Correlation Matrices and the Perron-Frobenius Theorem

Phelim Boyle

Wilfrid Laurier University

July 14 2014

Acknowledgements

Thanks to David Melkuev, Johnew Zhang, Bill Zhang, Ronnnie Feng and Jeyan Thangaraj for research assistance.

Outline

- Introduction and background
- The *Perron-Frobenius* theorem and extensions to negative elements
- Analysis of empirical data
- Theoretical results
- Numerical investigations
- Summary and conclusions

1 Introduction and background

Introduction

This is preliminary work. Comments welcome.

- Markowitz efficient portfolios selected by investors
- These portfolios have *desirable* properties
- Mean variance efficient
- Sharpe used these ideas to develop the CAPM
- Equilibrium model relating expected return to risk
- **Market portfolio is mean variance efficient**

Compatible $\boldsymbol{\Sigma}, x^{(m)}, \boldsymbol{\mu}$

- $\boldsymbol{\Sigma}$ covariance matrix; $\boldsymbol{\mu}$ return vector, $\boldsymbol{x}^{(m)}$ market weights
- These three entities have to be compatible since $\boldsymbol{x}^{(m)}$ on frontier
- Best and Grauer (1985)

$$\boldsymbol{\Sigma}\boldsymbol{x}^{(m)} = \gamma_1 \boldsymbol{\mu} + \gamma_2 \boldsymbol{e} \tag{1}$$

- Assume $\boldsymbol{\Sigma}$ known
- Assume- for now- we have a way to find $\boldsymbol{x}^{(m)}$
- Task is to find compatible $\boldsymbol{\mu}$

Picking the market portfolio

- Origins of idea from Sharpe and Ross (APT)
- Dominant common factor that influences stock returns
- PCA used to identify this factor
- Principal eigenvector of the *correlation matrix*
- Trzcinka (1986), Laloux et al(1999), Avellaneda and Lee (2010) & Allez and Bouchaud (2011)
- Market portfolio should have positive weights
- When will principal eigenvector have positive weights?

1 Introduction and background

2 Perron Frobenius

The Perron-Frobenius Theorem

Theorem (Perron-Frobenius)

A real $n \times n$ matrix, $\boldsymbol{A}$, with positive entries has a unique largest real eigenvalue and the corresponding eigenvector has strictly positive components.

- Provides sufficient conditions
- Result can be weakened
- Matrix $\boldsymbol{A}$ can *have some negative elements* and retain the Perron-Frobenius (PF) property.

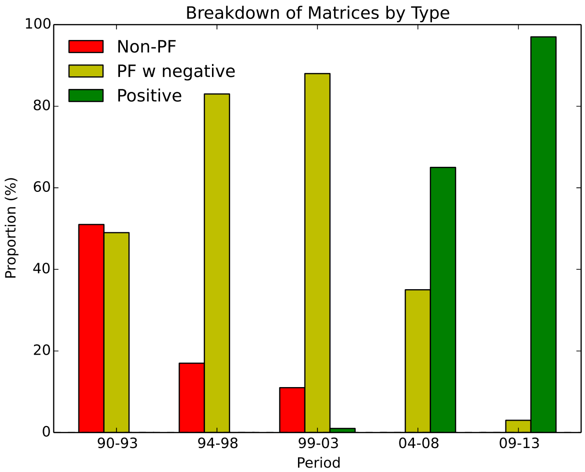
There are sometimes negative correlations between stock returns. Hence we are interested in correlation matrices which have the PF property

Empirical experiment

We obtain CRSP daily returns for S&P1500 components from 1990-2013.

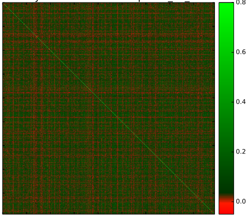
- Divide data to five-year periods
- Select 10,000 random samples of 50 stocks, compute **C**
  - All matrices positive-definite
- For non-positive matrices, we study the distribution of elements
- Test whether the type can be determined based on simple rules for the elements

