Random Correlation Matrices, Top Eigenvalue with Heavy Tails and Financial Applications

J.P Bouchaud
with: M. Potters, G. Biroli, L. Laloux, M. A. Miceli

CAPITAL FUND MANAGEMENT

http://www.cfm.fr

Portfolio theory: Basics

- Portfolio weights w_i

- Risk: variance of the portfolio returns

$$R^2 = \sum_{ij} w_i \sigma_i C_{ij} \sigma_j w_j$$

  where σ_i^2 is the variance of asset i and C_{ij} is the correlation matrix.

- If predicted gains are g_i then the expected gain of the portfolio is $G = \sum w_i g_i$.

Empirical Correlation Matrix

- Large set of Assets (N) and (comparable) set of data points (T)

- Empirical Variance

$$\sigma_i^2 = \frac{1}{T} \sum_t \left(X_i^t \right)^2$$

relative square-error is $(2 + \kappa)/T$

- Empirical Equal-Time Correlation Matrix

$$E_{ij} = \frac{1}{T} \sum_t \frac{X_i^t X_j^t}{\sigma_i \sigma_j}$$

order N^2 quantities estimated with NT datapoints. If $T < N$ $\mathbf{E}$ has rank $T < N$, not even invertible.

Markowitz Optimization

- Find the portfolio with maximum expected return for a given risk or equivalently, minimum risk for a given return (G)

- In matrix notation:

$$\mathbf{w}_C = G \frac{\mathbf{C}^{-1}\mathbf{g}}{\mathbf{g}^T \mathbf{C}^{-1}\mathbf{g}}$$

- Where all returns are measured with respect to the risk-free rate and $\sigma_i = 1$ (absorbed in g_i).

- Non-linear problem: $\sum_i |w_i| \leq A$ – a spin-glass problem!

Risk of Optimized Portfolios

- Let $\mathbf{E}$ be an noisy estimator of $\mathbf{C}$ such that $\langle \mathbf{E} \rangle = \mathbf{C}$

- "In-sample" risk

$$R_{\text{in}}^2 = \mathbf{w}_E^T \mathbf{E} \mathbf{w}_E = \frac{G^2}{\mathbf{g}^T \mathbf{E}^{-1} \mathbf{g}}$$

- True minimal risk

$$R_{\text{true}}^2 = \mathbf{w}_C^T \mathbf{C} \mathbf{w}_C = \frac{G^2}{\mathbf{g}^T \mathbf{C}^{-1} \mathbf{g}}$$

- "Out-of-sample" risk

$$R_{\text{out}}^2 = \mathbf{w}_E^T \mathbf{C} \mathbf{w}_E = \frac{G^2 \mathbf{g}^T \mathbf{E}^{-1} \mathbf{C} \mathbf{E}^{-1} \mathbf{g}}{(\mathbf{g}^T \mathbf{E}^{-1} \mathbf{g})^2}$$

Risk of Optimized Portfolios

- Using convexity arguments, and for large matrices:

$$R^2_{\text{in}} \leq R^2_{\text{true}} \leq R^2_{\text{out}}$$

- Importance of eigenvalue cleaning:

$$w_i \propto \sum_{kj} \lambda_k^{-1} V_i^k V_j^k g_j = g_i + \sum_{kj} (\lambda_k^{-1} - 1) V_i^k V_j^k g_j$$

  – Eigenvectors with $\lambda > 1$ are suppressed,

  – Eigenvectors with $\lambda < 1$ are enhanced. Potentially very large weight on small eigenvalues.

  – Must determine which eigenvalues to keep and which one to correct

CAPITAL FUND MANAGEMENT

J.Ph. Bouchaud

Spectrum of Wishart Ensemble

- Consider an Empirical Correlation Matrix of N assets using T data points both very large with $n = N/T$ finite.

$$E_{ij} = \frac{1}{T} \sum_{k=1}^{T} X_i^k X_j^k \qquad \text{where} \qquad \langle X_i^k X_j^l \rangle = C_{ij} \delta_{kl}$$

- We need to find the trace of the resolvent or Stieljes transform:

$$G(z) = \frac{1}{N} \mathrm{Tr} \left[(z\mathbf{I} - \mathbf{E})^{-1} \right]$$

$$\rho(\lambda) = \lim_{\epsilon \to 0} \frac{1}{\pi} \Im \left(G(\lambda - \mathrm{i}\epsilon) \right).$$

