uncovering latent risk factors and reducing dimensionality.
- Principal components reveal key patterns in how industries co-move and respond to market changes.

- Shrinkage estimators, guided by PCA, offer a promising method for enhancing stability in high-dimensional beta estimation.
- These findings contribute to better portfolio construction, risk management, and financial forecasting in high-dimensional asset spaces.

REFERENCE

Kenneth French Data Library
https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

“Honey, I Shrunk the Sample Covariance Matrix” - Ledoit-Wolf
http://www.ledoit.net/honey.pdf

“The Capital Asset Pricing Model: Theory and Evidence” - Fama-French
https://mba.tuck.dartmouth.edu/bespeneckbo/default/AFA611-Eckbo%20web%20site/AFA611-S6B-FamaFrench-CAPM-JEP04.pdf

Thank You

And a special thanks to Alberto Martin-Utrera for his guidance.