Random matrix theory: from nuclear physics and number theory to complex networks and finance

Nestor Parolya

TU Delft
Applied Mathematics
Group of Statistics
n.parolya@tudelft.nl

October 6, 2022

Outline

High-Dimensional Asymptotics and RMT (Motivation)
- Large Dimensional Data Analysis
- The Curse of Dimensionality

Random Matrix Theory (RMT)
- The roots of RMT
- Marchenko-Pastur-Silverstein's Equation, CLT

RMT and complex random networks

Logdeterminant of sample correlation matrix

Asymptotics

Two main types of asymptotics in multivariate statistics:

- *standard asymptotics*
  - fixed dimension p and large sample size $n \to \infty$;
  - classical limit theorems hold
- *large dimensional asymptotics*
  - both the dimension p and the sample size n tend to infinity;
  - the ratio p/n tends to a positive constant $c > 0$;
  - classical limit theorems do not hold anymore (the curse of dimensionality).

Which is the right one?

> ”The large dimensional asymptotics is closer to reality.”
>
> - Huber (1973)

The Curse of Dimensionality

Estimation problem

Given: a sequence of i.i.d. $p-$dimensional random vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ with population covariance matrix $Cov(\mathbf{y}_i) = \mathbf{\Sigma}$ and $E(\mathbf{y}_i) = 0$

- estimate covariance matrix $\mathbf{\Sigma}$
- estimate eigenvalues

Common estimator: sample covariance $\mathbf{S}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{y}_i \mathbf{y}_i^\top = \frac{1}{n}\mathbf{Y}_n\mathbf{Y}_n^\top$

- Classical theory: p is fixed and $n \to \infty$

$$\mathbf{S}_n \xrightarrow{a.s.} \mathbf{\Sigma}.$$

In particular, the p random eigenvalues of S_n converge to the eigenvalues of $\mathbf{\Sigma}$.

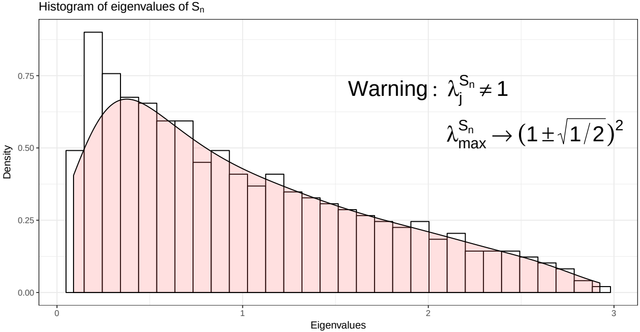
An effect of high-dimensions: example

Market application: 500 stocks ($= p$) and 1000 daily returns ($= n$)

Question: Can $\mathbf{S}_n$ be used estimate covariance in daily returns in this case ($p/n \to \frac{1}{2}$)? $\rightsquigarrow$ NO!

- $\mathbf{S}_n$ tends to under- or overestimate the true parameter
- eigenvalues of $\mathbf{S}_n$ do not consistently estimate eigenvalues of $\boldsymbol{\Sigma}$

Example: $\mathbf{y}_i \sim \mathcal{N}(0, I_p)$ i.i.d, $p/n \to \frac{1}{2}$ as $n \to \infty$



Curse of dimensionality: statistical inference

Logarithmic determinant: Let consider the important statistics in multivariate analysis

$$T_n = \log(\det(\mathbf{S}_n)) = \sum_{i=1}^{p} \log(\lambda_i)\,,$$

where λ_i are the eigenvalues of $\mathbf{S}_n$. When the dimension p is fixed it holds

$$T_n \xrightarrow{\text{a.s.}} 0 \text{ as } n \to \infty,$$

$$\sqrt{n/p}\,T_n \sim \mathcal{N}(0,2) \text{ as } n \to \infty\,.$$

If $p \to \infty$ and $n \to \infty$ such that $p/n \to c \in (0,1)$ then

$$\frac{1}{p} T_n \xrightarrow{\text{a.s.}} d(c) = \frac{c-1}{c}\log(1-c) - 1 < 0,$$

$$\sqrt{n/p}\,T_n \sim d(c)\sqrt{np} \to -\infty.$$

Warning! The statistical inference is not reliable anymore!

First Steps in Random Matrix Theory

John Wishart (1928)

- finite-dimensional matrices
- Wishart distribution

Eugen Wigner (1958)

- infinite matrices
- empirical spectral distribution

Wigner's motivation: bypass the Schrödinger equation and explain the statistics of experimentally measured atomic energy levels in terms of the limiting spectrum of the large random matrices.

Empirical spectral distribution (e.s.d)

For any symmetric $p \times p$ matrix A

$$F_n^A(x) = \frac{1}{p} \sum_{i=1}^{p} \mathbb{1}\{\lambda_i \leq x\}$$

is called empirical spectral distribution function of A.

- analyse the eigenvalue structure of A via the measure F_n^A
- find limiting spectral distribution F, i.e. $F_n^A \to F$ as $p \to \infty$

Wigner's semi-circle law

This law was first observed by Wigner (1955) for certain special classes of random matrices arising in quantum mechanical investigations.

