3. Free Random Matrices and applications

J.P Bouchaud

CAPITAL FUND MANAGEMENT

http://www.cfm.fr

Free Random Matrices

- Freeness

  – Freeness is the generalisation of independence for matrices. Two matrices $\mathbf{A}$, $\mathbf{B}$ are said to be free essentially if the eigenvectors of $\mathbf{A}$ are a random rotation of those of $\mathbf{B}$.

  – Examples: $\mathbf{A}$, $\mathbf{B}$ sym. and fixed, $\mathbf{H}$ random GOE matrix, $\mathbf{O}$ a random rotation

$$\mathbf{A} \text{ and } \mathbf{H}; \qquad \mathbf{A} \text{ and } \mathbf{O}^t\mathbf{B}\mathbf{O}; \qquad \mathbf{H_1} \text{ and } \mathbf{H_2}$$

  – Rectangular matrix examples: $\mathbf{A}$ $N \times T$ fixed, $\mathbf{C}$ $N \times N$ fixed, $\mathbf{H}$ $N \times T$ IID Gaussian:

$$\mathbf{A}\mathbf{A}^t \text{ and } \mathbf{H}\mathbf{H}^t; \qquad \mathbf{H_1}\mathbf{H_1^t} \text{ and } \mathbf{H_2}\mathbf{H_2^t} \qquad \mathbf{H_1}\mathbf{H_1^t} \text{ and } \mathbf{C}$$

Free Random Matrices

- Two powerful composition theorems

  - If $\mathbf{A}$, $\mathbf{B}$ are sym. and free, then the spectrum of $\mathbf{A}+\mathbf{B}$ is such that:

$$R_{A+B}(z) = R_A(z) + R_B(z)$$

  - If $\mathbf{A}$, $\mathbf{B}$ are sym., non negative and free, then the spectrum of $\mathbf{AB}$ is such that:

$$S_{AB}(z) = S_A(z) S_B(z)$$

Convergence to the semi-circle

- CLT for matrix spectrum

- Take a 'small' matrix H with a centred spectrum and expand $G(z)$ in $1/z$:

$$G(z) = \frac{1}{z} + 0 + \epsilon^2 \frac{1}{z^3} + O(\epsilon^3/z^4) \to \frac{1}{z} \approx G - \epsilon^2 G^3$$

$$B(z) \approx \frac{1}{z - \epsilon^2 z^3} \to R(z) = B(z) - \frac{1}{z} \approx \epsilon^2 z + O(\epsilon^3 z^2)$$

- Now add M such free matrices with $\epsilon = M^{-1/2}$ and $M \to \infty$, then

$$R_M(z) = M\epsilon^2 z + O(M\epsilon^3 z^2) \to_{M \to \infty} z$$

Convergence to the semi-circle

The sum of M ‘small’ centred matrices has a Wigner spectrum in the large M limit, with computable corrections

The Marcenko-Pastur distribution

- Consider the following empirical $N \times N$ correlation matrix

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

- When $\mathbf{C} = \mathbf{1}$, E_{ij} is a sum of rotationally invariant projectors $(X_i^k X_j^k)/T$

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

- Inverting $G_k(z)$ to first order in $1/N$,

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

which is the R-transform of the MP distribution

The EMA Marcenko-Pastur distribution

- Consider the case where $\mathbf{C} = \mathbf{1}$ with an Empirical matrix computed using an exponentially weighted moving average with $\alpha = 1 - q/N$:

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

- In law, E_{ij} satisfies $E_{ij} = \alpha E_{ij} + (1-\alpha) X_i^0 X_j^0$.