Changes in correlation through time

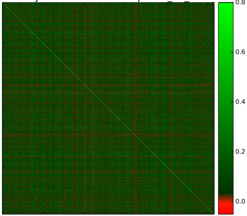


Visualizing changes in correlation through time

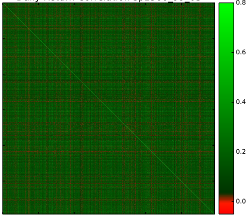
Daily Return Correlation sp1500_90_93



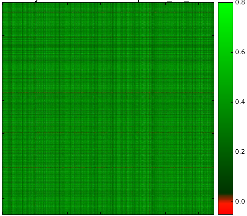
Daily Return Correlation sp1500_94_98



Daily Return Correlation sp1500_99_03

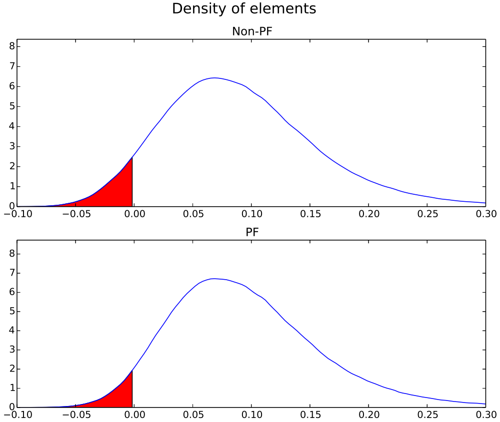


Daily Return Correlation sp1500_04_08



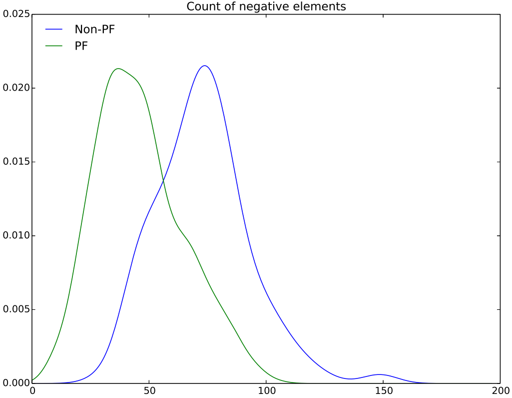
Distribution of elements

Fatter left tails for Non-PF correlation matrices. E.g. 1994-1998:



Count of negative elements

Estimated density for the count of negative elements. E.g. 1994-1998:



Characteristics of stock correlations

- Which stocks are most prevalent in Non-PF matrices?
  - rank stocks by average correlation with the rest (ascending)
  - take the top three; these are "low correlation" stocks
- Low correlation stocks appear as (almost) full rows of negative elements
- When these stocks are selected we end up with Non-PF matrices
  - Consistent with proposition on strictly negative rows (discussed later)

Element distribution: empirical vs simulated

Statistics of negative elements

Stat	Empirical		Simulated	
	PF	Non-PF	PF	Non-PF
Count*	45	72	N/A	612
Mean	-0.016	-0.019	N/A	-0.112
Std dev	0.014	0.015	N/A	0.084
Min	-0.171	-0.171	N/A	-0.661
25%	-0.022	-0.027	N/A	-0.162
50%	-0.012	-0.015	N/A	-0.096
75%	-0.005	-0.006	N/A	-0.045
Max	-0.000	-0.000	N/A	-0.000

* average number per matrix

Visualizing negative rows

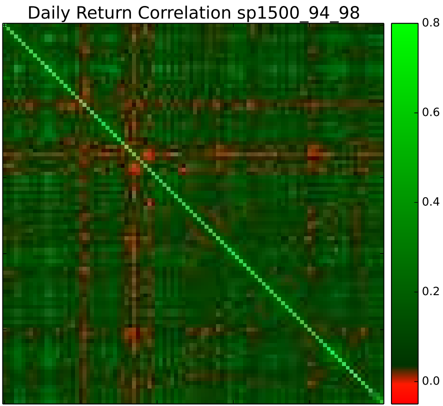


Figure: Correlation matrix visualization. 100 stocks daily returns, 1994-1998.

Visualizing negative rows

Different sampling frequency (e.g. weekly) can sometimes reduce the number of negative correlation

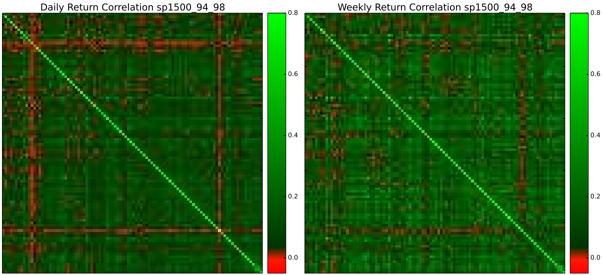


Figure: Correlation matrix visualization. Daily (left) returns vs weekly (right) returns