Theorem (Arnold, L. (1967))

Let $\mathbf{W}_n = \frac{1}{\sqrt{n}}\mathbf{X}_n$*, where* $\mathbf{X}_n$ *is a symmetric random matrix with i.i.d. real random variables* $x_{i,j}$ *which have zero means and* $E(\xi_{i,j}^2) = 1$ *for* $i \neq j$ *and* $E(\xi_{i,i}^2) = 2$ *(Wigner matrix or GOE for Gaussian case). The empirical spectral distribution* F_n' *of the eigenvalues of* $\mathbf{W}_n$ *almost surely tends to*

$$F'(x) = \frac{1}{2\pi}\sqrt{4 - x^2} \times \mathbb{1}_{|x|\leq 2}\,.$$

Semi-circular law

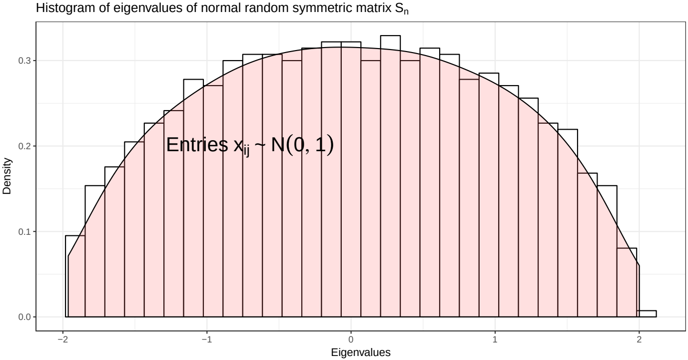


Figure: Limiting spectral distribution of GOE matrices

Stieltjes transform

For a function with bounded variation G the Stieltjes transform is defined as

$$m_G(z) = \int_{-\infty}^{+\infty} \frac{1}{\lambda - z} dG(\lambda) \tag{1}$$

$$\forall z \in \mathbb{C}^+ = \{z \in \mathbb{C} : Im(z) > 0\}.$$

The fundamental connection to random matrices: The Stieltjes transform of the e.s.d. $F_n(\lambda)$ of $\mathbf{S}_n$ is given by

$$\begin{aligned} m_{F_n}(z) &= \frac{1}{p}\sum_{i=1}^{p}\int_{-\infty}^{+\infty}\frac{1}{\lambda - z}\delta(\lambda-\lambda_i)d\lambda = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{\lambda_i - z} \\ &= \frac{1}{p}\mathrm{tr}\{(\mathbf{S}_n - z\mathbf{I})^{-1}\}. \end{aligned} \tag{2}$$

Volodymyr Marčenko

Leonid Pastur

In *Distribution of eigenvalues for some sets of random matrices*, Matematicheskii Sbornik 114(4) (1967), the limiting distribution of the eigenvalues of a sample covariance matrix as the size of the matrix grows was derived

Theorem

Silverstein's (Marčenko-Pastur) equation

Assume that (A1) holds, $p/n \to c \in (0, +\infty)$ and that $F^{\mathbf{\Sigma}_n}$ converges weakly to a cumulative distribution function (c.d.f.) H.
Then the e.d.f. $F^{1/n\mathbf{Y}_n\mathbf{Y}_n^\top}$ converges weakly almost surely to some deterministic c.d.f.s F, which Stieltjes transformation $m_F(z)$ is the unique solution of the following equation

$$m_F(z) = \int_{-\infty}^{+\infty} \frac{dH(\tau)}{\tau(1 - c - czm_F(z)) - z} . \quad (3)$$

Marchenko-Pastur Law, $\Sigma_n = \sigma^2 \mathbf{I}$

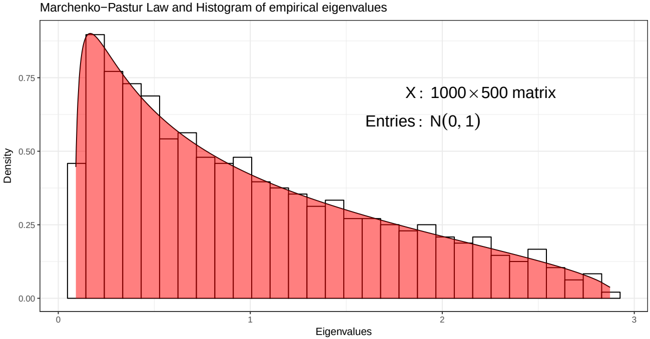
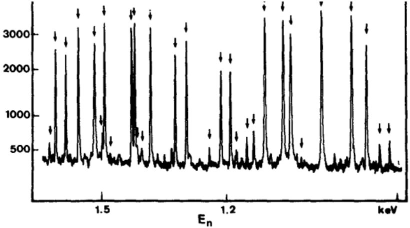


Figure: Marchenko-Pastur law for a Gaussian random matrix.

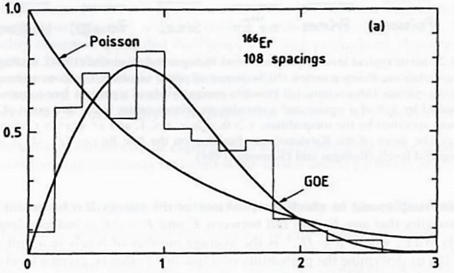
$$F'(\lambda) = \begin{cases} (1 - \frac{1}{c})\delta_0(\lambda) + f(\lambda) & \text{for} \quad c > 1 \\ f(\lambda) = \frac{1}{2\sigma\pi} \frac{\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})}}{c\lambda} & \text{for} \quad c \leq 1 \,, \end{cases} \qquad (4)$$

with $\lambda_{max} = \sigma^2(1 + \sqrt{c})^2$ and $\lambda_{min} = \sigma^2(1 - \sqrt{c})^2$.

