Plan of the talk:

- Motivation: Portfolio selection
- Problem: EV of emprical covariance matrix
- Method: Random Matrix Theory
- Application: Some examples
- Summary

Portfolio selection (H. Markowitz)

Profit/Risk , Diversification

$$X = \sum_{i=1}^{N} p_i X_i, \quad \sum_{i=1}^{N} p_i = 1,$$

Revenue: $\langle X \rangle = \sum_{i=1}^{N} p_i \langle X_i \rangle$

Risk: $\sigma^2(X) = \sum_{ij} p_i C_{ij} p_j = \sum_i \lambda_i v_i^2$

Portfolio selection rule is to minimize risk for given expected revenue

How to compute: C_{ij}?

Historical data: x_{it} value of X_i at t

$$c_{ij} = \frac{1}{T} \sum_{t=1}^{T} x_{it} x_{jt}$$

$$\rho_c(\lambda) \quad \longrightarrow \quad \rho_C(\lambda)$$

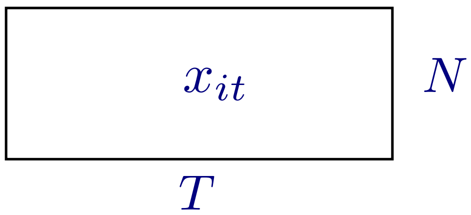
Genuine vs empirical covariance matrix

Statistical system of N-degrees of freedom: X_i, $i = 1, \ldots, N$

Covariance matrix: $C_{ij} = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle$

Additional assumption: $\langle X_i \rangle = 0 \longrightarrow C_{ij} = \langle X_i X_j \rangle$

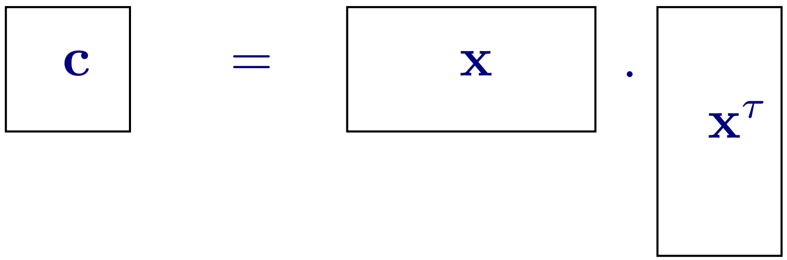
Empirical covariance matrix: T measurements of X_i: x_{it}, $t = 1, \ldots, T$



$r = N/T$, ($r < 1$)

$$c_{ij} = \frac{1}{T} \sum_{t=1}^{T} x_{it} x_{jt}$$

$$\mathbf{c} = \frac{1}{T} \mathbf{x} \mathbf{x}^{\tau}$$



Eigenvalue spectrum

Example: i.i.d. random numbers:

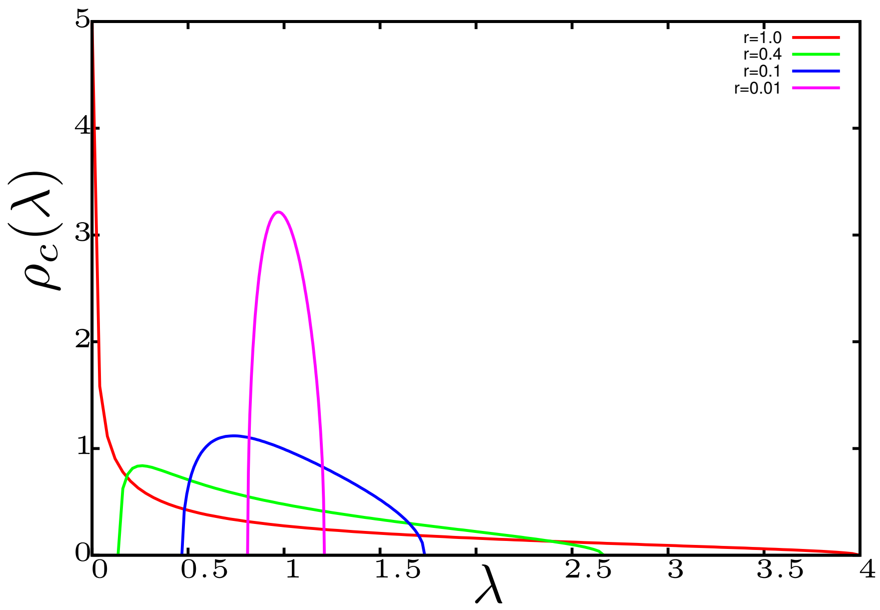
$$C_{ij} = \langle X_i X_j \rangle = \delta_{ij} \implies \rho_C(\lambda) = \delta(\lambda - 1)$$

Question: $\rho_c(\lambda)$?

