Random Matrix Theory

Evgenii Chzhen

16 Jan 2016

Contents

This paper is based on lectures given by Mr. Jamal Najim at Universit?? Paris-Est Marne-la-Vall??e. ## 1. Introduction In this paper we want to proceed experimental tests of two theorems in random matrix theory (RMT). We start with an introduction of basic objects in RMT.

Definition 1. *Consider a symmetric (Hermitian) matrix Z_N of $N \times N$ dimension, each element of the matrix $(Z_N)_{ij}$ is a random real (complex) valued variable. An empirical spectral measure of Z_N is defined as*

$$\mathcal{L}_N(A) = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}(A) = \frac{\#\{\textit{eigenvalues of } Z_N \textit{ in } A\}}{N}, \textit{ for } A \subseteq \mathbb{R}.$$

Remark 1. *Since Z_N is a symmetric (Hermitian) matrix, therefore by Spectral theorem there exist $\lambda_1, ..., \lambda_N \in \mathbb{R}$, where $\lambda_1, ..., \lambda_N$ are eigenvalues of matrix Z_N counted with multiplicity.*

Remark 2. $\forall f \in \mathbb{C}_b(\mathbb{R})$ *holds* $\int f(x)\,\mathcal{L}_N(dx) = \frac{1}{N}\sum\limits_{i=1}^{N} f(\lambda_i)$.

2. Wigner theorem

In this section we define a Wigner matrix and state a Wigner theorem.

Definition 2. *Z_N is called a Wigner Matrix if Z_N is a Hermitian matrix such that for all $i \leq j$ $(Z_N)_{ij} = \frac{Y_{ij}}{\sqrt{N}}$ where Y_{ij} are i.i.d random variables with $\mathbb{E}Y_{ij} = 0$ and $VarY_{ij} = \sigma^2$.*

Definition 3. *Random variable X has a semicircle distribution if its' probability density function is defined as it follows*

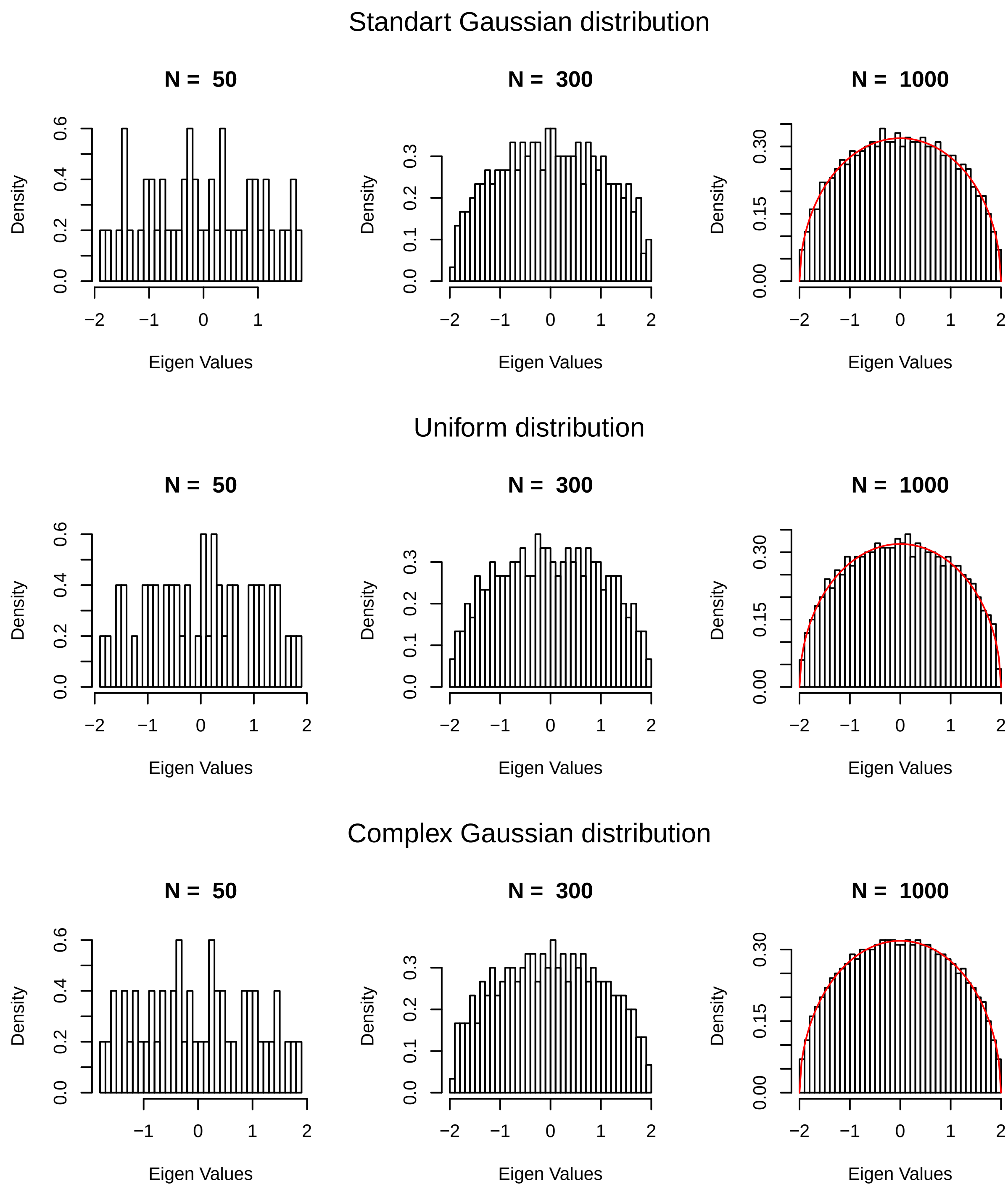
$$\mathbb{P}_{sc}(dx) = \frac{\sqrt{4\sigma^2 - x^2}}{2\pi\sigma^2}\mathbb{1}_{[-2\sigma, 2\sigma]}(x)dx.$$