Nuclear Physics = Economy



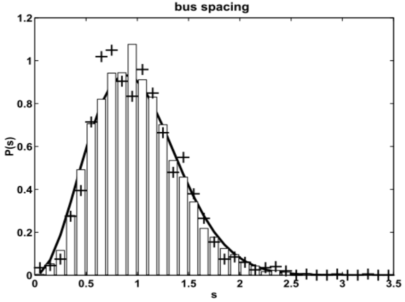
Scattering cross-section of neutron scattering for the Gadolinium-156 nucleus (it determines how likely the neutron is to bounce off the nucleus.) (Source: Coceva and Stefanon, Nuclear Physics A, 1979)



The probability density for the nearest neighbor spacings in slow neutron resonance levels.

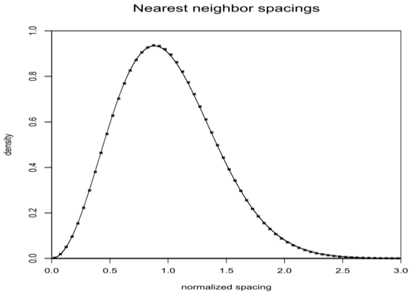
(Source: Liou et al, Phys.rev C5, 1972)

Nuclear Physics = Arriving of the bus = Number theory



A histogram of spacings between bus arrival times in Cuernavaca (Mexico)

(Source: Krbalek-Seba, J. Phys. A., 2000)



Spacing distribution for a billion zeroes of the Riemann zeta function and RMT prediction

(Source: Andrew Odlyzko, Contemp. Math. 2001)

Quote of the day

"Why should the same equation describe both the structure of an atomic nucleus and a sequence at the heart of number theory? And what do random matrices have to do with either of those realms? In recent years, the plot has thickened further, as random matrices have turned up in other unlikely places, such as games of solitaire, one-dimensional gases and chaotic quantum systems. Is it all just a cosmic coincidence, or is there something going on behind the scenes? ."

- Brian Hayes, "The spectrum of Riemannium",
American Scientist, 2003

Spectral properties of model networks

- It is well-known that the properties of networkor graphs can be characterized by spectrum of associated *adjacency* and Laplacian matrix (Chang, Spectral Graph Theory (1997)).
- For an unweighted graph, adjacency matrix **A** is defined in following way: $A_{ij} = 1$, if i and j nodes are connected and zero otherwise.
- For undirected networks, adjacency and Laplacian both are symmetric matrices and consequently have real eigenvalues.
- The rich information about the topological structure can be extracted from the spectral analysis of the networks.

Nearest Neighbour Spacing Distribution (NNSD)

Denote the eigenvalues of a network by $\lambda_i, i = 1, \ldots, N$ with N - size of the network.

Definition (NNSD)

Let $\bar{\lambda}_i = \bar{N}(\lambda_i) = NF(\lambda_i)$ be the unfolded eigenvalue. Then the *Nearest Neighbour Spacing* is defined as

$$s_i = \bar{\lambda}_{i+1} - \bar{\lambda}_i.$$

The NNSD is defined as a probability distribution $P(s)$ of s_i's.

Poisson density: $P(s) = e^{-s}$

GOE density: $P(s) = \frac{\pi}{2} s e^{-\frac{\pi s^2}{4}}$

Brody distribution: $P(s) = As^{\beta} e^{-\alpha s^{\beta+1}}$

with $A = (1+\beta)\alpha$ and $\alpha = \left[\Gamma\left(\frac{\beta+2}{\beta+1}\right)\right]^{\beta+1}$

NNSD reflects local correlations between eigenvalues.

Spectral rigidity Δ_3

Definition (Δ_3 statistic)

For any x from the spectrum denote

$$\Delta_3(L, x) = \frac{1}{L} \min_{a,b} \int_{x}^{x+L} [F_N(\lambda) - a\lambda - b]^2 \, d\lambda$$

where a and b are obtained from the OLS fit. Then the average over several values of x gives the spectral rigidity $\Delta_3(L)$.

It measures the least-square deviation of the spectral staircase function F_N from the best straight line fitting for a finite interval L of the spectrum.

"Picket fence" strongly correlated $\Delta_3(L) = 1/12$. The most rigid with all spacing equal (e.g., 1-D harmonic oscillator)

Uncorrelated $\Delta_3(L) = L/15$ (strong fluctuations around spectral density F')

GOE $\Delta_3(L) \sim \frac{1}{\pi^2} \log(L)$ (intermediate case)

Δ_3 reflects the long-range correlations among the eigenvalues.

Erdös-Rényi network

Starting with N nodes, random connections between pairs of nodes are made with probability p

average graph degree $k = p(N-1) \sim pN$

degree distribution $P(k) = C_k^{N-1} p^k (1-p)^{N-1-k}$

adjacent matrix A would have $2pN^2$ entries equal to one and rest entries zeros.