Null hypothesis $\mathbf{C} = \mathbf{I}$

- E_{ij} is a sum of (rotationally invariant) matrices $E^k_{ij} = (X^k_i X^k_j)/T$

- Free random matrix theory: Find the additive R-transform $R(x) = B(x) - 1/x$; $B(G(z)) = z$)

$$G_k(z) = \frac{1}{N}\left(\frac{1}{z-n} + \frac{N-1}{z}\right)$$

- defining $n = N/T$, inverting $G_k(z)$ to first order in $1/N$,

$$R_k(x) = \frac{1}{T(1-nx)} \quad \text{by additivity} \quad R_E(x) = \frac{1}{(1-nx)}$$

$$G_E(z) = \frac{(z+n-1) - \sqrt{(z+n-1)^2 - 4zn}}{2zn}$$

Null hypothesis $\mathbf{C} = \mathbf{I}$

$$\rho(\lambda) = \frac{\sqrt{4\lambda n - (\lambda + n - 1)^2}}{2\pi\lambda n} \qquad \lambda \in [(1-\sqrt{n})^2, (1+\sqrt{n})^2]$$

Marcenko-Pastur (1967) (and many rediscoveries)

- Any eigenvalue beyond the Marcenko-Pastur band can be deemed to contain some information (but see below)

General $\mathbf{C}$ Case

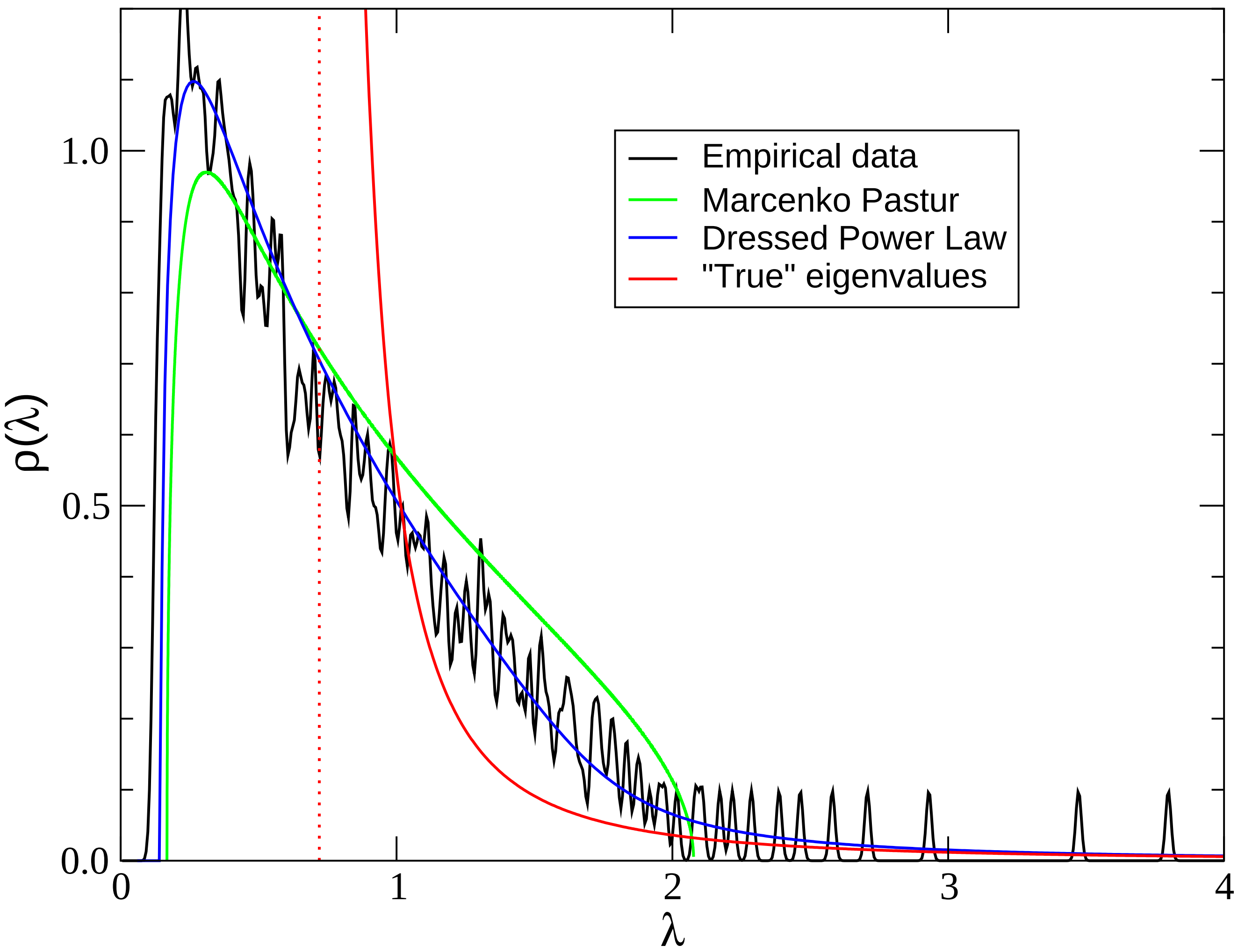
- The general case for $\mathbf{C}$ cannot be directly written as a sum of "Blue" functions.