$c_{ij} \sim \delta_{ij} + \epsilon_{ij}$ where $\epsilon_{ij} \sim O(1/\sqrt{T})$

Limit: $N \to \infty$ and $r = N/T =$ const

Wishart: $\rho_c(\lambda) = \frac{1}{2\pi r \lambda}\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}$ where $\lambda_\pm = (1 \pm \sqrt{r})^2$



Problems

Direct problem: $\rho_C(\lambda) \Longrightarrow \rho_c(\lambda)$

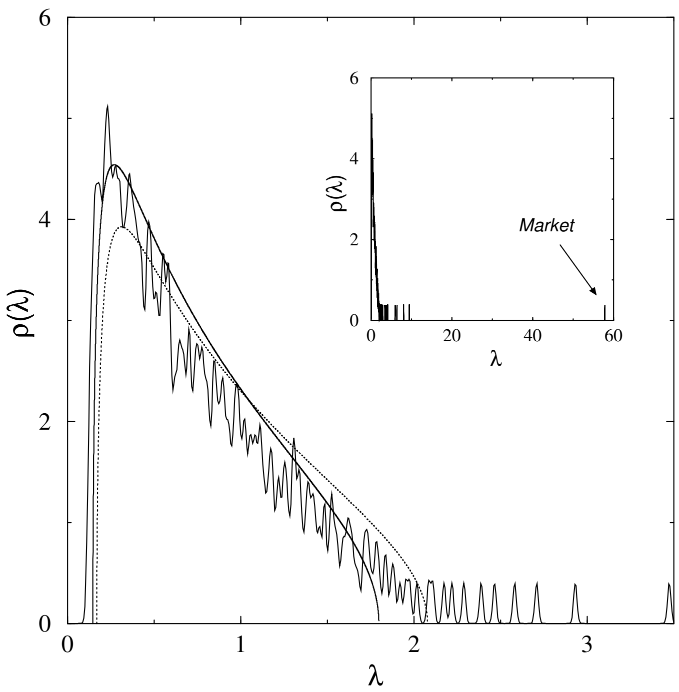
Inverse problem: $\rho_c(\lambda) \Longrightarrow \rho_C(\lambda)$

Real inverse problem: $\lambda_1, \ldots, \lambda_N \Longrightarrow \rho_C(\lambda)$

Answer depends on $r = \frac{N}{T}$: e.g. $r \to 0 \quad \Rightarrow \quad \rho_c(\lambda) \to \rho_C(\lambda)$

3 important observations

$$c_{ij} = \frac{1}{T}\sum_{t=1}^{T} x_{it}x_{jt} \quad \text{where} \quad x_{it} = \frac{R_{it} - \langle R_i \rangle}{\sigma_i}$$



from PRL 83(7), 1467 (1999) by Bouchaud, Cizeau, Laloux, Potters

Conclusions: $\mathbf{C}$ not $\mathbf{c}$, spikes = sectors, **universality** of the bulk (Wishart)

Random Matrix Theory

- many-body quantum systems **E. Wigner**
- mesoscopic systems
- localization theory
- glassy systems
- chaos
- QCD - Dirac operator
- color expansion $1/N$ (planar diagrammatics)
- counting (knots, pseudoknots enumeration)
- 2d quantum gravity, non-critical strings
- Riemann hypothesis
- sui generis branch of mathematical physics
- multivariate analysis

Physics and portfolio

- J. P. Bouchaud, P. Cizeau, L. Laloux, M. Potters, *Int. J. Theor. App. Finance* **3** (2000) 391.
- V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, H. E. Stanley, *Phys. Rev.* **E 65** (2002) 066126.
- S. Drożdż, J. Kwapień, F. Grummer, E. Ruf, J. Speth, *Physica* **A 299** (2001) 144.
- F. Lillo and R. Mantegna, cond-mat/0305546.
- P. Repetowicz and P. Richmond, math-ph/0411020.
- A. Utsugi, K. Ino and M. Oshikawa, cond-mat/0312643.
- G. Papp, S. Pafka, M.A. Nowak, I. Kondor, *Acta Phys. Polon.* **B 36** (2005) 2757.
- T. Guhr, B. Kälber, *J. Phys.* **A 36** (2003) 3009.
- Y. Malevergna, D. Sornette, *Physica* **A 331** (2004) 660.
- Z. Burda, A. Görlich, A. Jarosz, J. Jurkiewicz, *Physica* **A 343** (2004) 295.