- The R-transform of the extra piece is

$$R_0(x) = \frac{q}{N(1-qx)}$$

EWMA Empirical Correlation Matrices

- Now, using: $R_{aA}(x) = aR_A(ax)$,

$$R_E(x) = R_{\alpha E}(x) + R_0(x) = (1-q/N)R_E((1-q/N)x) + \frac{q}{N(1-qx)}$$

- To first order in $1/N$

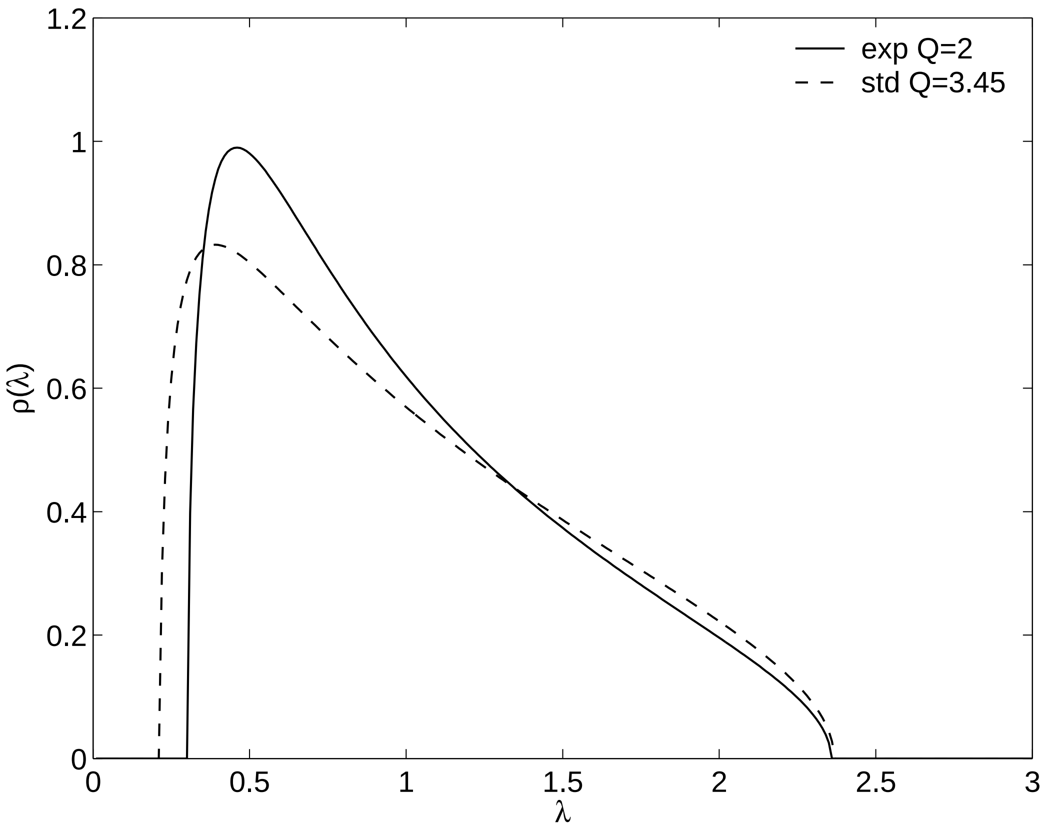
$$R(x) + xR'(x) + \frac{q}{1-qx} = 0 \quad \text{sol:} \quad R(x) = -\frac{\log(1-qx)}{qx}$$

- Going back to the resolvent to find the density

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

- $\rho(\lambda \to 0) \sim \exp(-1/q)$ when $q \to 0$.

EWMA Empirical Correlation Matrices



Spectrum of the exponentially weighted random matrix with $q \equiv (N(1-\alpha)) = 1/2$ and the spectrum of the standard random matrix with $q \equiv N/T = 1/3.45$.