We pick $p = 0.01$ and $N = 2000$, which leads to a connected graph with average degree 20 and ca. 1% of nonzero values in matrix **A**. We generate 10 of such networks and plot the averages of the quantities of interest: spectral density, NNSD and Δ_3.

Erdös-Rényi network contd.

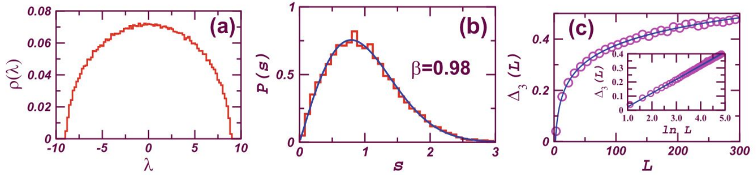


Fig. 6.1 (**a**) Density distribution, (**b**) NNSD and (**c**) Δ_3 statistics for random network with size $N = 2{,}000$ and average degree $\langle k \rangle = 20$. In *middle panel*, the histograms are numerical results and the *solid lines* represent fitted Brody distribution. For *last panel*, the *circles* are numerical results and the *solid curve* is the GOE prediction of RMT. *Inset* plots the $\Delta_3(L)$ in semi-logarithmic scale, in this scale it has the slope 0.0975. All figures are plotted for the average over ten realizations of the networks

Rai, A., Jalan, S. (2015). Application of Random Matrix Theory to Complex Networks. In: Banerjee, S., Rondoni, L. (eds) Applications of Chaos and Nonlinear Dynamics in Science and Engineering - Vol. 4. Understanding Complex Systems. Springer

Scale-free network

Scale-free networks are mostly modeled using algorithm provided by Barabási and Albert (2002, Rev. Mod. Phys.):

- Starting with a small number, m_0 of the nodes, a new node with $m \leq m_0$ connections is added at each time step.
- This new node connects with an already existing node i with probability $\pi(k_i) \propto k_i$ with k_i being degree of node i.
- After τ time steps the model leads to a network with $N = \tau + m_0$ nodes and $m\tau$ connections.

This ensures

degree distribution $P(k) = k^{-\gamma}$ (power law). $\gamma = 3$ for $\pi(k_i) \propto k_i$.

adjacent matrix A would have $2pN^2$ entries equal to one and rest entries zeros.

We keep $N = 2000$ with average degree of 20.

Scale-free network contd.

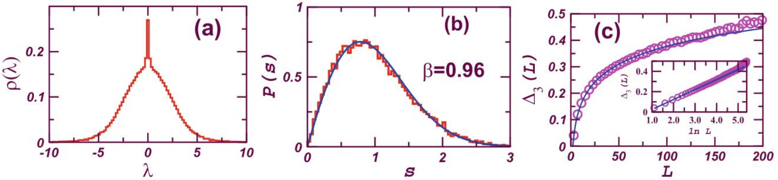


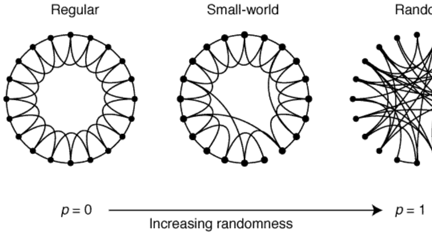
Fig. 6.2 (**a**) Density distribution, (**b**) NNSD and (**c**) Δ_3 statistics for scale-free network with $N = 2{,}000$ and $\langle k \rangle = 20$

Rai, A., Jalan, S. (2015). Application of Random Matrix Theory to Complex Networks. In: Banerjee, S., Rondoni, L. (eds) Applications of Chaos and Nonlinear Dynamics in Science and Engineering - Vol. 4. Understanding Complex Systems. Springer

Small-World networks

Small-world networks are constructed using the following algorithm of Watts and Strogatz (1998, Nature):

- Starting with an one-dimensional ring lattice of N nodes in which every node is connected to its $k = 2$ nearest neighbors, each connection of the lattice is rewired randomly with the probability p such that self-connections and multiple connections are excluded.
- Thus, $p = 0$ gives a regular network and $p = 1$ comletely random



  network.
- For $N = 2000$ and average node degree of 20, the typical small-world behavior is observed around $p = 0.005$

Small-World network contd.

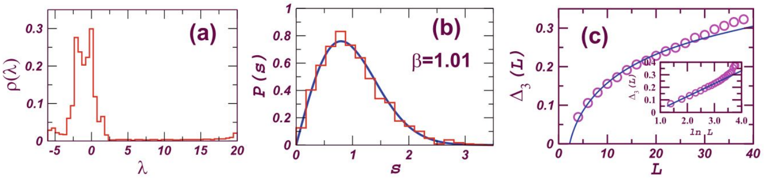


Fig. 6.3 For small-world network with $N = 2{,}000$ and $\langle k \rangle = 20$. (**a**) illustrates the density distribution depicting multi-peak structure. (**b**) depicts that nearest-neighbor spacing distribution (NNSD) $P(s)$ of the adjacency matrices of small-world network follows GOE statistics. The histograms are numerical results and the *solid lines* represent fitted Brody distribution. (**c**) plots $\Delta_3(L)$ statistics for eigenvalues spectra of the random network. The *circles* are numerical results and the *solid curve* is GOE prediction of RMT. *Inset* shows the $\Delta_3(L)$ in semi-logarithmic scale

Rai, A., Jalan, S. (2015). Application of Random Matrix Theory to Complex Networks. In: Banerjee, S., Rondoni, L. (eds) Applications of Chaos and Nonlinear Dynamics in Science and Engineering - Vol. 4. Understanding Complex Systems. Springer

Modular networks

Modular networks carry signature of community structure.