Notation

- $\mathbb{PF}(n)$ is the set of $n \times n$ correlation matrices possessing the strong Perron-Frobenius property
- $\mathbb{PF}^{+}(n)$ is the set of matrices in $\mathbb{PF}(n)$ that have only positive elements
- $\mathbb{PF}^{-}(n)$ is the set of matrices in $\mathbb{PF}(n)$ that have at least one negative element
- $\mathbb{PD}(n)$ is the set of $n \times n$ positive-definite correlation matrices
- $\mathbb{PF}^{PD}(n)$ is the set $\mathbb{PF}(n) \cap \mathbb{PD}(n)$

Analytic result for three by three case

- Consider the 3×3 correlation matrix

$$\mathbf{C} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix}$$

$\mathbf{C} \in \mathbb{PF}(3)$ when the following condition holds

$$\begin{cases} \rho_{12} + \rho_{13} & > 0 \\ \rho_{12} + \rho_{23} & > 0 \\ \rho_{13} + \rho_{23} & > 0 \end{cases}$$

- When $\mathbf{A}, \mathbf{B} \in \mathbb{PF}^{PD}(3)$, then convex combination always preserve the above condition
- $\mathbb{PF}^{PD}(3)$ is convex

Negative row

Assume $\mathbf{C} \in \mathbb{PF}(n)$, and λ and $\boldsymbol{v}$ are dominant eigenpair. Then $\lambda > 1$ and $v_i > 0, i = 1, \ldots, n$. For each row i,

$$v_i + \sum_{j \neq i} v_j \rho_{i,j} = \lambda v_i \tag{2}$$

$$v_1(\lambda - 1) = \sum_{j \neq i} v_j \rho_{i,j} \tag{3}$$

The LHS of (3) is positive. Thus, $\rho_{i,j}$ cannot be all negative. In other words, PF matrix cannot have rows of only negative off diagonals!

Constant correlation matrices

Suppose **P** is a correlation matrix. **P** has all off-diagonal entries equal to ρ

- Whenever $\rho > -\frac{1}{n-1}$, $\mathbf{P} \in \mathbb{PD}(n)$
- Furthermore, $\mathbf{P} \in \mathbb{PF}^{PD}(n)$ if and only if $\rho > 0$

Numerical investigations

- Simulate 10,000 random positive-definite correlation matrices using Harry Joe's method (2006)
- Study the distribution of Non-PF and PF matrices among various dimensions
- Within the set of PF matrices, we test convexity properties
- Better understand eventually positive condition

Simulated proportions by type

Dimension	$\mathbb{PF}^+(n)$	$\mathbb{PF}^-(n)$	**Non-PF**
3	14.78%	10.04%	75.18%
4	3.29%	9.31%	87.40%
5	0.45%	5.59%	93.96%
6	0.08%	3.02%	96.90%
7	0.01%	1.55%	98.44%
8	0.00%	0.79%	99.21%

Table: Proportion of sample correlation matrices by type from dimension 3 to 8

- The set of PF matrices shrinks while matrix dimension increases
- The decreasing portion of positive matrices proposes a limitation of simulation method - low likelihood of getting positive matrices

Eventually positive matrices

Definition

An $n \times n$ matrix $\mathbf{A}$ is said to be eventually positive if there exists a positive integer k_0 such that $\mathbf{A}^k > 0$ for all $k > k_0$.

Theorem (Noutsos)

For any symmetric $n \times n$ matrix $\mathbf{A}$ the following properties are equivalent.

1. $\mathbf{A}$ possesses the strong Perron-Frobenius property.
2. $\mathbf{A}$ is eventually positive

Summary

- Interest in finding compatible $\boldsymbol{x}^{(m)}$ and $\boldsymbol{\mu}$
- Market portfolio has positive weights
- Proxied by dominant eigenvector of correlation matrix
- When does dominant eigenvector have positive weights
- Perron-Frobenius property
- Explored this question in three ways
  1. Using empirical data
  2. Theoretical analysis
  3. Numerical simulation

Conclusions

Empirical results

- Negative correlation has declined during last 20 years
- Negative correlations tend to occur in rows
- Has implications for PF property

Analytical results

- A row of negative correlations destroys the PF property
- Obtained a simple characterization of 3×3 matrices
- Constant correlation matrices have simple classification

Simulation results

- PF matrices rare in high dimensions for random matrices
- Failure of PF related to negative elements
- Related to number, size and position of negative elements