RMT and covariance cleaning

Given:

$$\langle x_{it} x_{jt'} \rangle = C_{ij} \delta_{tt'}$$

Searched:

$$\rho_c(\lambda) = \left\langle \frac{1}{N} \sum_i \delta(\lambda - \lambda_i) \right\rangle$$

RMT and covariance cleaning

Given:

$$\langle x_{it} x_{jt'} \rangle = C_{ij} \delta_{tt'}$$

Searched:

$$\rho_c(\lambda) = \left\langle \frac{1}{N} \sum_i \delta(\lambda - \lambda_i) \right\rangle$$

Method:

$$g(z) = \frac{1}{N} \left\langle \mathrm{Tr} \frac{1}{z - c} \right\rangle = \frac{1}{N} \left\langle \sum_{i=1}^{N} \frac{1}{z - \lambda_i} \right\rangle$$

$$\frac{1}{x + i0^+} = PV \frac{1}{x} - i\pi\delta(x)$$

$$\rho_c(x) = -\frac{1}{\pi} \mathrm{Im}\, g(x + i0^+)$$

RMT and covariance cleaning

Given:

$$\langle x_{it} x_{jt'} \rangle = C_{ij} \delta_{tt'}$$

Searched:

$$\rho_c(\lambda) = \left\langle \frac{1}{N} \sum_i \delta(\lambda - \lambda_i) \right\rangle$$

Method:

$$g(z) = \frac{1}{N} \left\langle \mathrm{Tr} \frac{1}{z - c} \right\rangle = \frac{1}{N} \left\langle \sum_{i=1}^{N} \frac{1}{z - \lambda_i} \right\rangle$$

$$\frac{1}{x + i0^+} = PV \frac{1}{x} - i\pi\delta(x)$$

$$\rho_c(x) = -\frac{1}{\pi} \mathrm{Im}\, g(x + i0^+)$$

Universality: Gaussian theory $\Longrightarrow$ perturbation theory $\Longrightarrow$ Feynman diagrams $\Longrightarrow$ planar diagrams for large N

Moments generating functions:

$$M(z) = \sum_{k=1}^{\infty} \frac{M_{\mathbf{C}k}}{z^k}, \quad m(z) = \sum_{k=1}^{\infty} \frac{m_{\mathbf{c}k}}{z^k},$$

$$M(z) = zG(z) - 1 \quad \text{and} \quad m(z) = zg(z) - 1$$

Solution (Conformal map)

$$m(z) = M(Z)$$

where

$$Z = \frac{z}{1 + rm(z)}$$

or

$$z = Z(1 + rM(Z))$$

Noise dressing of moments

$$\begin{aligned} m_1 &= M_1 \\ m_2 &= M_2 + r M_1^2 \\ m_3 &= M_3 + 3r M_1 M_2 + r^2 M_1^3 \\ \ldots \end{aligned}$$

$$\begin{aligned} M_1 &= m_1 \\ M_2 &= m_2 - r m_1^2 \\ M_3 &= m_3 - 3r m_1 m_2 + 2r^2 m_1^3 \\ \ldots & . \end{aligned}$$

Summary

Portfolio/RMT

Exact relation:

$$m(z) = M(Z) \quad \text{where} \quad Z = \frac{z}{1 + rm(z)}$$

Generalizations: temporal correlations and heavy tails

Applications:

- multivariate statistics
- quantitative finance
- telecommunication
- biology
- physics (lattice MC)
- ...

Acknowledgments

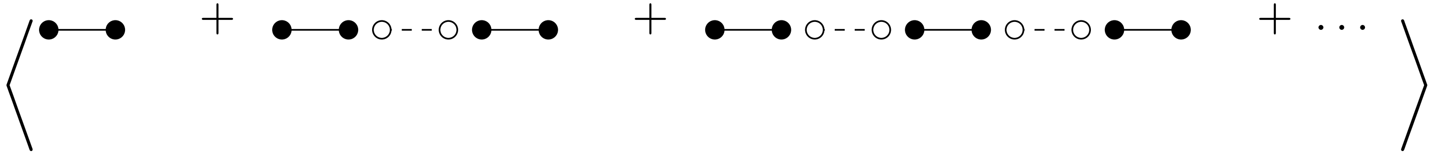
Andrzej Görlich, Andrzej Jarosz, Jerzy Jurkiewicz, Maciej A. Nowak, Gabor Papp, Bartlomiej Waclaw, Ismail Zahed

