Application of Statistical Physics in Time Series Analysis

Duan Wang
Center for Polymer Studies, Boston University
wangduan@bu.edu

Advisor: H. Eugene Stanley
Outline

Basic Concepts in Financial Time Series

Random Matrix Theory
- Meaning of eigenvalues and eigenvectors
- Eigenvalue distribution for random correlation matrix
- Interpreting empirical eigenvalue distribution

Extensions of Random Matrix Theory
- Autoregressive random matrix theory (ARRMT)
- Time-lag random matrix theory (TLRMT)
- Global factor model (GFM)

(Application of RMT in portfolio optimization)
Price, Return and Return Magnitude

Price ($S_{i,t}$): index $i$ at time $t$

Return:
\[ R_{i,t} \equiv \log S_{i,t} - \log S_{i,t-1} \]
\[ = \log \left( \frac{S_t}{S_{t-1}} \right) \]
\[ = \log \left( \frac{S_{t-1} - S_t}{S_{t-1}} \right) \approx \frac{\Delta S_{t-1}}{S_{t-1}} \]

Magnitude of return:
\[ |r^*_i| \equiv \left| R_{i,t} - \langle R_{i,t} \rangle \right| \]
Measure of Dependence: Cross–Correlation, Autocorrelation, and Time–Lag Cross–Correlation

For time series \( \{X_t\} \) and \( \{Y_t\} \),

- **Cross–correlation**
  \[
  C_{XY} \equiv \frac{\langle X_t Y_t \rangle - \langle X_t \rangle \langle Y_t \rangle}{\sigma_X \sigma_Y}
  \]

  Denotes average

  Standard deviations

- **Autocorrelation**
  \[
  A_X(\Delta t) \equiv \frac{\langle X_t X_{t+\Delta t} \rangle - \langle X_t \rangle \langle X_{t+\Delta t} \rangle}{\sigma_X \sigma_X}
  \]

  Time lag

- **Time–lag Cross–correlation**
  \[
  C_{XY}(\Delta t) \equiv \frac{\langle X_t Y_{t+\Delta t} \rangle - \langle X_t \rangle \langle Y_{t+\Delta t} \rangle}{\sigma_X \sigma_Y}
  \]

**Properties:**

- \( 0 < C < 1 \) correlated
- \( C = 0 \) uncorrelated
- \( -1 < C < 0 \) anti–correlated

Example: highly correlated France and Italy indices
Examples of Return Autocorrelations and Magnitude Autocorrelations

Simulate return autocorrelations: AR(1)

\[ X_t = \phi X_{t-1} + \epsilon_t \]

Simulate magnitude autocorrelations: GARCH(1,1)

\[ \sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \]
Multiple Time Series

Annual Return of 6 US stocks

<table>
<thead>
<tr>
<th>Date</th>
<th>GE</th>
<th>MSFT</th>
<th>JNJ</th>
<th>K</th>
<th>BA</th>
<th>IBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-Jan-94</td>
<td>56.44%</td>
<td>-1.50%</td>
<td>6.01%</td>
<td>-9.79%</td>
<td>58.73%</td>
<td>21.51%</td>
</tr>
<tr>
<td>3-Jan-95</td>
<td>18.23%</td>
<td>33.21%</td>
<td>41.56%</td>
<td>7.46%</td>
<td>-0.24%</td>
<td>6.04%</td>
</tr>
<tr>
<td>2-Jan-96</td>
<td>56.93%</td>
<td>44.28%</td>
<td>57.71%</td>
<td>37.76%</td>
<td>65.55%</td>
<td>27.33%</td>
</tr>
<tr>
<td>2-Jan-97</td>
<td>42.87%</td>
<td>79.12%</td>
<td>22.94%</td>
<td>-5.09%</td>
<td>54.34%</td>
<td>41.08%</td>
</tr>
<tr>
<td>2-Jan-98</td>
<td>47.11%</td>
<td>38.04%</td>
<td>17.62%</td>
<td>32.04%</td>
<td>37.11%</td>
<td>2.63%</td>
</tr>
<tr>
<td>4-Jan-99</td>
<td>34.55%</td>
<td>85.25%</td>
<td>26.62%</td>
<td>-10.74%</td>
<td>22.00%</td>
<td>-2.11%</td>
</tr>
<tr>
<td>3-Jan-00</td>
<td>28.15%</td>
<td>11.20%</td>
<td>3.41%</td>
<td>-48.93%</td>
<td>43.53%</td>
<td>23.76%</td>
</tr>
<tr>
<td>2-Jan-01</td>
<td>4.61%</td>
<td>-47.19%</td>
<td>10.69%</td>
<td>11.67%</td>
<td>28.29%</td>
<td>21.76%</td>
</tr>
<tr>
<td>2-Jan-02</td>
<td>-19.74%</td>
<td>4.27%</td>
<td>-7.00%</td>
<td>19.90%</td>
<td>-15.09%</td>
<td>4.55%</td>
</tr>
<tr>
<td>2-Jan-03</td>
<td>-44.78%</td>
<td>-29.47%</td>
<td>-5.67%</td>
<td>10.88%</td>
<td>-23.23%</td>
<td>15.54%</td>
</tr>
<tr>
<td>2-Jan-04</td>
<td>35.90%</td>
<td>18.01%</td>
<td>-1.27%</td>
<td>15.49%</td>
<td>39.82%</td>
<td>31.80%</td>
</tr>
</tbody>
</table>