General $\mathbf{C}$ Case

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

- But the spectrum of $\mathbf{X}\mathbf{C}\mathbf{X}^{\mathrm{t}}$ is the same as that of $\mathbf{C}\mathbf{X}^{\mathrm{t}}\mathbf{X}$

- Using S-transforms:

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

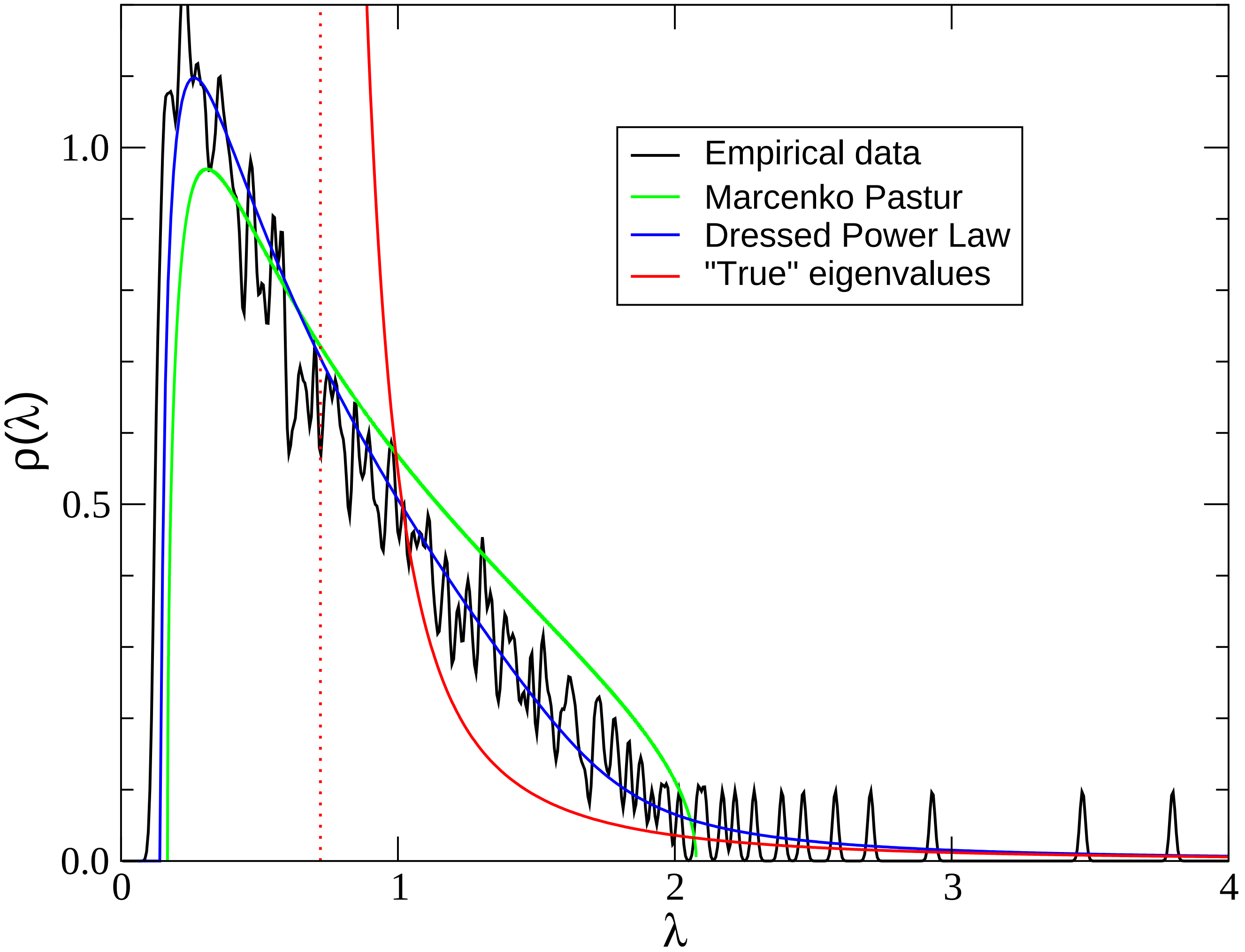
- Other equivalent expression:

$$G_E(z) = \int d\lambda\, \rho_C(\lambda) \frac{1}{z - \lambda(1 - q + qzG_E(z))}$$

