Seeking Gold in Sand

Financial applications of Random Matrix Theory in stock market data

Mike S. Wang

Faculty of Mathematics // Department of Chemistry
University of Cambridge

October, 2016

Preview of the talk

1. What is Random Matrix Theory?
2. Data Overview
3. Random Matrix Theory: The Marčenko-Pastur Law
4. Mode & Clustering Analyses
5. A Multi-layer Structured Correlation Model & Its Predictions
6. Effect of Layer Division & an Excellent Match
7. Summary & Further Developments
8. Q&A Time

1 What is Random Matrix Theory?

Random Matrix Theory (RMT) is the study of *matrices with random variable (r.v.) entries* , e.g.

$$\begin{bmatrix} X_{11} & X_{12} & \cdots \\ X_{21} & X_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

In particular, it concerns the *emergent behaviours* of random matrices in the asymptotic limit.

Introduced by Wishart (1928), RMT gained prominence when Wigner (1950s) applied the theory to spacing of energy levels in nuclear physics.

1 What is Random Matrix Theory?

Random Matrix Theory (RMT) is the study of *matrices with random variable (r.v.) entries* , e.g.

$$\begin{bmatrix} X_{11} & X_{12} & \cdots \\ X_{21} & X_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

In particular, it concerns the *emergent behaviours* of random matrices in the asymptotic limit.

Introduced by Wishart (1928), RMT gained prominence when Wigner (1950s) applied the theory to spacing of energy levels in nuclear physics.

1 What is Random Matrix Theory?

Random Matrix Theory (RMT) is the study of *matrices with random variable (r.v.) entries* , e.g.

$$\begin{bmatrix} X_{11} & X_{12} & \cdots \\ X_{21} & X_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

In particular, it concerns the *emergent behaviours* of random matrices in the asymptotic limit.

Introduced by Wishart (1928), RMT gained prominence when Wigner (1950s) applied the theory to spacing of energy levels in nuclear physics.

2 Data overview: S&P 500

Procedure

Price indices $\rightarrow$ logarithmic returns $\rightarrow$ de-meaned and normalised data.

$$\begin{array}{cc} & P \text{ stocks} \\ T \text{ days} & \underbrace{\begin{bmatrix} 126.8 & 30.5 & \cdots \\ 126.3 & 30.7 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}}_{\text{raw data}} \end{array} \mapsto \underbrace{\begin{bmatrix} -0.9 & 1.7 & \cdots \\ 1.5 & 0.3 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}}_{\text{standardised data } X} \mapsto \underbrace{\begin{bmatrix} 1.0 & 0.2 & \cdots \\ 0.2 & 1.0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}}_{\text{correlation matrix } E}$$

The log return is

$$R_i = \log \frac{p_i}{p_{i-1}} (\approx \frac{p_i - p_{i-1}}{p_{i-1}}), \quad i > 1$$

where p_i is the i-th trading day price index. The empirical correlation matrix is

$$E = \frac{1}{T} X^t X.$$

3 Random Matrix Theory: the Marčenko-Pastur law

Covariance-correlation matrices are of fundamental importance to modern portfolio theory.

They belong a class of random matrices called the Wishart ensemble.

An important universality law for this ensemble in RMT:

The Marčenko-Pastur law

If $X : T \times P$ has independently identically distributed (i.i.d.) r.v. entries with mean 0 and variance 1, then the limiting eigenvalue density function (e.d.f.) of matrix $E = T^{-1}X^tX$ is

$$f(\lambda) = \frac{1}{2\pi}\frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{r\lambda}$$

as $P, T \to \infty$ and $P/T \to r \in (0, 1)$, where $\lambda_\pm = (1 \pm \sqrt{r})^2$.

3 Random Matrix Theory: the Marčenko-Pastur law

Covariance-correlation matrices are of fundamental importance to modern portfolio theory.

They belong a class of random matrices called the Wishart ensemble.

