Colle e

Sample Covariance Matrices of Multiple Time Series

- **Covariance matrix of stationary stochastic process $x_t = (x_t^a)$, $t \in \mathbb{Z}$, $1 \leq a \leq p$:**

$$C_{ij}^{ab} = \frac{1}{M} \sum_{t=1}^{M} x_{i+t}^{a} x_{j+t}^{b} = \frac{1}{M} (X X^T)_{ij}^{ab} .$$

Here $X = (x_{it})$ is $pN \times M$ matrix with entries $x_{it} = x$

Compare with Wishart–Laguerre Ensemble

- Empirical covariances for N data, evaluated on the basis of M measurements for each variable. Use $N \times M$ matrices $X = (x_{it})$ with i.i.d. entries x_{it} to compute:

$$C_{ij} = \frac{1}{M}(X X^T)_{ij} = \frac{1}{M}\sum_{t=1}^{M} x_{it}x_{jt} .$$

Expect finite sample fluctuation around mean.

$$C_{ij} = \langle x_i x_j \rangle \pm \mathcal{O}(1/\sqrt{M}) = \delta_{ij} \pm \mathcal{O}(1/\sqrt{M})$$

- Spectrum of C as $N \to \infty$, $M \to \infty$ @ fixed $\alpha = N/M$?

$\Rightarrow$ **Marčenko Pastur-Law**

$$\rho_\alpha(\lambda) = \frac{1}{2\pi\alpha\lambda}\sqrt{4\alpha - (\lambda - (1+\alpha))^2}$$

Performing the Average

- **Standard Approach – Replica Method**

$$\left\langle \ln Z_{N_p} \right\rangle = \lim_{n\to 0} \frac{1}{n} \ln \left\langle Z^n_{N_p} \right\rangle$$

- **For integer n, $Z^n_{N_p}$ is partition function of n identical copies of the system (n-th power of Gaussian integral)**

- **Experience: final result has structure of replica-symmetric high-temperature solution ⇔ annealed calculation (n**

Performing the Average (contd.)

- **Insert definition of C, and $\alpha_p = \alpha p$,**

$$\langle$$

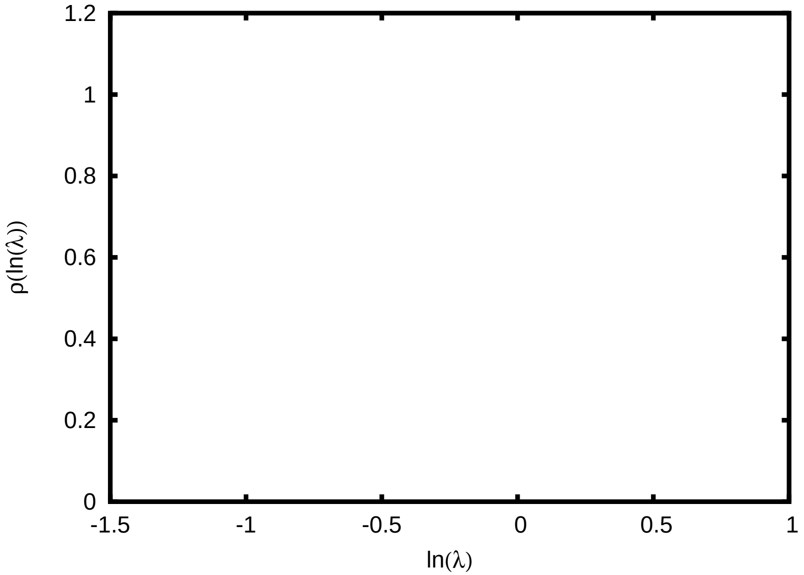
Numerical Tests

- **Spectral density for $x_n \sim \mathcal{N}(0,1)$ i.i.d. @ α = 0.1**

AR-1 Process @ $\alpha = 0.1$, $p = 1$

$$x_n = a_1 x_{n-1} + \sqrt{1 - a_1^2}\, \xi_n$$

- (Logarithmic) Spectral density for AR-1 process @ $\alpha = 0.1$



AR-1 Process @ $\alpha = 0.1$, $p = 2$ and $p = 3$

- (Logarithmic) Spectral density for AR-1 process @ $\alpha = 0.1$

$$x_n = A\, x_{n-1} + \sigma$$

Summary

- **Computed DOS of sample covariance matrices for multiple time-series using annealed calculation.**

- **Key ingredient: Szegö's theorem for (block) Toeplitz matrices**

- **Rectangular window and decorrelation approximation $\Rightarrow$ Closed form approximation.**

- **Use of Szegös theorem suggests a scaling form for DOS.**

  - **scaling is requires knowledge of a function on $\mathbb{R}^p$! DOS for i.i.d. data is insufficient.**
  - **currently working on effective methods to evaluate scaling function for $p > 1$.**

- **Lots of possible applications.**