6 Stocks
6x6 Correlation matrix

Symmetric matrix
Main diagonal=1

Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>GE</th>
<th>MSFT</th>
<th>JNJ</th>
<th>K</th>
<th>BA</th>
<th>IBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>1</td>
<td>0.5791</td>
<td>0.5383</td>
<td>-0.056</td>
<td>0.905</td>
<td>0.2819</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.5791</td>
<td>1</td>
<td>0.56</td>
<td>-0.0532</td>
<td>0.3502</td>
<td>-0.0406</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.5383</td>
<td>0.56</td>
<td>1</td>
<td>0.2838</td>
<td>0.3925</td>
<td>0.0002</td>
</tr>
<tr>
<td>K</td>
<td>-0.056</td>
<td>-0.0532</td>
<td>0.2838</td>
<td>1</td>
<td>-0.1234</td>
<td>-0.1427</td>
</tr>
<tr>
<td>BA</td>
<td>0.905</td>
<td>0.3502</td>
<td>0.3925</td>
<td>-0.1234</td>
<td>1</td>
<td>0.5821</td>
</tr>
<tr>
<td>IBM</td>
<td>0.2819</td>
<td>-0.0406</td>
<td>0.0002</td>
<td>-0.1427</td>
<td>0.5821</td>
<td>1</td>
</tr>
</tbody>
</table>

Eigenvalue decomposition

\[ C = QAQ^{-1} \]

\[ \Lambda = diag\{2.80, 1.45, 0.96, 0.43, 0.34, 0.02\} \]

Meaning of eigenvalues?
Explained by random matrix theory (RMT)
- **Original RMT**
  - Aim: find existence of collective behavior
  - Method: compare the eigenvalue distribution between
    1. a **symmetric random matrix** and
    2. a **Hamiltonian matrix**

- **RMT in Econophysics**
  - Aim: find existence of cross-correlation in multiple time series
  - Method: compare the eigenvalue distribution of
    1. a **Wishart matrix (random correlation matrix)** and
    2. empirical **cross-correlation matrix**

*Wishart matrix: sample correlation matrix for uncorrelated time series*
Sample cross-correlations between independent time series are not zero if time series are of finite length.

Mathematically: $C_{ij}=0$

Statistically: $C_{ij}$ falls in $(-1.96\sqrt{\frac{1}{T}},1.96\sqrt{\frac{1}{T}})$, with 95% probability.

Cross-correlation distribution between 2 uncorrelated time series, each with length $T=2000$.