An important universality law for this ensemble in RMT:

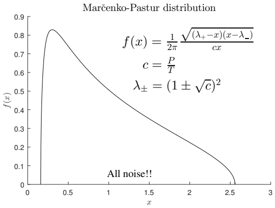
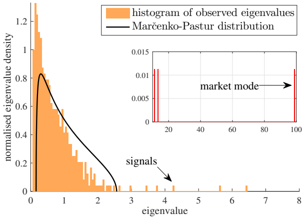
The Marčenko-Pastur law

If $X : T \times P$ has independently identically distributed (i.i.d.) r.v. entries with mean 0 and variance 1, then the limiting eigenvalue density function (e.d.f.) of matrix $E = T^{-1}X^tX$ is

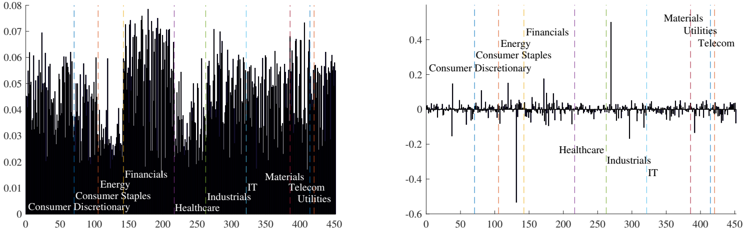
$$f(\lambda) = \frac{1}{2\pi}\frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{r\lambda}$$

as $P, T \to \infty$ and $P/T \to r \in (0, 1)$, where $\lambda_\pm = (1 \pm \sqrt{r})^2$.

3 The Marčenko-Pastur law: a crude prediction

underlying correlation matrix C		predicted e.d.f. of E
I_P	$\Rightarrow$	
?	$\Rightarrow$	

4 Mode analysis



The market mode (L: IPR 3.98×10^{-5})
and
the lowest mode (R: IPR 0.149).

Localisation

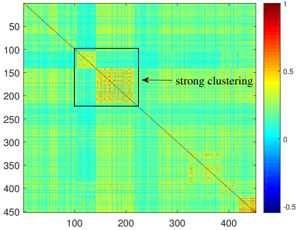
The inverse participation ratio is defined by $\mathrm{IPR}(\mathbf{v}) = \sum_{i=1}^{P} |\tilde{v}_i|^4$ where $\tilde{\mathbf{v}}$ is the vector $\mathbf{v}$ demeaned and normalised.

4 Clustering analysis

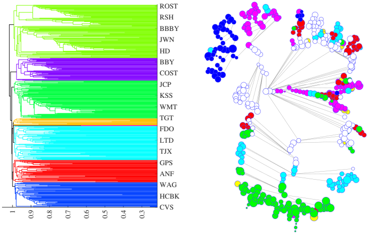
Market mode removal: $E' = E - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^t$ ($\lambda_1, \mathbf{v}_1$ largest eigenvalue pair);

Dissimilarity distance: $d_{ij} = 1 - \mathrm{corr}(i, j)$;

Average linkage: $D_{IJ} = \frac{1}{|I||J|} \sum_{i \in I, j \in J} d_{ij}$.

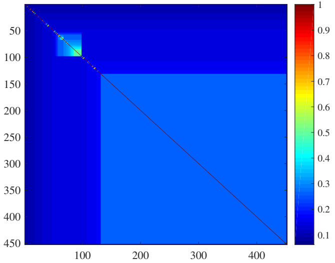


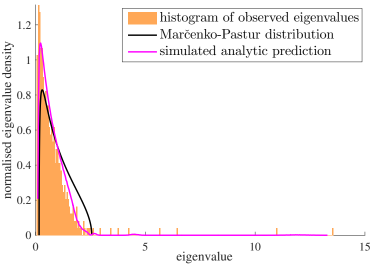
Heatmap



Dendrogram (L)
Minimum spanning tree (R)

5 A multi-layered correlation model: preview visualisation





Heatmap (L) and analytic prediction (R)
for the empirical e.d.f of a multi-layered model.

5 A multi-layered correlation model: analytic prediction

Model: $\nu(\lambda) = P^{-1}\sum_{i=1}^{P}\delta(\lambda-\lambda_i)$, the e.d.f. of underlying correlation matrix C. **Prediction**: $f(\lambda)$, the limiting e.d.f. of empirical correlation matrix E.