- Solution using different techniques (replicas, diagrams, S-transform:

$$zG_E(z) = ZG_C(Z) \qquad \text{where} \qquad Z = \frac{z}{1 + n(zG_E(z) - 1)}$$

- For stocks, one large eigenvalue – the "market" – and several sectors

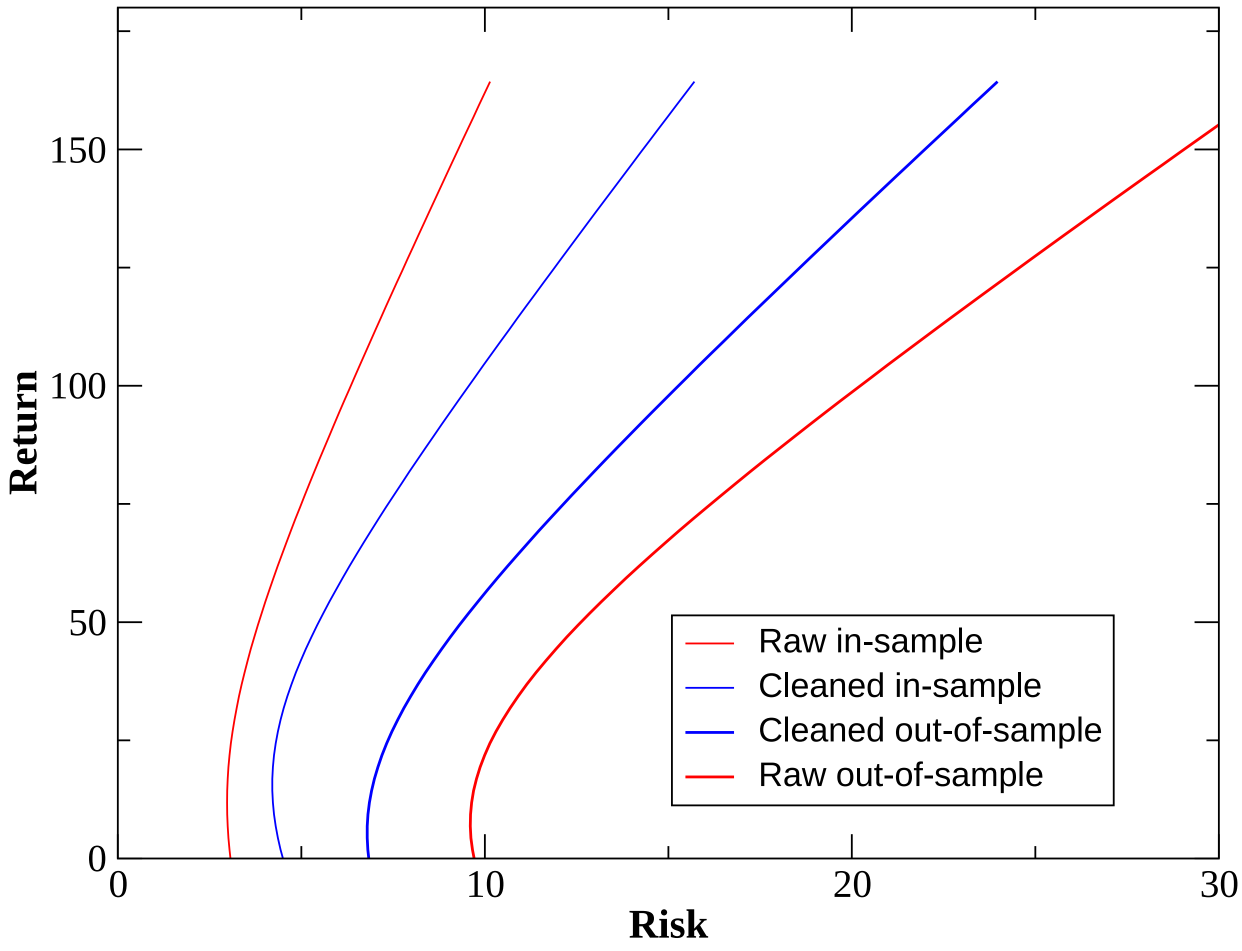
Empirical Correlation Matrix



CAPITAL FUND MANAGEMENT

J.Ph. Bouchaud

Matrix Cleaning



General Correlation matrices

- Non equal time correlation matrices

$$E_{ij}^{\tau} = \frac{1}{T} \sum_t \frac{X_i^t X_j^{t+\tau}}{\sigma_i \sigma_j}$$

$N \times N$ but not symmetrical: 'leader-lagger' relations

- General rectangular correlation matrices

$$G_{\alpha i} = \frac{1}{T} \sum_{t=1}^{T} Y_\alpha^t X_i^t$$

N 'input' factors X; M 'output' factors Y

– Example: $Y_\alpha^t = X_j^{t+\tau}$, $N = M$

Singular values and relevant correlations

- Singular values: Square root of the non zero eigenvalues of GG^T or G^TG, with associated eigenvectors u_α^k and v_i^k $\rightarrow$ $1 \geq s_1 > s_2 > ... s_{(M,N)^-} \geq 0$

- Interpretation: $k = 1$: best linear combination of input variables with weights v_i^1, to optimally predict the linear combination of output variables with weights u_α^1, with a cross-correlation $= s_1$.