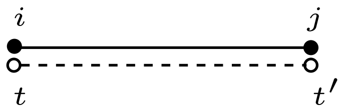
- Z. Burda, J. Jurkiewicz, M.A. Nowak, *Is Econophysics a Solid Science?*, Acta Phys. Polon. **B34** (2003) 87.
- Z. Burda, A. Görlich, A. Jarosz, J. Jurkiewicz, *Signal and Noise in Correlation Matrix*, Physica **A343** (2004) 295.
- Proceedings of the Conference: Applications of Random Matrices to Economy and Other Complex Systems, Krakow, 25-28.05.2005 *Acta Phys. Polon.* **B 36** (2005) 2757.

Planar diagrammatics

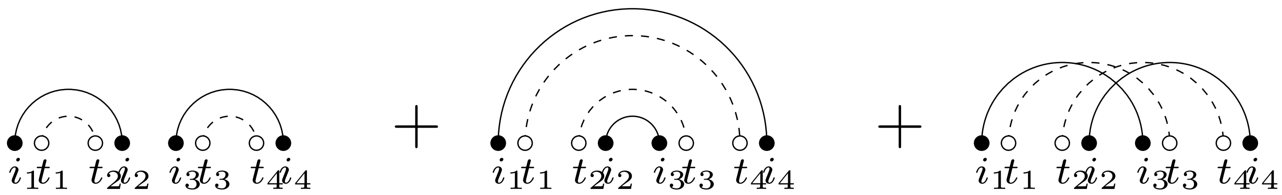
$$\mathbf{g}(z) = \left\langle \frac{1}{z-\mathbf{c}} \right\rangle = \left\langle \sum_{k\geq 0} \frac{\mathbf{c}^k}{z^{k+1}} \right\rangle$$

$$= \left\langle \frac{1}{z} + \frac{1}{z}\mathbf{x}\frac{1}{T}\mathbf{x}^\tau\frac{1}{z} + \frac{1}{z}\mathbf{x}\frac{1}{T}\mathbf{x}^\tau\frac{1}{z}\mathbf{x}\frac{1}{T}\mathbf{x}^\tau\frac{1}{z} + \ldots \right\rangle$$