General $\mathbf{C}$ Case

- Check: $\rho_C(\lambda) = \delta(\lambda - 1)$: $qzG_E^2 - (z + q - 1)G_E + 1 = 0$, OK

Empirical Correlation Matrix



The Student (Elliptic) Ensemble

- On a given day, the volatility is a random variable:

$$E_{ij} = \frac{1}{T}\sum_{k=1}^{T} \sigma_k^2 X_i^k X_j^k$$

- When $\mathbf{C} = \mathbf{1}$, E_{ij} is a sum of rotationally invariant projectors $(X_i^k X_j^k)/T$

$$G_k(z) = \frac{1}{N}\left(\frac{1}{z - \sigma_k^2 q} + \frac{N-1}{z}\right)$$

- Write $\sigma^2 \equiv \mu/s$, and $P(s) = s^{\mu/2-1}e^{-s}/\Gamma(\mu/2)$. Find the additive R-transform, from which the Blue function is found:

$$B(x) = \frac{1}{x} + \frac{1}{T}\sum_t \frac{\frac{\mu}{s_t}}{(1 - \frac{qx\mu}{s_t})} = \frac{1}{x} + \int ds P(s)\frac{\frac{\mu}{s}}{(1 - \frac{qx\mu}{s})}$$

The Student (Elliptic) Ensemble

- Inverting this relation in terms of G leads to:

$$\lambda = \frac{G_R}{G_R^2 + \pi^2\rho^2} + \int ds P(s) \frac{\mu(s - \mu G_R/Q)}{(s - \mu G_R/Q)^2 + \pi^2\rho^2} \quad (1)$$

$$0 = \rho\left(-\frac{1}{G_R^2\pi^2\rho^2} + \int ds P(s) \frac{\mu^2/Q}{(s - \mu G_R/Q)^2 + \pi^2\rho^2}\right) \quad (2)$$

The Student (Elliptic) Ensemble

- When $P(s) = s^{\mu/2-1} e^{-s}/\Gamma(\mu/2)$, one finds $\rho(\lambda) \sim \lambda^{-1-\mu/2}$.

- There is a lower edge to the spectrum

- Fit *very* well but...not the good model

- For the Maximum Likelihood Estimator of the correlation matrix for the Student ensemble, one recovers the MP spectrum

- The case $\mathbf{C} \neq \mathbf{1}$ can be treated using S-transforms.

More 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 $v_i^k \to$ $1 \geq s_1 > s_2 > ...s_{(M,N)^-} \geq 0$

- 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

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:

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

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

CAPITAL FUND MANAGEMENT

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

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

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

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

- $$S_{GG}(x) = S_A(x)S_B(x) = \frac{(1+x)^2}{(x+n)(x+m)}.$$

CAPITAL FUND MANAGEMENT

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

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

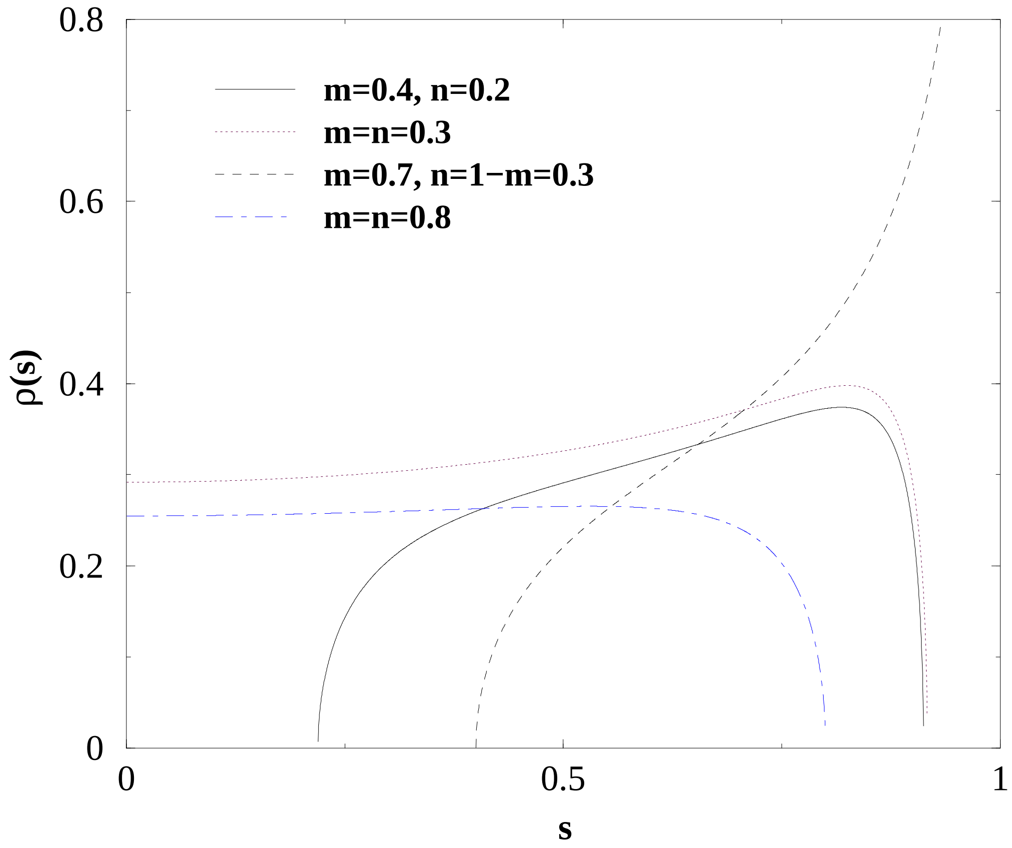
- Analogue of the Marcenko-Pastur result for rectangular correlation matrices, but different from MP^2.