Sample correlation distribution: $C_{ij}\sim\text{N}(0,1/T)$
Gaussian distribution with mean zero and variance $1/T$
Meaning of Eigenvalues and Eigenvectors

**Principal Components**
- Linear combination of individual time series
- Orthogonal transformation of correlated time series

**Eigenvectors (Factor Loadings)**
- Weight of each individual time series in a principal component

**Eigenvalues**
- Variance of each principal component
- Measures the significance of a factor

\[ N \text{ time series with length } T \rightarrow \text{principal components} \]

Not all of them are significant

Compare them with the eigenvalues of Wishart matrix to find out the number of significant factors
Eigenvalue Distribution for Wishart Matrix

- Eigenvalue distribution of a Wishart matrix when
- Only determine by $\sigma$ and $Q \equiv \frac{T}{N}$
- Has an upper and lower bound
Empirical Eigenvalue Distribution


Eigenvalue distribution of Wishart matrix

Largest eigenvalue (not shown):
Average behavior
(Behavior of the whole market)

\[ \lambda_{\pm} = 1 + \frac{1}{Q} \pm 2 \sqrt{\frac{1}{Q}} \]

\[ Q \equiv \frac{L}{N} (> 1) \]

Other large eigenvalues:
Subgroups
(Sectors, industries, ...)

Largest eigenvalue of uncorrelated time series

Significant factors

Noise

\( \Delta t = 1 \text{ day} \)

1990–96

Probability density \( P(\lambda) \)
Random Matrix Theory: Summary

**Aim**

1. Test significance of correlations in multiple time series
2. Find number of significant factors
3. Reduce noise in correlation matrix

**Procedure**

1. Wishart matrix:
2. Empirical eigenvalues:
3. Significant factors:
4. Noise:
Efficient Frontier
H. Markowitz (1952)

Given:
Individual return $\mu$, Covariance matrix $\Sigma$, Portfolio return $\mu^*$

Calculate:
Weigh $w_{\text{eff}}$ that minimize the portfolio risk $w^T \Sigma w$

Calculate efficient frontier:
$$w_{\text{eff}} = \arg\min_w w^T \Sigma w$$
Subject to $$w^T \mu = \mu^* , \ w^T 1 = 1$$

Solution:
$$A = \mu^T \Sigma^T 1 1$$
$$B = \mu^T \Sigma^T \mu$$
$$C = 1^T \Sigma^T 1 1$$
$$D = BC - A$$
$$w_{\text{eff}} = \{B \Sigma^T 1 - A \Sigma^T \mu + \mu^* (C \Sigma^T 1 - \mu - A \Sigma^T 1 1)/D \}$$
Minimum Variance Portfolio: Cumulative Return

Portfolio:
404 stocks from S&P 500, rebalancing at the end of each year

After Applying RMT in constructing the minimum variance portfolio, we increased returns and reduced risk.
Autoregressive Random Matrix Theory (ARRMT)

Problem
- Wishart matrix assumes no autocorrelation
- Autocorrelation can impact eigenvalue distribution
- We should not compare empirical eigenvalue distribution of autocorrelated time series with

Action
- Quantify autocorrelation in empirical time series
- Compare its eigenvalues with the largest eigenvalue from simulated time series with no crosscorrelations but same autocorrelations as the empirical time series
Quantify Autocorrelation:
First Order Autoregressive Model (AR(1))

AR(1) Model:

\[ X_t = \phi X_{t-1} + \epsilon_t \]

Autocorrelation:
Correlation distribution for different AR(1) Coefficient

\[ \frac{1}{T} \left[ 1 + 2 \sum_{\Delta t=1}^{\Delta t^+} A(\Delta t)A'(\Delta t) \right] \]
Eigenvalue distribution for different AR(1) Coefficient

Autocorrelation in time series may influence the eigenvalue distribution for uncorrelated time series.

AR(1) Process
\[ X_{t} = \phi X_{t-1} + \epsilon_{t} \]

Variance of correlation
\[ \text{Var}(r) = \frac{1}{T} \frac{1+\phi^2}{1-\phi^2} \]

Equivalent length
\[ T^* = T \frac{1-\phi^2}{1+\phi^2} \]

Equivalent \( Q \)
\[ Q^* = \frac{T^*}{N} \]

Wishart Matrix:
\[ \lambda^+ = (1+\sqrt{1/Q})^2 \]
ARRMT: Summary

Aim

Adjust RMT for autocorrelated time series

Procedure

(1) Fit AR models, get
(2) Simulate time series using the estimated
(3) Calculate largest eigenvalue
(4) Repeat (1)–(3)
(5) Distribution of largest eigenvalue
(6) For small data sets, choose 95\textsuperscript{th} percentile
(7) Compare with empirical eigenvalues
Empirical Data: Is Autocorrelation Significant?