- Let us assume, for simplicity, that communities are modeled by random networks.
- Random matrices corresponding to unweighted random networks have entries 0 and 1, where number of 1's in a row follows a Gaussian distribution with mean p and variance $p(1-p)$.
- Take m random networks with connection probability p (each block GOE)
- Introduce random connections among sub-networks with probability q

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_q$$

- The ratio q/p can be considered as the relative strength of $\mathbf{A}_q$ and $\mathbf{A}_0$
- $q = 0$ corresponds to two completely separate blocks and $q = p$ to a random network.
- As the coupling between the two blocks increases ($q > 0$), the density distribution manifest a transition to the semicircular form at $q = p$:

$$\rho(\lambda) = \frac{2}{\pi\lambda_0^2}\sqrt{\lambda_0^2 - \lambda^2} \text{ with } \lambda_0 = (\lambda_{max} - \lambda_{min})/2 \,.$$

- We take $m = 2$, for $N = 500$ (each).

Modular networks contd.

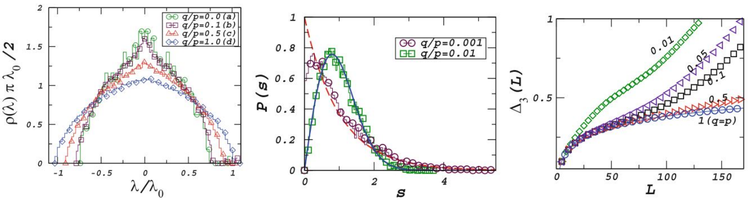


Fig. 6.4 Density distribution (*left panel*), NNSD (*middle panel*), Δ_3 statistics (*right panel*) of the two random sub-networks connecting each other with probability (a) $q = 0$ and $q/p = 0$; (b) $q = 0.001$ and hence $q/p = 0.1$; (c) $q = 0.005$, hence $q/p = 0.5$ and (d) $q/p = 1$ which corresponds to $q = 0.01$. Each block (random network) has size $N = 500$. The axes are scaled in such a way that the semicircle corresponding to $q = p$ has unit radius (see text). All graphs are plotted for 20 realizations of random sets of connections among the two sub-networks

Rai, A., Jalan, S. (2015). Application of Random Matrix Theory to Complex Networks. In: Banerjee, S., Rondoni, L. (eds) Applications of Chaos and Nonlinear Dynamics in Science and Engineering - Vol. 4. Understanding Complex Systems. Springer

Sample correlation matrix

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a sequence of independent p-dimensional random vectors from some common distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The corresponding correlation matrix $\mathbf{R}$ is defined by

$$\mathbf{R} = \mathrm{diag}(\boldsymbol{\Sigma})^{-1/2} \times \boldsymbol{\Sigma} \times \mathrm{diag}(\boldsymbol{\Sigma})^{-1/2},$$

Sample correlation matrix

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a sequence of independent p-dimensional random vectors from some common distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The corresponding correlation matrix $\mathbf{R}$ is defined by

$$\mathbf{R} = \text{diag}(\boldsymbol{\Sigma})^{-1/2} \times \boldsymbol{\Sigma} \times \text{diag}(\boldsymbol{\Sigma})^{-1/2},$$

Similarly, let $\mathbf{S}$ be the sample covariance matrix with the corresponding sample correlation matrix defined by

$$\hat{\mathbf{R}} = \text{diag}(\mathbf{S})^{-1/2} \times \mathbf{S} \times \text{diag}(\mathbf{S})^{-1/2}.$$

Sample correlation matrix

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a sequence of independent p-dimensional random vectors from some common distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The corresponding correlation matrix $\mathbf{R}$ is defined by

$$\mathbf{R} = \text{diag}(\boldsymbol{\Sigma})^{-1/2} \times \boldsymbol{\Sigma} \times \text{diag}(\boldsymbol{\Sigma})^{-1/2},$$

Similarly, let $\mathbf{S}$ be the sample covariance matrix with the corresponding sample correlation matrix defined by

$$\hat{\mathbf{R}} = \text{diag}(\mathbf{S})^{-1/2} \times \mathbf{S} \times \text{diag}(\mathbf{S})^{-1/2}.$$

In this talk we will present the central limit theorem (CLT) for $\log |\hat{\mathbf{R}}|$ in case

- $p \to \infty$, $n \to \infty$ and $p/n \to \gamma \in (0, 1]$ and $p \leq n$,
- the spectral norm of $\mathbf{R}$ is uniformly bounded,
- finite fourth moments, i.e., $\mathbb{E}(x_{ij}^4) < \infty$.

Why logarithmic determinant of sample correlation matrix?

The sample correlation matrix $\hat{\mathbf{R}}$ is a popular target in multivariate analysis. See, for example, the classical books by Anderson (1958), Muirhead (1982) and Eaton (1983).

Why logarithmic determinant of sample correlation matrix?

The sample correlation matrix $\hat{\mathbf{R}}$ is a popular target in multivariate analysis. See, for example, the classical books by Anderson (1958), Muirhead (1982) and Eaton (1983).