The Stieltjes transform pair

$$G(z) = \int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{f(\lambda)}{\lambda - z}, \qquad f(\lambda) = \lim_{\epsilon\to 0} \mathrm{Im}\{G(\lambda + i\epsilon)\}$$

The Marčenko-Pastur equation

$$-\frac{1}{G(z)} = z - r\int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{\lambda\nu(\lambda)}{1+\lambda G(z)}$$

A polynomial equation

$$[1+zG(z)]\prod_{i=1}^{P}[1+\lambda_i G(z)] = \frac{1}{r}\sum_{i=1}^{P}\lambda_i G(z)\prod_{j\neq i}^{P}[1+\lambda_j G(z)].$$

5 A multi-layered correlation model: analytic prediction

Model: $\nu(\lambda) = P^{-1} \sum_{i=1}^{P} \delta(\lambda - \lambda_i)$, the e.d.f. of underlying correlation matrix C. **Prediction**: $f(\lambda)$, the limiting e.d.f. of empirical correlation matrix E.

The Stieltjes transform pair

$$G(z) = \int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{f(\lambda)}{\lambda - z}, \qquad f(\lambda) = \lim_{\epsilon \to 0} \mathrm{Im}\{G(\lambda + i\epsilon)\}$$

The Marčenko-Pastur equation

$$-\frac{1}{G(z)} = z - r \int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{\lambda \nu(\lambda)}{1 + \lambda G(z)}$$

A polynomial equation

$$[1 + zG(z)] \prod_{i=1}^{P} [1 + \lambda_i G(z)] = \frac{1}{T} \sum_{i=1}^{P} \lambda_i G(z) \prod_{j \neq i}^{P} [1 + \lambda_j G(z)].$$

5 A multi-layered correlation model: analytic prediction

Model: $\nu(\lambda) = P^{-1} \sum_{i=1}^{P} \delta(\lambda - \lambda_i)$, the e.d.f. of underlying correlation matrix C. **Prediction**: $f(\lambda)$, the limiting e.d.f. of empirical correlation matrix E.

The Stieltjes transform pair

$$G(z) = \int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{f(\lambda)}{\lambda - z}, \qquad f(\lambda) = \lim_{\epsilon \to 0} \mathrm{Im}\{G(\lambda + i\epsilon)\}$$

The Marčenko-Pastur equation

$$-\frac{1}{G(z)} = z - r \int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{\lambda \nu(\lambda)}{1 + \lambda G(z)}$$

A polynomial equation

$$[1 + zG(z)] \prod_{i=1}^{P} [1 + \lambda_i G(z)] = \frac{1}{T} \sum_{i=1}^{P} \lambda_i G(z) \prod_{j \neq i}^{P} [1 + \lambda_j G(z)] .$$

5 A multi-layered correlation model: analytic prediction

Model: $\nu(\lambda) = P^{-1} \sum_{i=1}^{P} \delta(\lambda - \lambda_i)$, the e.d.f. of underlying correlation matrix C. **Prediction**: $f(\lambda)$, the limiting e.d.f. of empirical correlation matrix E.

The Stieltjes transform pair

$$G(z) = \int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{f(\lambda)}{\lambda - z}, \qquad f(\lambda) = \lim_{\epsilon \to 0} \mathrm{Im}\{G(\lambda + i\epsilon)\}$$

The Marčenko-Pastur equation

$$-\frac{1}{G(z)} = z - r \int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{\lambda \nu(\lambda)}{1 + \lambda G(z)}$$

A polynomial equation

$$[1 + zG(z)] \prod_{i=1}^{P} [1 + \lambda_i G(z)] = \frac{1}{T} \sum_{i=1}^{P} \lambda_i G(z) \prod_{j \neq i}^{P} [1 + \lambda_j G(z)].$$

5 A multi-layered correlation model: analytic prediction

Model: $\nu(\lambda) = P^{-1}\sum_{i=1}^{P}\delta(\lambda-\lambda_i)$, the e.d.f. of underlying correlation matrix C. **Prediction**: $f(\lambda)$, the limiting e.d.f. of empirical correlation matrix E.