Data: Change of air pressure in 95 US cities

Histogram of AR(1) Coefficients
RMT and ARRMT for Daily Air Pressure of 95 US Cities

RMT
• 11 significant factors

ARRMT
• 8 significant factors
Time Lag Random Matrix Theory (TLRMT)

(Podobnik, Wang et. al., 2010, EPL.)

- Aim: extend RMT to time lag correlation matrix
- Unsymmetrical matrices, eigenvalues are complex numbers
- Use singular value decomposition (SVD) instead of eigenvalue
- Largest singular value: strength of correlation

\[ M = U \Sigma V^* \]

Correlation matrix for lag=0 and lag=1

<table>
<thead>
<tr>
<th>Lag=0</th>
<th>GE</th>
<th>MSFT</th>
<th>JNJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>1</td>
<td>0.579</td>
<td>0.538</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.579</td>
<td>1</td>
<td>0.56</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.538</td>
<td>0.56</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lag=1</th>
<th>GE</th>
<th>MSFT</th>
<th>JNJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>0.378</td>
<td>0.499</td>
<td>0.551</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.528</td>
<td>0.339</td>
<td>0.295</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.635</td>
<td>0.659</td>
<td>0.499</td>
</tr>
</tbody>
</table>
Time lag RMT:
Singular value vs time-lag cross-correlations
TLRMT Applied to 48 Stock Indices

We find:

1. Short-range return cross-correlations (after time lag=2)
2. Long-range magnitude cross-correlation (scaling exponent=0.25)
Global factor model (GFM)

Wang et al., 2011, PRE.

- Aim: explain cross-correlations using one single process.
- Assumption: Each individual index fluctuates in response to one common process, the “global factor” $M_t$.

Measures dependence of $R_{i,t}$ on $M_t$

$$R_{i,t} = \mu_i + b_i M_t + \epsilon_{i,t}.$$  

individual return  
mean return  
Global factor  
unsystematic noise

What is the global factor?
- A linear combination of all individual index returns
  $$M_t = \omega_1 R_{1,t} + \omega_2 R_{2,t} + \ldots + \omega_N R_{N,t}$$
- The weight of each index return is calculated using statistical method.
GFM Properties:

- Variance of global factor—- Cross-correlation among individual indices (holds for both returns and magnitudes)
  \[
  \text{Cov}(r_{i,t}, r_{j,t}) = b_i b_j \text{Var}(M_t). \quad \text{Return cross-correlation}
  \]
  \[
  \text{Cov}(r_{i,t}^2, r_{j,t}^2) = b_i^2 b_j^2 \text{Var}(M_t^2). \quad \text{Magnitude cross-correlation}
  \]

- Autocorrelation of global factor—- Time lag cross-correlation among individual indices
  \[
  \text{Cov}(r_{i,t}, r_{j,t}; \Delta t) = b_i b_j A_M(\Delta t).
  \]
  \[
  \text{Cov}(r_{i,t}^2, r_{j,t}^2; \Delta t) = b_i^2 b_j^2 A_{M^2}(\Delta t).
  \]
  \[A_M = \text{autocorrelation of global factor}\]

**Conclusion:**

Cross-correlation among individual indices decay in the same pattern as the autocorrelation of the global factor.

*GFM explains the reason for the long range magnitude cross-correlations.*
Estimate of the global factor:

Principal Component Analysis (PCA)

- **Aim:** Linear regression with unobservable $M_t$
  \[ R_{i,t} = \mu_i + b_i M_t + \epsilon_{i,t}. \]

- **Basis:** Least total squared errors

- **Calculation:** Eigenvalue decomposition ($C=U^+DU$)
  1. The principal components are related with the eigenvectors.
  2. $M_t =$ the first principal component

- **Procedure:**
  1. Find $M_t$ using eigenvalue decomposition.
  2. Linear regression, find $\mu, b_i, \epsilon$. 