- $|\hat{\mathbf{R}}|$ is the likelihood ratio test statistic for testing that the p entries of $\mathbf{x}_1$ are independent/uncorrelated.

Why logarithmic determinant of sample correlation matrix?

The sample correlation matrix $\hat{\mathbf{R}}$ is a popular target in multivariate analysis. See, for example, the classical books by Anderson (1958), Muirhead (1982) and Eaton (1983).

- $|\hat{\mathbf{R}}|$ is the likelihood ratio test statistic for testing that the p entries of $\mathbf{x}_1$ are independent/uncorrelated.
- If $\mathbf{R} = \mathbf{I}$, the density of $\hat{\mathbf{R}}$ is given by

$$\text{Constant} \times |\mathbf{R}|^{(n-p-2)/2} d\mathbf{R}\,.$$

Then in case $p = n - 2$ or $p/n \to 1$ one can test on the (approximate) uniformity of the entries of a large random correlation matrix.

Why logarithmic determinant of sample correlation matrix?

The sample correlation matrix $\hat{\mathbf{R}}$ is a popular target in multivariate analysis. See, for example, the classical books by Anderson (1958), Muirhead (1982) and Eaton (1983).

- $|\hat{\mathbf{R}}|$ is the likelihood ratio test statistic for testing that the p entries of $\mathbf{x}_1$ are independent/uncorrelated.
- If $\mathbf{R} = \mathbf{I}$, the density of $\hat{\mathbf{R}}$ is given by

$$\text{Constant} \times |\mathbf{R}|^{(n-p-2)/2} d\mathbf{R}\,.$$

  Then in case $p = n - 2$ or $p/n \to 1$ one can test on the (approximate) uniformity of the entries of a large random correlation matrix.
- The volume of a hyperellipsoid constructed from standardized vectors is proportional the determinant of the sample correlation matrix.

What is already known?

In case $\mathbf{R} = \mathbf{I}$ and $\mathbf{x}_1$ is multivariate normal it is much known about $\log |\hat{\mathbf{R}}|$, i.e.,

- Empirical distribution of the eigenvalues of $\hat{\mathbf{R}}$ satisfies Marchenko-Pastur law (Jiang (2004))

What is already known?

In case $\mathbf{R} = \mathbf{I}$ and $\mathbf{x}_1$ is multivariate normal it is much known about $\log |\hat{\mathbf{R}}|$, i.e.,

- Empirical distribution of the eigenvalues of $\hat{\mathbf{R}}$ satisfies Marchenko-Pastur law (Jiang (2004))
- The largest eigenvalue of $\hat{\mathbf{R}}$ asymptotically satisfies the Tracy-Widom law (Bao, Pan and Zhou (2012))

What is already known?

In case $\mathbf{R} = \mathbf{I}$ and $\mathbf{x}_1$ is multivariate normal it is much known about $\log|\hat{\mathbf{R}}|$, i.e.,

- Empirical distribution of the eigenvalues of $\hat{\mathbf{R}}$ satisfies Marchenko-Pastur law (Jiang (2004))
- The largest eigenvalue of $\hat{\mathbf{R}}$ asymptotically satisfies the Tracy-Widom law (Bao, Pan and Zhou (2012))
- The quantity $\log|\hat{\mathbf{R}}|$ satisfies the CLT (Jiang and Yang (2013) and Jiang and Qi (2015)).

In case $\mathbf{R} \neq \mathbf{I}$ and $\mathbf{x}_1$ is multivariate normal there is the CLT proved by Jiang (2019, AoAP).

What is already known?

In case $\mathbf{R} = \mathbf{I}$ and $\mathbf{x}_1$ is multivariate normal it is much known about $\log |\hat{\mathbf{R}}|$, i.e.,

- Empirical distribution of the eigenvalues of $\hat{\mathbf{R}}$ satisfies Marchenko-Pastur law (Jiang (2004))
- The largest eigenvalue of $\hat{\mathbf{R}}$ asymptotically satisfies the Tracy-Widom law (Bao, Pan and Zhou (2012))
- The quantity $\log |\hat{\mathbf{R}}|$ satisfies the CLT (Jiang and Yang (2013) and Jiang and Qi (2015)).

In case $\mathbf{R} \neq \mathbf{I}$ and $\mathbf{x}_1$ is multivariate normal there is the CLT proved by Jiang (2019, AoAP).

Nothing is known about $\log |\hat{\mathbf{R}}|$ in case $\mathbf{R} \neq \mathbf{I}$ and $\mathbf{x}_1$ with generic non-normal distribution.

What is already known?

In case $\mathbf{R} = \mathbf{I}$ and $\mathbf{x}_1$ is multivariate normal it is much known about $\log|\hat{\mathbf{R}}|$, i.e.,

- Empirical distribution of the eigenvalues of $\hat{\mathbf{R}}$ satisfies Marchenko-Pastur law (Jiang (2004))
- The largest eigenvalue of $\hat{\mathbf{R}}$ asymptotically satisfies the Tracy-Widom law (Bao, Pan and Zhou (2012))
- The quantity $\log|\hat{\mathbf{R}}|$ satisfies the CLT (Jiang and Yang (2013) and Jiang and Qi (2015)).