The Stieltjes transform pair

$$G(z) = \int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{f(\lambda)}{\lambda - z}, \qquad f(\lambda) = \lim_{\epsilon\to 0} \mathrm{Im}\{G(\lambda + i\epsilon)\}$$

The Marčenko-Pastur equation

$$-\frac{1}{G(z)} = z - r\int_{-\infty}^{\infty} \mathrm{d}\lambda \frac{\lambda\nu(\lambda)}{1+\lambda G(z)}$$

A polynomial equation

$$[1+zG(z)]\prod_{i=1}^{P}[1+\lambda_i G(z)] = \frac{1}{T}\sum_{i=1}^{P}\lambda_i G(z)\prod_{j\neq i}^{P}[1+\lambda_j G(z)]\,.$$

6 The layer division process: fundamental structures

Layer division:

$$M := M_m(1, \alpha) \quad \text{a single cluster}$$

$$\Downarrow$$

$$M' := \begin{bmatrix} M_{m_1}(1, \alpha_1) & B \\ B^t & M_{m_2}(1, \alpha_2) \end{bmatrix} \quad \text{two smaller clusters}$$

where $m = m_1 + m_2$.

Fundamental structures:

$$M_n(x, y) \equiv \underbrace{\begin{bmatrix} x & y & \cdots & y \\ y & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & y \\ y & \cdots & y & x \end{bmatrix}}_{n}, \quad B = \beta\underbrace{(1, \ldots, 1)}_{m_1}^t \underbrace{(1, 1, \ldots, 1)}_{m_2}.$$

Useful techniques and results:

- elementary operations;
- the identity

$$\begin{pmatrix} S & T \\ U & V \end{pmatrix} \begin{pmatrix} I & 0 \\ -V^{-1}U & I \end{pmatrix} = \begin{pmatrix} S - TV^{-1}U & T \\ 0 & V \end{pmatrix}$$

  where V is invertible;

- the *Sherman-Morrison formula*:

$$(A + \mathbf{uv}^T)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{uv}^T A^{-1}}{1 + \mathbf{v}^T A^{-1}\mathbf{u}}$$

  where A is invertible and $1 + \mathbf{v}^T A^{-1}\mathbf{u} \neq 0$.

6 The layer division process: eigenvalue splitting

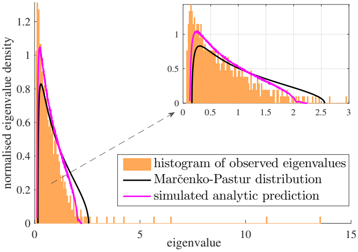
Findings:

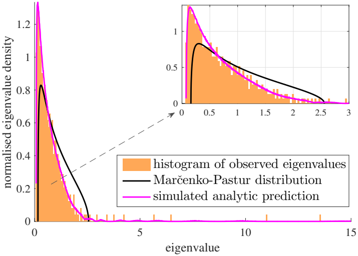
- M has an eigenvalue $\lambda_1 = 1 - \alpha$ of multiplicity $m - 1$ and an eigenvalue $\lambda_2 = 1 + (m-1)\alpha$ of multiplicity 1.
- M' has eigenvalues $1 - \alpha_{1,2}$ of multiplicity $m_{1,2} - 1$, and the remaining eigenvalues $\lambda'_{1,2}$ are roots of a quadratic polynomial with

$$\lambda'_1 + \lambda'_2 = \lambda_1 + \lambda_2.$$

The interaction correlation β perturbs these two eigenvalues, causing them to separate and repel.

6 The layer division process: an excellent match





Comparison of the predicted empirical spectral density functions of a 10-layer correlation model (L) and a 148-layer one (R) with the market mode.

7 Summary & further developments

Summary.

More considerations:

- edge asymptotics with the *Tracy-Widom law*;
- time evolution;
- fine-tuning;
- ...

7 Summary & further developments

Summary.

More considerations:

- edge asymptotics with the *Tracy-Widom law*;
- time evolution;
- fine-tuning;
- . . .

Q&A

Acknowledgements

Many thanks to Dr Lucy Colwell and her PhD student Chongli Qin at the Molecular Informatics Centre, Department of Chemistry. This project was generously supported by the Bridgwater Scheme.

More information about this project at
`http://people.ds.cam.ac.uk/sw664/SUROP%20Project%202016/SUROP.html`