- s_1: measure of the predictive power of the set of Xs with respect to Ys

- Other singular values: orthogonal, less predictive, linear combinations

CAPITAL FUND MANAGEMENT

Benchmark: no cross-correlations

- Null hypothesis: No correlations between Xs and Ys – $\langle G \rangle = \mathbf{0}$

- But arbitrary correlations *among* Xs, C_X, and Ys, C_Y, are possible

- Consider exact normalized principal components for the sample variables Xs and Ys:

$$\hat{X}_i^t = \frac{1}{\sqrt{\lambda_i}} \sum_j U_{ij} X_j^t; \qquad \hat{Y}_\alpha^t = ...$$

and define $\hat{G} = \hat{Y}\hat{X}^T$.

Benchmark: no cross-correlations

- Tricks:
  - Non zero eigenvalues of $\widehat{G}\widehat{G}^T$ are the same as those of $\widehat{X}^T\widehat{X}\widehat{Y}^T\widehat{Y}$
  - $A = \widehat{X}^T\widehat{X}$ and $B = \widehat{Y}^T\widehat{Y}$ are mutually free, with n (m) eigenvalues equal to 1 and $1-n$ $(1-m)$ equal to 0
  - "S-transforms" are multiplicative

Technicalities

- $$\eta_A(y) \equiv \frac{1}{T}\mathrm{Tr}\frac{1}{1+yA}.$$

- $$\Sigma_A(x) \equiv -\frac{1+x}{x}\eta_A^{-1}(1+x).$$

- $$\eta_A(y) = 1 - n + \frac{n}{1+y}, \qquad \eta_B(y) = 1 - m + \frac{m}{1+y}.$$