In case $\mathbf{R} \neq \mathbf{I}$ and $\mathbf{x}_1$ is multivariate normal there is the CLT proved by Jiang (2019, AoAP).

Nothing is known about $\log|\hat{\mathbf{R}}|$ in case $\mathbf{R} \neq \mathbf{I}$ and $\mathbf{x}_1$ with generic non-normal distribution.

Till now...

Non-centered case

First, we start with the (non-centered) sample correlation matrix **R** given by

$$\hat{\mathbf{R}} = \mathrm{diag}(\mathbf{S})^{-1/2}\,\mathbf{S}\,\mathrm{diag}(\mathbf{S})^{-1/2}\,,$$

where $\mathbf{S} = (1/n)\mathbf{X}\mathbf{X}^\top = (1/n)\boldsymbol{\Sigma}^{1/2}\mathbf{Z}\mathbf{Z}^\top\boldsymbol{\Sigma}^{1/2}$ with the noise matrix $\mathbf{Z} = (\mathbf{z}_1, \ldots, \mathbf{z}_n)$.

Non-centered case

First, we start with the (non-centered) sample correlation matrix $\mathbf{R}$ given by

$$\hat{\mathbf{R}} = \mathrm{diag}(\mathbf{S})^{-1/2}\,\mathbf{S}\,\mathrm{diag}(\mathbf{S})^{-1/2},$$

where $\mathbf{S} = (1/n)\mathbf{X}\mathbf{X}^\top = (1/n)\mathbf{\Sigma}^{1/2}\mathbf{Z}\mathbf{Z}^\top\mathbf{\Sigma}^{1/2}$ with the noise matrix $\mathbf{Z} = (\mathbf{z}_1, \ldots, \mathbf{z}_n)$.

Theorem (Logarithmic law in the non-centered case)

Assume that z_{ij} are independent random variables with mean zero, variance one and finite fourth moment $\mathbb{E}|z_{11}|^4 < \infty$. If the spectral norm of $\mathbf{R}$ is uniformly bounded and $p/n \to \gamma \in (0,1)$, and let

$$\mu_n = \log\det(\mathbf{R}) + (p - n + \frac{1}{2})\log(1 - \frac{p}{n}) - p + \frac{p}{n} + \frac{1}{2}\frac{p}{n}(\mathbb{E}|z_{11}|^4 - 3)\left(C_{\mathbf{R}^{1/2}} - 1\right)$$

$$\sigma_n^2 = -2\log(1 - \frac{p}{n}) - 2\frac{p}{n} + 2\frac{p}{n}\,tr(\mathbf{R} - \mathbf{I})^2/p$$

then

$$\frac{\log\det(\hat{\mathbf{R}}) - \mu_n}{\sigma_n} \xrightarrow{d} \mathcal{N}(0,1),$$

where $C_{\mathbf{R}^{1/2}} = \frac{1}{p}||\mathbf{R}^{1/2} \circ \mathbf{R}^{1/2}||_F^2 = \frac{1}{p}tr\left[\left(\mathbf{R}^{1/2} \circ \mathbf{R}^{1/2}\right)^2\right]$ and '$\circ$' denotes the Hadamard product.

Centered case

One has rather to consider rather a centered sample corelation matrix

$$\hat{\mathbf{R}}_c = \mathrm{diag}(\mathbf{S}_c)^{-1/2}\,\mathbf{S}_c\,\mathrm{diag}(\mathbf{S}_c)^{-1/2},$$

where $\mathbf{S}_c$ is the centered (by the sample mean) sample covariance matrix given by

$$\mathbf{S}_c = \frac{1}{n-1}(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1}^\top)(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1}^\top)^\top \quad \text{with } \bar{\mathbf{x}} = 1/n\mathbf{X}\mathbf{1} \text{ the sample mean}$$

and $\mathbf{1} = (1, \ldots, 1)$ denotes the n-dimensional vector of ones. We have the following corollary.

Centered case

One has rather to consider rather a centered sample corelation matrix

$$\hat{\mathbf{R}}_c = \mathrm{diag}(\mathbf{S}_c)^{-1/2}\,\mathbf{S}_c\,\mathrm{diag}(\mathbf{S}_c)^{-1/2},$$

where $\mathbf{S}_c$ is the centered (by the sample mean) sample covariance matrix given by

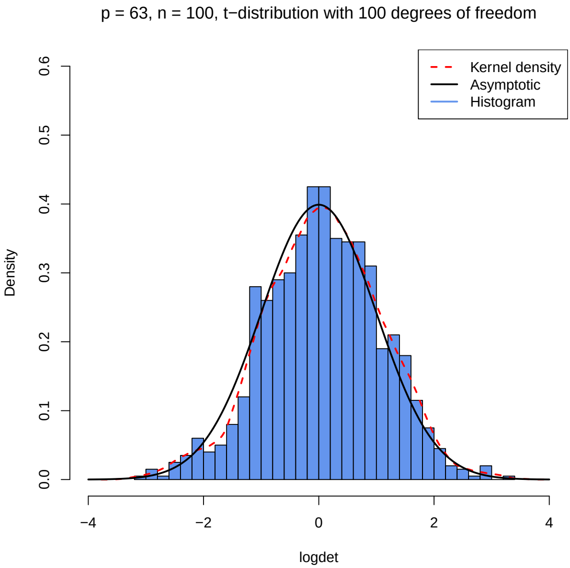
$$\mathbf{S}_c = \frac{1}{n-1}(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1}^\top)(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1}^\top)^\top \quad \text{with } \bar{\mathbf{x}} = 1/n\mathbf{X}\mathbf{1} \text{ the sample mean}$$

and $\mathbf{1} = (1, \ldots, 1)$ denotes the n-dimensional vector of ones. We have the following corollary.

