Correlation Matrices and the Perron-Frobenius Theorem

Phelim Boyle

Wilfrid Laurier University

July 14 2014
Thanks to David Melkuev, Johnew Zhang, Bill Zhang, Ronnnie Feng and Jeyan Thangaraj for research assistance.
Correlation Matrices and the Perron-Frobenius Theorem

Phelim Boyle

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

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

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

(1)

- Assume $\Sigma$ known
- Assume- for now- we have a way to find $x^{(m)}$
- Task is to find compatible $\mu$
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
- 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, $A$, 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 $A$ 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.
1 Introduction and background

2 Perron Frobenius

3 Analysis of empirical data
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
Correlation Matrices and the Perron-Frobenius Theorem

Phelim Boyle

Introduction and background

Perron Frobenius

Analysis of empirical data

Theoretical results

Numerical investigations

Summary and conclusions

Visualizing changes in correlation through time

<table>
<thead>
<tr>
<th>Date Range</th>
<th>Correlation Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>sp1500 90 93</td>
<td>Daily Return</td>
</tr>
<tr>
<td>sp1500 94 98</td>
<td>Daily Return</td>
</tr>
<tr>
<td>sp1500 99 03</td>
<td>Daily Return</td>
</tr>
<tr>
<td>sp1500 04 08</td>
<td>Daily Return</td>
</tr>
</tbody>
</table>

[Graphs showing correlation matrices]
Fatter left tails for Non-PF correlation matrices. E.g. 1994-1998:
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

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

* average number per matrix
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
Introduction and background

Perron Frobenius

Analysis of empirical data

Theoretical results

Summary and conclusions
Notation

- $\mathbf{PF}(n)$ is the set of $n \times n$ correlation matrices possessing the strong Perron-Frobenius property
- $\mathbf{PF}^+(n)$ is the set of matrices in $\mathbf{PF}(n)$ that have only positive elements
- $\mathbf{PF}^-(n)$ is the set of matrices in $\mathbf{PF}(n)$ that have at least one negative element
- $\mathbf{PD}(n)$ is the set of $n \times n$ positive-definite correlation matrices
- $\mathbf{PF}^{PD}(n)$ is the set $\mathbf{PF}(n) \cap \mathbf{PD}(n)$
Consider the $3 \times 3$ correlation matrix

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

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

$\mathbb{PF}^{PD}(3)$ is convex
Assume $C \in \mathbb{PF}(n)$, and $\lambda$ and $\nu$ are dominant eigenpair. Then $\lambda > 1$ and $\nu_i > 0$, $i = 1, \ldots, n$. For each row $i$,

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

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

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!
Suppose $P$ is a correlation matrix. $P$ has all off-diagonal entries equal to $\rho$

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

2 Perron Frobenius

3 Analysis of empirical data

4 Theoretical results

5 Numerical investigations
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

<table>
<thead>
<tr>
<th>Dimension</th>
<th>$\text{PF}^+(n)$</th>
<th>$\text{PF}^-(n)$</th>
<th>Non-PF</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>14.78%</td>
<td>10.04%</td>
<td>75.18%</td>
</tr>
<tr>
<td>4</td>
<td>3.29%</td>
<td>9.31%</td>
<td>87.40%</td>
</tr>
<tr>
<td>5</td>
<td>0.45%</td>
<td>5.59%</td>
<td>93.96%</td>
</tr>
<tr>
<td>6</td>
<td>0.08%</td>
<td>3.02%</td>
<td>96.90%</td>
</tr>
<tr>
<td>7</td>
<td>0.01%</td>
<td>1.55%</td>
<td>98.44%</td>
</tr>
<tr>
<td>8</td>
<td>0.00%</td>
<td>0.79%</td>
<td>99.21%</td>
</tr>
</tbody>
</table>

**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
Definition

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

Theorem (Noutsos)

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

1. $A$ possesses the strong Perron-Frobenius property.
2. $A$ is eventually positive.
Introduction and background

Perron Frobenius

Analysis of empirical data

Theoretical results

Numerical investigations

Summary and conclusions
Interest in finding compatible $x^{(m)}$ and $\mu$

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 $3 \times 3$ 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