- $$\Sigma_{GG}(x) = \Sigma_A(x)\Sigma_B(x) = \frac{(1+x)^2}{(x+n)(x+m)}.$$

Benchmark: Random SVD

- Final result:([LL,MAM,MP,JPB])

$$\rho(s) = (1-n,1-m)^+\delta(s)+(m+n-1)^+\delta(s-1)+\frac{\sqrt{(s^2-\gamma_-)(\gamma_+-s^2)}}{\pi s(1-s^2)}$$

with

$$\gamma_\pm = n+m-2mn \pm 2\sqrt{mn(1-n)(1-m)}, \quad 0 \le \gamma_\pm \le 1$$

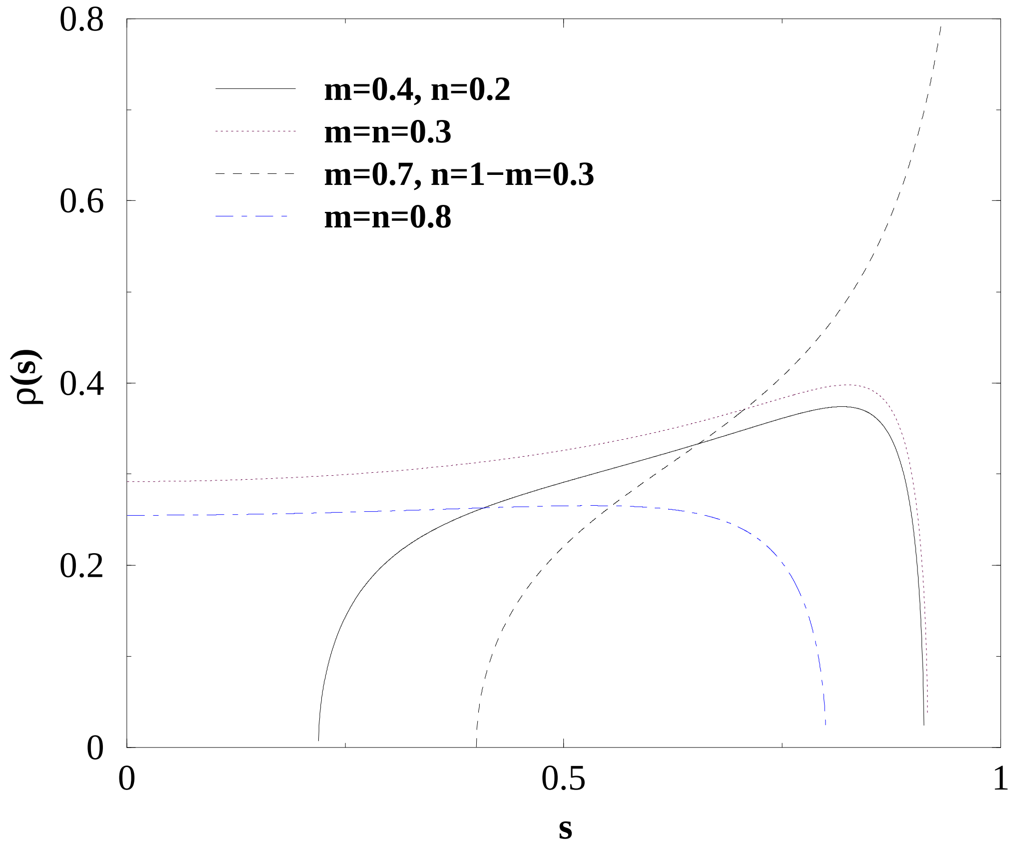
- Analogue of the Marcenko-Pastur result for rectangular correlation matrices

- Many applications; finance, econometrics ('large' models), genomics, etc.

Benchmark: Random SVD

- Simple cases:
  - $n = m,\ s \in [0, 2\sqrt{n(1-n)}]$
  - $n, m \to 0,\ s \in [|\sqrt{m} - \sqrt{n}|, \sqrt{m} + \sqrt{n}]$
  - $m = 1,\ s \to \sqrt{1-n}$
  - $m \to 0,\ s \to \sqrt{n}$

