

KING'S
College
LO

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:

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

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 $N \rightarrow \infty, M \rightarrow \infty$ @ fixed $\beta = N/M$?
 \Rightarrow **Marčenko Pastur-Law**

$$\rho(\lambda) = \frac{1}{2} \left[4 - (\lambda - (1 + \beta))^2 \right]$$

Performing the Average

- Standard Approach – Replica Method

$$\langle \ln Z_{N_p} \rangle = \lim_{n \rightarrow 0} \frac{1}{n} \ln \langle Z_{N_p}^n \rangle$$

- For integer n , $Z_{N_p}^n$ 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 \Leftrightarrow annealed calculation ($n = 1$).

$\langle \ln Z_{N_p} \rangle \simeq \ln \langle Z_{N_p} \rangle \Rightarrow$ Do annealed calculation from the start

$$\langle Z_{N_p} \rangle = \left\langle \int \prod_{k,a} \frac{du_{ka}}{2\pi i} \exp \left\{ -\frac{i}{2} \sum_{k,a,b} u_{ka} \left(\sum_{ab} C_k^{ab} \right) u_{kb} \right\} \right\rangle$$

Performing the Average (contd.)

- Insert definition of C , and $p = p$,

<

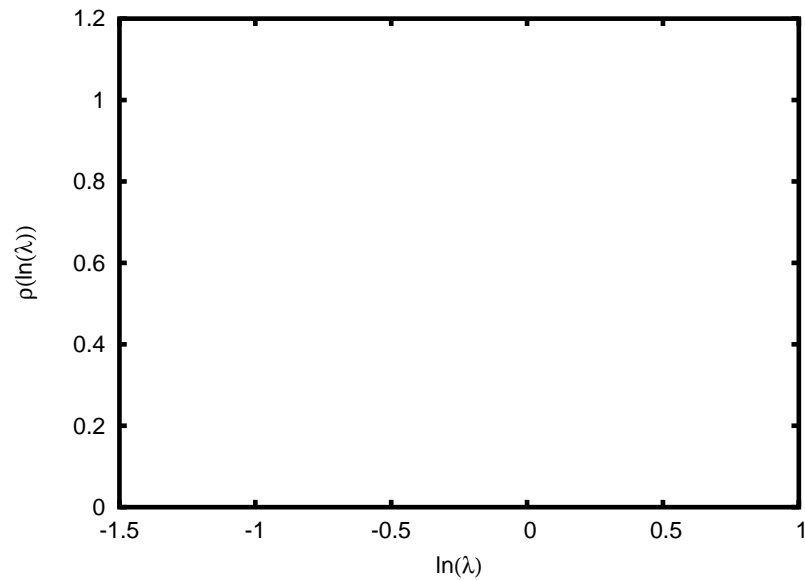
Numerical Tests

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

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

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

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



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

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

$$x_n = A x_{n-1} +$$

Summary

- Computed DOS of sample covariance matrices for multiple time-series using annealed calculation.
- Key ingredient: Szegő's theorem for (block) Toeplitz matrices
- Rectangular window and decorrelation approximation \Rightarrow Closed form approximation.
- Use of Szegős 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.**