Unknown, linear combination of $R_{i,t}$
Correlation matrix
Eigenvector matrix
Eigenvalue matrix
Estimate of the global factor: Result

Global Factor (M)

Autocorrelation of M

Autocorrelation of abs(M)
**Significant test of Global Factor**

- Only 3 eigenvalues above Wishart maximum (significant).
- The largest eigenvalue (global factor) constitute 31% of total variance, and 75% of total variance of 3 significant principal components.
- Conclusion: the single global factor is sufficient in explaining correlations.

*Eigenvalues above and below Wishart maximum.*

![Graph showing eigenvalues and principal components](image)
Application: risk forecasting

- **Define risk: volatility (time dependent standard deviation)**
  
  volatility = expected standard deviation of a time series at time t given information till time t-1.

- **How to calculate risk: Apply GARCH to global factor**
  
  Coefficients estimated using maximum likelihood (MLE)

  \[
  \sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2
  \]

  Forecasted using recursion with estimated coefficients

- **Historical risk and expectation of future risk**

![Graph showing time series and conditional volatility](image-url)
Application: Multiple Global Factors
Wang et. al., working paper.

- PC1: American and EU countries
- PC2: Asian and Pacific countries
- PC3: Middle east countries

Statistically, the world economy can be grouped as 3 sectors.

\[
r_j = a_j + b_{j1}F_1 + b_{j2}F_2 + \cdots + b_{jn}F_n + \epsilon_j
\]

3 significant principal components (PCs)
Multiple Global Factors: Western, Asian and Pacific, and Middle East

1. Large correlation between Western and Asian economies
2. Small correlation between Middle East Economy and the other 2 groups
3. Each group has its own financial crises.
4. The 2008 market crash influenced all 3 groups, indicating globalization (large volatility correlation).
Thanks for listening!

- Duan Wang
- wangduan@bu.edu
Eigenvectors: Implied Participation Number

Definition

\[ Nlp \overset{\text{def}}{=} \frac{1}{\sum_{i=1}^{N} v_{i}} \]

Interpretation

A rough estimation of how many individuals contribute to the PC.

Two extreme cases

1. \( v_{i} = \frac{1}{\sqrt{N}} \), for all \( i \)’s. \( Nlp = N \).
2. \( v_{1} = 1, v_{i} \neq 1 = 0 \). \( Nlp = 1 \).

For Wishart matrix
Empirical Implied Participation Number

422 stocks from S&P 500, 1737 daily returns

Wishart Matrix
Similar participation number for all PCs

Empirical Correlation Matrix
- First PC has 370 participants (87.7% of all stocks)
  Global market factor
- First few PCs have large $N/p$
  Sectors or industry groups
- Last few PCs have small $N/p$
  Small subgroups
Empirical Implied Participation Number

**First Eigenvector**
- All positive, similar weights
- Indicating global factor

**Last Eigenvector**
- Small number of participants
- Indicating small subgroup
**Efficient Frontier**  
**H. Markowitz (1952)**

**Given:**  
Individual return $\mu$, Covariance matrix $\Sigma$, Portfolio return $\mu^*$

**Calculate:**  
Weigh $w^{\text{eff}}$ that minimize the portfolio risk $w^T \Sigma w$

Calculate efficient frontier:  
$$w^{\text{eff}} = \arg\min_w w^T \Sigma w$$
Subject to  
$$w^T \mu = \mu^*, \quad w^T 1 = 1$$

**Solution:**  
$$A = \mu^T \Sigma \uparrow 1 1$$
$$B = \mu^T \Sigma \uparrow \mu$$
$$C = 1^T \Sigma \uparrow 1 1$$
$$D = BC - A$$
$$w^{\text{eff}} = \{B\Sigma \uparrow -1 1 - A\Sigma \uparrow -1 \mu + \mu^* (C\Sigma \uparrow -1 \mu - A\Sigma \uparrow -1 1) \}/D$$
Example: 404 stocks from S&P 500

Rebalance every end of year

Use data from past 5 years to calculate $\mathbf{w}^{\text{eff}}$
At the end of 2006, use 2002-2006

Rebalance portfolio using $\mathbf{w}^{\text{eff}}$

What is realized frontier in 2007?
Example: 404 stocks from S&P 500

Rebalance every end of year

Use data from past 5 years to calculate $w^{\text{eff}}$
At the end of 2006, use 2002-2006

Rebalance portfolio using $w^{\text{eff}}$

What is realized frontier in 2007?