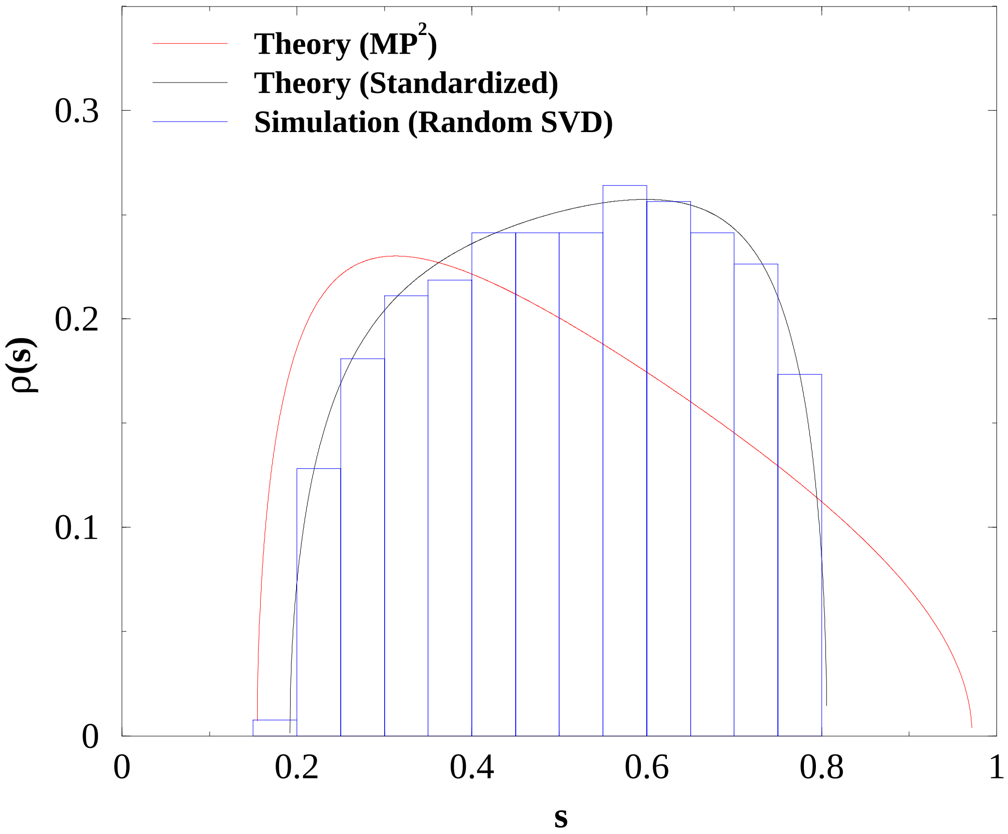
RSVD: Numerical illustration



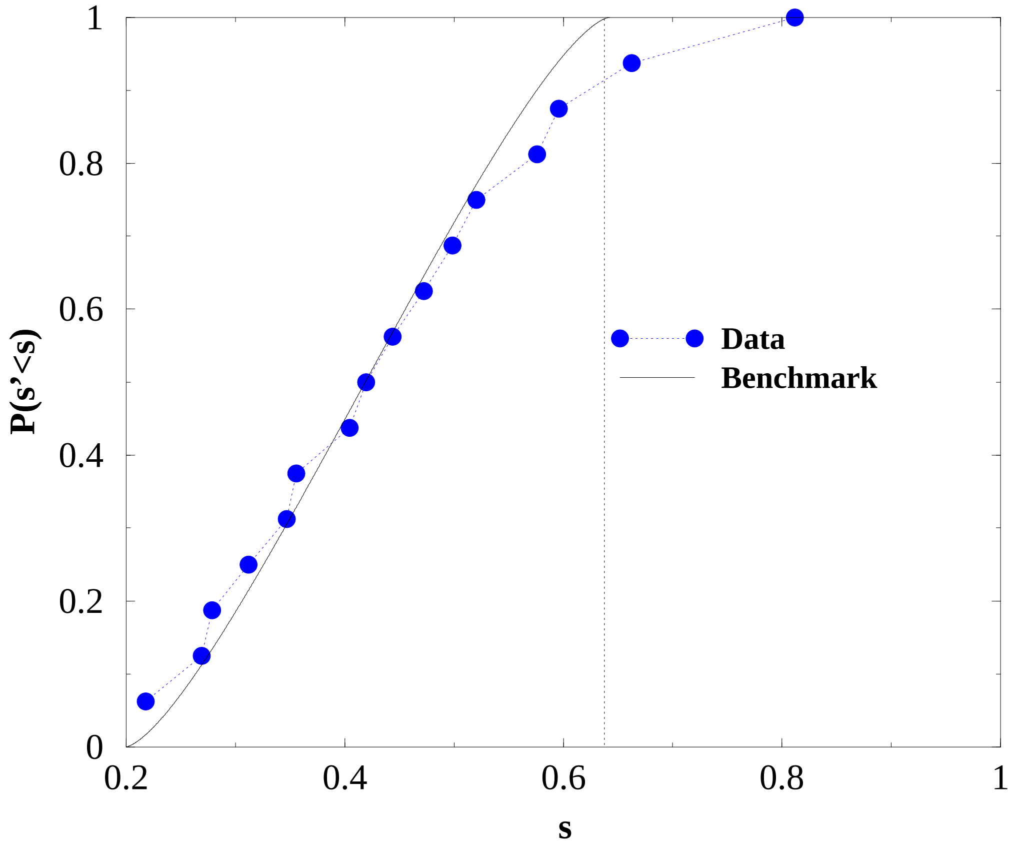
CAPITAL FUND MANAGEMENT

J.Ph. Bouchaud

RSVD: Numerical illustration



Inflation vs. Economic indicators



$N = 50,\ M = 16,$

$T = 265.$

CAPITAL FUND MANAGEMENT

J.Ph. Bouchaud

Statistics of the Top Eigenvalue

- All previous results are true when $N, M, T \to \infty$ with fixed n, m
- How far is the top eigenvalue expected to leak out at finite N?
- Precise answer when matrix elements are iid Gaussian: Tracy-Widom statistics
- Width of the smoothed edge: $N^{-2/3}$
- Relation with the directed polymer problem + many others

Statistics of the Top Eigenvalue

- Exceptions
  - 'Strong' Rank One Perturbation $\to$ emergence of an isolated eigenvalue with *Gaussian*, $N^{-1/2}$ fluctuations (Baik, Ben-Arous, Péché)
  - E.g.: $E_{ij} \to E_{ij} + \rho(1 - \delta_{ij})$ leads to a *market mode* $\lambda_{\text{max}} \approx N\rho$
  - Fat tailed distribution of matrix elements

Fat tails and Top Eigenvalue: Wigner Case

- Eigenvalue statistics of large real symmetric matrices with iid elements X_{ij}, $P(x) \sim |X|^{-1-\mu}$

- Eigenvalue density:
  - $\mu > 2 \to$ Wigner semi-circle in $[-2, 2]$