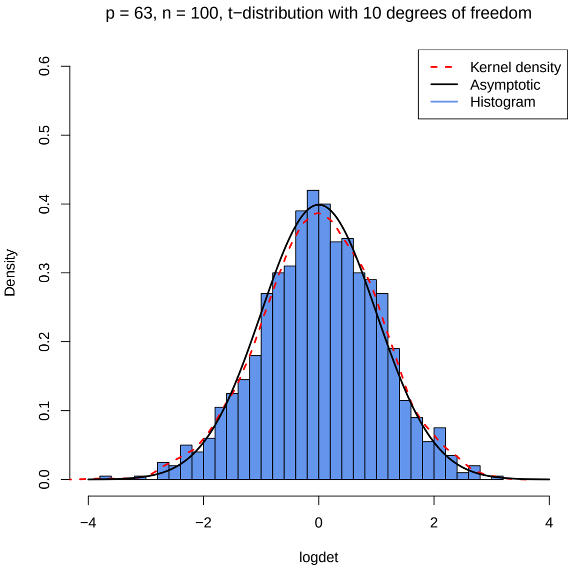
Corollary (Logarithmic law in the centered case (substitution principle))

Under conditions of main Theorem and for centered sample correlation matrix $\mathbf{R}_c$ the obtained CLT still holds if one replaces everywhere in formulas of μ_n and σ_n^2 the sample size n by $n-1$.

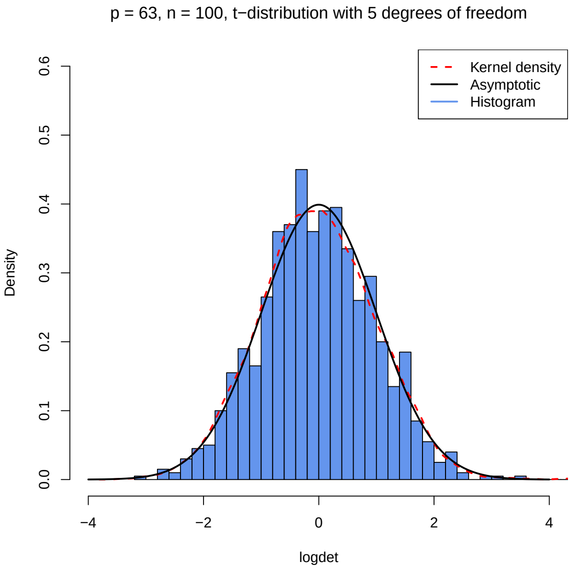
Finite sample example I



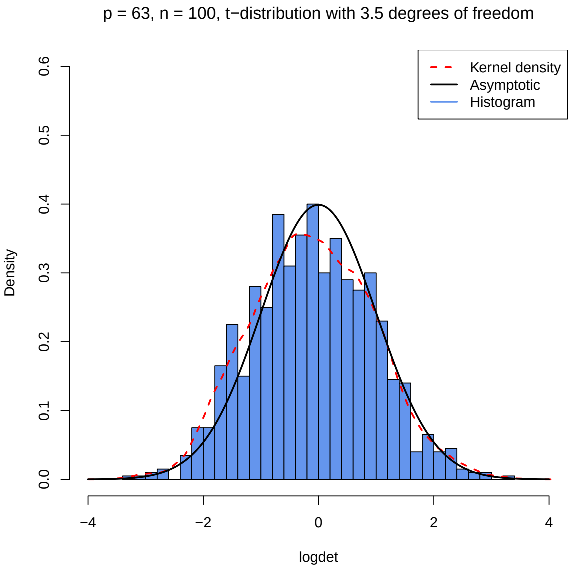
Finite sample example II



Finite sample example III



Finite sample example IV



Thank you very much for your attention!

- Rai, A., Jalan, S. (2015). Application of Random Matrix Theory to Complex Networks. In: Banerjee, S., Rondoni, L. (eds) Applications of Chaos and Nonlinear Dynamics in Science and Engineering - Vol. 4. Understanding Complex Systems. Springer
- Bai Z.D., and J. W. Silverstein (2004), CLT for linear spectral statistics of large-dimensional sample covariance matrices, *Annals of Probability* 32, 553-605.
- Bai Z.D., and J. W. Silverstein (2010) Spectral Analysis of Large Dimensional Random Matrices, Springer: New York; Dordrecht; Heidelberg; London.
- Bodnar, T., Dette, H., and Parolya N. (2019). Testing for independence of large dimensional vectors. *The Annals of Statistics* 47(5), 2977-3008.
- Heiny, J., Parolya, N. (2022). Log determinant of large correlation matrices under infinite fourth moment, under revision in *Annales de l'Institut Henri Poincaré*
- Parolya, N., Heiny. J., Kurowicka, D. (2022). Logarithmic law of large random correlation matrix (submitted for publication)
- Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Sbornik: Mathematics* 1, 457-483.