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

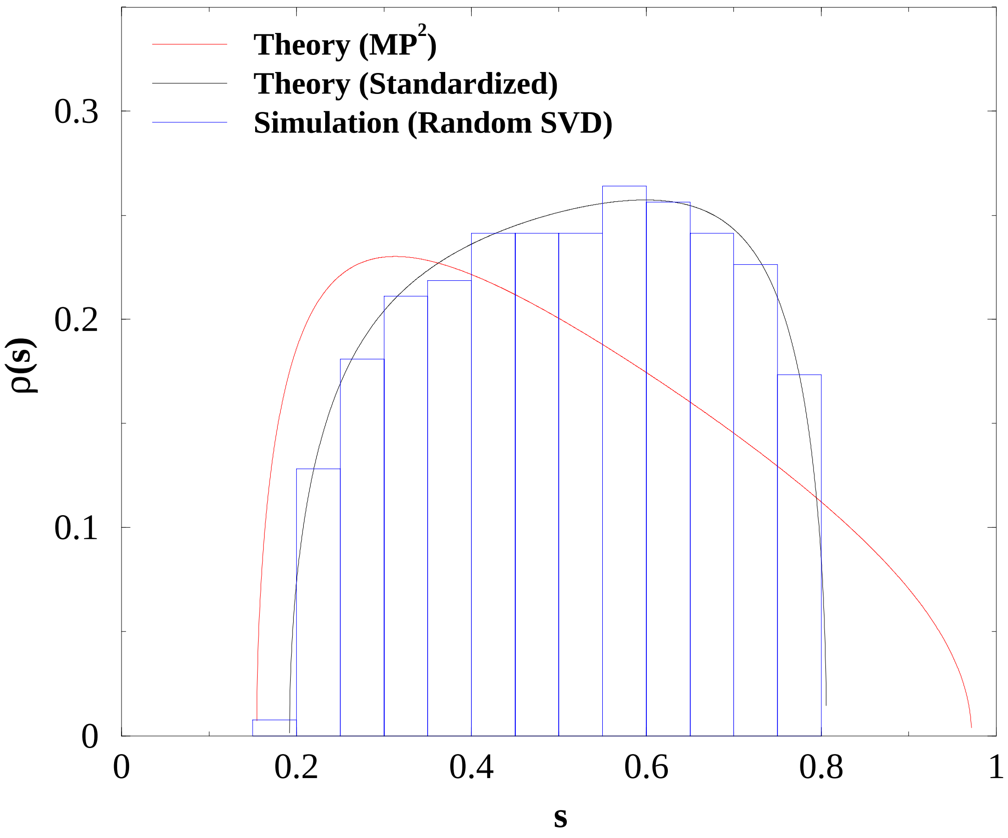
Benchmark: Random SVD

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

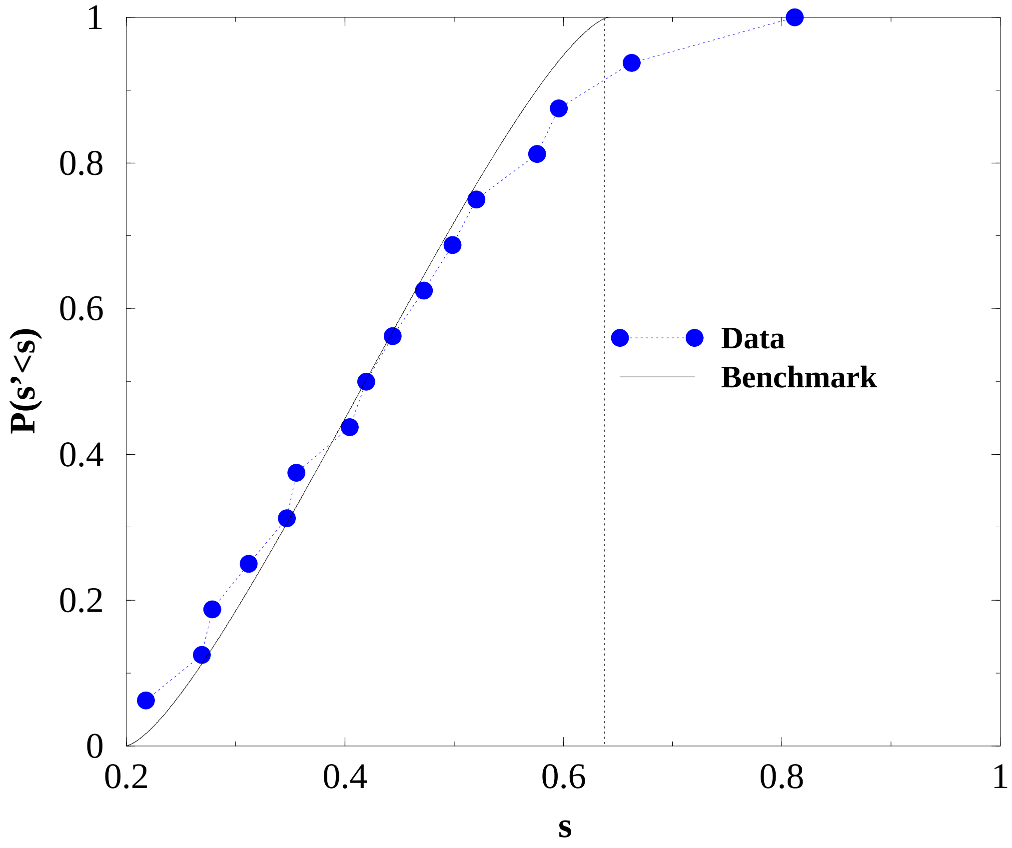
RSVD: Numerical illustration



RSVD: Numerical illustration



Inflation vs. Economic indicators



$N = 50, M = 16, T = 265$