  - $\mu < 2 \to$ unbounded density with tails $\rho(\lambda) \sim \lambda^{-1-\mu}$

- Note: $\mu < 2$ non trivial statistics of eigenvectors (localized/delocalized) (Cizeau,JPB)

Fat tails and Top Eigenvalue: Wigner Case

- Largest Eigenvalue statistics ([GB,MP,JPB])
  - $\mu > 4$: $\lambda_{\max} - 2 \sim N^{-2/3}$ with a Tracy-Widom distribution (max of strongly correlated variables)
  - $2 < \mu < 4$: $\lambda_{\max} \sim N^{\frac{2}{\mu}-\frac{1}{2}}$ with a *Fréchet* distribution (although the density goes to zero when $\lambda > 2$!!)
  - $\mu = 4$: $\lambda_{\max} \geq 2$ but remains $O(1)$, with a new distribution:

$$P_>(\lambda_{\max}) = w\theta(\lambda_{\max} - 2) + (1-w)F(s) \quad \lambda_{\max} = s + \frac{1}{s}$$

- Note: The case $\mu > 4$ still has a power-law tail for finite N, of amplitude $N^{2-\mu/2}$

Fat tails and Correlation Matrices

- $$E_{ij} = \frac{1}{T} \sum_t X_i^t X_j^t$$

- $\mu > 4$: $\lambda_{\max} - (1 + \sqrt{n})^2 \sim N^{-2/3}$ (but with a power-law tail as above)

- $\mu < 4$: $\lambda_{\max} \sim N^{\frac{4}{\mu}-1} n^{1-2/\mu}$

- Fat tails induce fictitious 'strong' correlations – important for applications in finance where $\mu \approx 3 - 5$.