Problem: in sample estimate of the weights does not minimize the out of sample error
Reason for Instability

\[
\mathbf{w}_{\text{minvar}} = \mathbf{\Sigma}^{-1} \frac{1}{\sqrt{p}} \mathbf{\Sigma}^{-1/2} \mathbf{1}
\]

Eigenvalue decomposition

- \[\mathbf{\Sigma} = \mathbf{V} \mathbf{D} \mathbf{V}^{-1}\]
- \[\mathbf{\Sigma}^{-1} = \mathbf{V} \mathbf{D}^{-1} \mathbf{V}^{-1}\]
- \[\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_N)\]. Eigenvalues \[\lambda_1 > \lambda_2 > ... > \lambda_N\].
- \[\mathbf{D}^{-1} = \text{diag}(\lambda_1^{-1}, \lambda_2^{-1}, ..., \lambda_N^{-1})\].

Smallest eigenvalues play larger roles in calculating \(\mathbf{w}_{\text{minvar}}\)

Smaller PCs are dominated by noise

How to solve the problem?
Use RMT to determine the number of eigenvalues we should keep
RMT in Portfolio Optimization: Procedure

Use RMT to decide number of significant factors $M$

Dimension reduction

- $D = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_N)$. Eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_N$.
- $D_{adj} = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_M, \lambda_{r}, \lambda_{r}, ..., \lambda_{r})$. $\lambda_{r} \overset{\text{def}}{=} \frac{1}{N} - M \sum_{i=M+1}^{N} \lambda_i$
- $\Sigma_{adj} = V D_{adj} \ V^T$

Use adjusted covariance matrix to calculate frontier and weights
Realized Frontier: RMT Applied
Minimum Variance Portfolio: Cumulative Return

After Applying RMT in constructing the minimum variance portfolio, we increased returns and reduced risk.
Minimum Variance Portfolio: Turnover Rate
Other New Features in RMT

- Time-lag RMT (TLRMT)
- Global Factor Model (GFM)
- Variance Crosscorrelation
  - Conditional Variance Adjusted Regression Model (CVARM)
  - Conditional Heteroskedasticity Adjusted Regression Model (CHARM)

\[ R_{i,t} = \mu_i + b_i M_t + \epsilon_{i,t} \]

\[ \epsilon_{i,t} = \sigma_{i,t} \eta_{i,t} \]
\[ \sigma_{i,t}^2 = \alpha_0 + \alpha_1 \epsilon_{i,t-1}^2 + \beta_1 \sigma_{i,t-1}^2 + \gamma_1 M_{t-1}^2 \]
\[ M_t = \tilde{\sigma}_{t} \tilde{\eta}_t \]
\[ \tilde{\sigma}_t^2 = \tilde{\alpha}_0^2 + \tilde{\alpha}_1 M_{t-1}^2 + \tilde{\beta}_1 \sigma_{t-1}^2 \]
\[ \text{Var}(R_{i,t}) = (b_i^2 + \frac{\gamma_1}{1 - \alpha_1 - \beta_1}) \text{Var}(M_t) + \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \]
Crosscorrelation Distribution vs AR(1) Coefficients

Variance

\[ \frac{1}{T} \left[ 1 + 2 \sum_{\Delta t=1}^{\Delta t_{\text{max}}} A(\Delta t)A'(\Delta t) \right] \]
Largest Eigenvalue vs Autocorrelation

\[ X_t = \phi X_{t-1} + \epsilon_t \]

\[ \lambda_+ = \sqrt{\frac{1 + \phi^2}{1 - \phi^2}} \lambda_+ \]

\[ N=2000 \]
\[ T=8000 \]
RMT for Small-Sized Data

N=20, T=40

95th Percentile
Other Possible Extensions:

- RMT with existence of multicollinearity
- RMT with N > T

(Show Figures Here!)