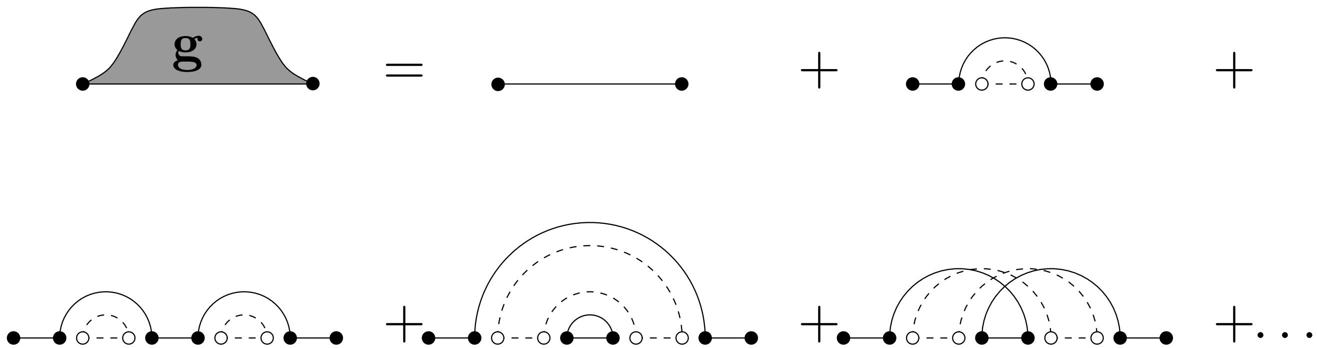
$=$ 

$\langle x_{it} x_{jt'} \rangle = C_{ij}\delta_{tt'} =$ 

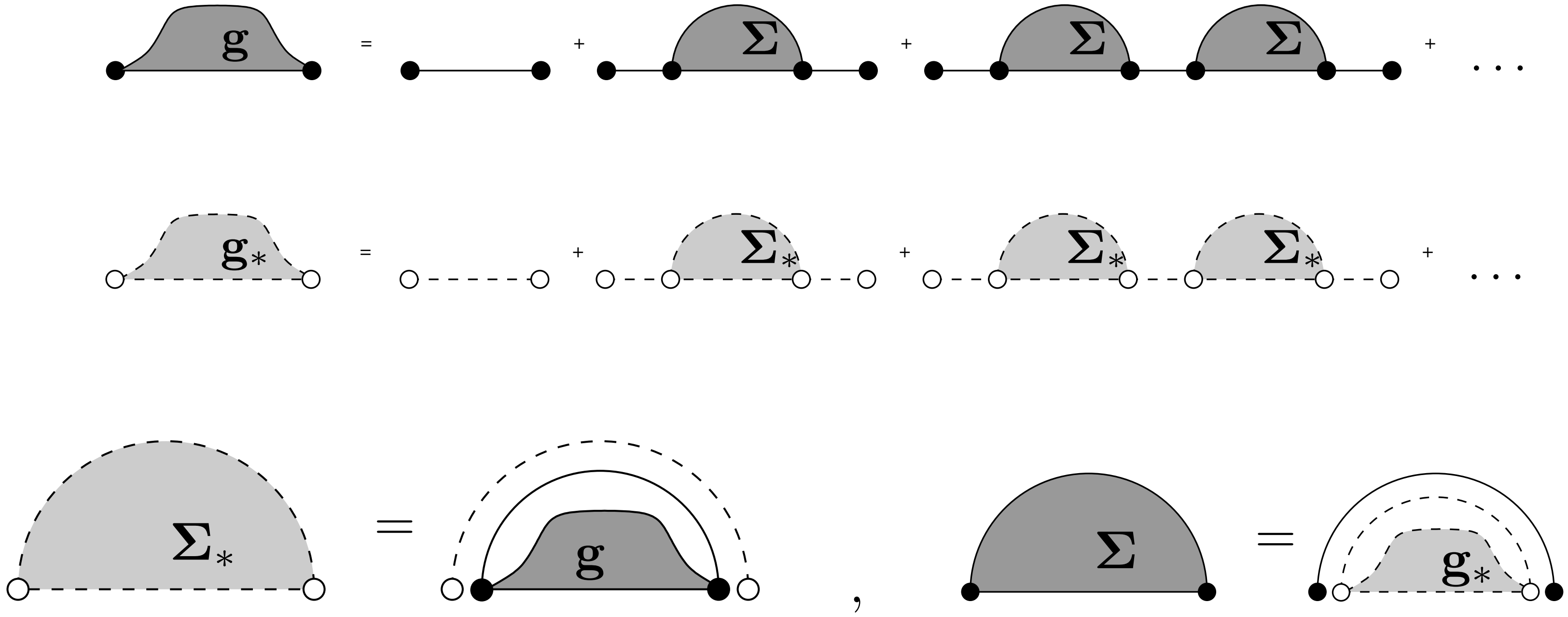
$$\begin{aligned} \langle x_{i_1t_1} x_{i_2t_2} x_{i_3t_3} x_{i_4t_4} \rangle &= \langle x_{i_1t_1} x_{i_2t_2} \rangle \langle x_{i_3t_3} x_{i_4t_4} \rangle \\ &+ \langle x_{i_1t_1} x_{i_3t_3} \rangle \langle x_{i_2t_2} x_{i_4t_4} \rangle \qquad (1) \\ &+ \langle x_{i_1t_1} x_{i_4t_4} \rangle \langle x_{i_2t_2} x_{i_3t_3} \rangle \end{aligned}$$



Diagrammatic representation of $\mathbf{g(z)}$



Closed set of equations



$$
\begin{aligned}
\mathbf{g}(z) &= \frac{1}{z - \mathbf{\Sigma}(z)} \\
\mathbf{g}_*(z) &= \frac{1}{T - \mathbf{\Sigma}_*(z)} \\
\mathbf{\Sigma}(z) &= \mathbf{C}\mathrm{Tr}\left[\mathbf{g}_*(z)\right] \\
\mathbf{\Sigma}_*(z) &= \mathrm{Tr}\left[\mathbf{g}(z)\mathbf{C}\right]
\end{aligned}
$$

Additional information in the inverse problem

K sektors: $M(Z) = \sum_{k=1}^{K} \frac{p_k \Lambda_k}{Z - \Lambda_k} \Longrightarrow M_k \Longrightarrow m_k$

$$\chi^2 = \sum_{k=1}^{L} \left(\frac{m_k^{th}(p_j, \Lambda_j) - m_k^{exp}}{\Delta_k} \right)^2$$

Example: $T = 333$, $N = 100$, $K = 3$,

$p_1 = p_2 = p_3 = 1/3$,

$\Lambda_1 = 1$, $\Lambda_2 = 2$, $\Lambda_3 = 3$ (unknown)