EWMA Empirical Correlation Matrices

- Consider the case where the Empirical matrix is often computed using an exponentially weighted moving average (EWMA) with $\epsilon = 1/T$

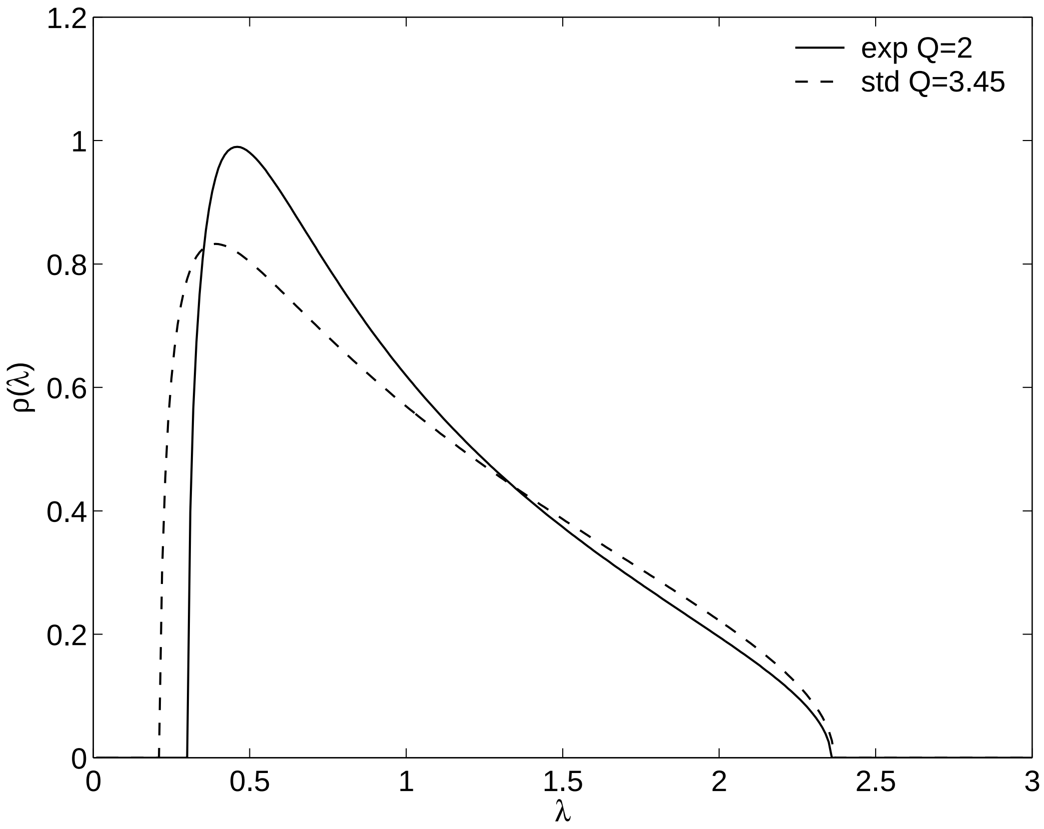
$$E_{ij} = \epsilon \sum_{k=0}^{\infty} (1-\epsilon)^k X_i^k X_j^k \qquad \text{where} \qquad \langle X_i^k X_j^l \rangle = \delta_{ij}\delta_{kl}$$

- Above trick based R-functions still works:

$$\rho(\lambda) = \frac{1}{\pi}\Im G(\lambda) \quad \text{where } G(\lambda) \text{ solves} \quad \lambda n G = n - \log(1 - nG)$$

CAPITAL FUND MANAGEMENT

EWMA Empirical Correlation Matrices



Spectrum of the exponentially weighted random matrix with $n = 1/2$ and the spectrum of the standard random matrix with $n \equiv N/T = 1/3.45$.

Dynamics of the top eigenvector

- Specific dynamics of large top eigenvalue and eigenvector: Ornstein-Uhlenbeck processes (on the unit sphere for $\mathbf{V}^1$)

- The angle obeys the following SDE:

$$d\theta \approx -\frac{\epsilon}{2}\sin 2\theta dt + \zeta_t\, dW_t$$

with

$$\zeta_t^2 \approx \epsilon^2 \left[\frac{1}{2}\sin^2 2\theta_t + \frac{\Lambda_1}{\Lambda_0}\cos^2 2\theta_t\right]$$

- Eigenvector dynamics:

$$\left\langle \langle \psi_{0t+\tau}|\psi_{0t}\rangle \right\rangle \approx E(\cos(\theta_t - \theta_{t+\tau})) \approx 1 - \epsilon\frac{\Lambda_1}{\Lambda_0}(1 - \exp(-\epsilon\tau))$$

The variogram of the top eigenvector