Theorem 1. *Let Z_N be a Wigner Matrix and $\mathcal{L}_N$ is an empirical spectral measure of Z_N, therefore almost surely*

$$\mathcal{L}_N \xrightarrow{\mathcal{D}} \mathbb{P}_{sc}$$

.

Now we want to make a numerical experiments, let Z_N be a Wigner Matrix, for $N = 50, 300, 1000$ we make a simulation of the matrix from a given distribution and compute eigenvalues after we plot a histogram of obtained values and compare them with theoretical results. Red line corresponds to semicircle distribution.



3. Marchenko-Pastur theorem

In this section we state a Marchenko-Pastur theorem.

Definition 4. *For $c > 0$ a distribution of Marchenko-Pastur is defined as*

$$\mathbb{P}_{MP}(dx) = \begin{cases} \frac{\sqrt{(\lambda^+ - x)(x-\lambda^-)}}{2\pi x \sigma^2 c} \mathbb{1}_{[\lambda^-,\lambda^+]}(x)dx, & \text{if } 0 < c \leq 1 \\ (1-\frac{1}{c})\delta_0(dx) + \frac{\sqrt{(\lambda^+ - x)(x-\lambda^-)}}{2\pi x \sigma^2 c} \mathbb{1}_{[\lambda^-,\lambda^+]}(x)dx, & \text{if } c > 1 \end{cases},$$

where $\lambda^+ = \sigma^2(1+\sqrt{c})^2$ and $\lambda^- = \sigma^2(1-\sqrt{c})^2$.

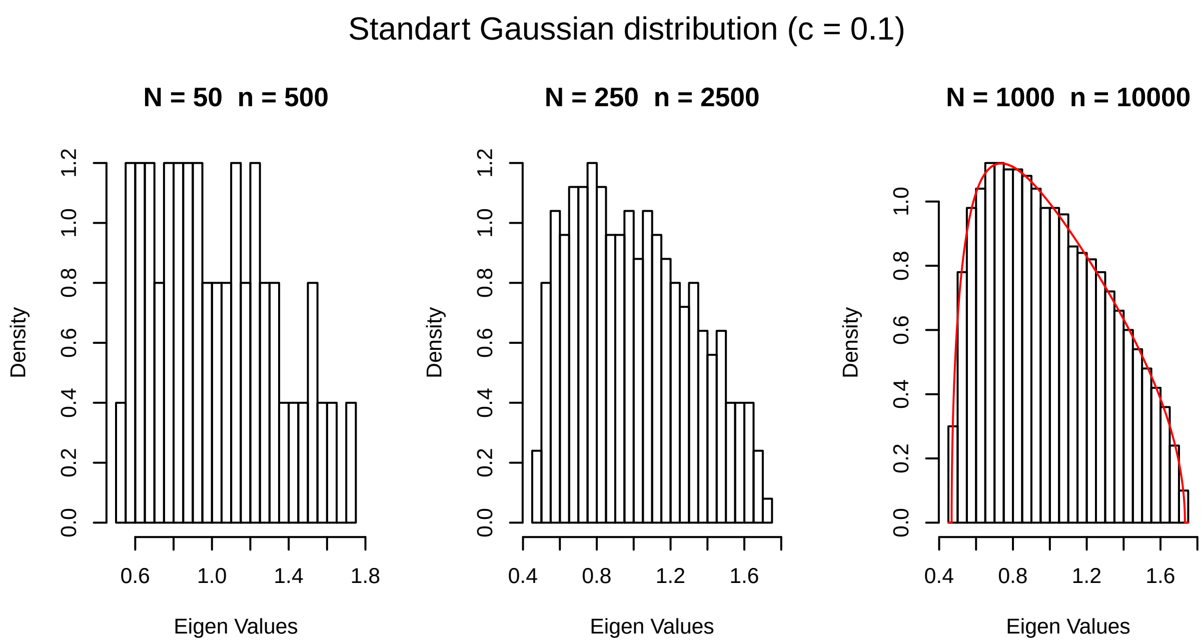
Theorem 2. *Let $Z_N = \frac{1}{n}X_N X_N^*$, where X_N is a matrix of dimension $N \times n$ and $(X_{ij})_{ij}$ are i.i.d. random variables such that $\mathbb{E}X_{ij} = 0$ and $Var X_{ij} = \sigma^2$. Let $\mathcal{L}_N$ be an empirical spectral measure of Z_N. If $\frac{N}{n} \to c > 0$ therefore almost surely*

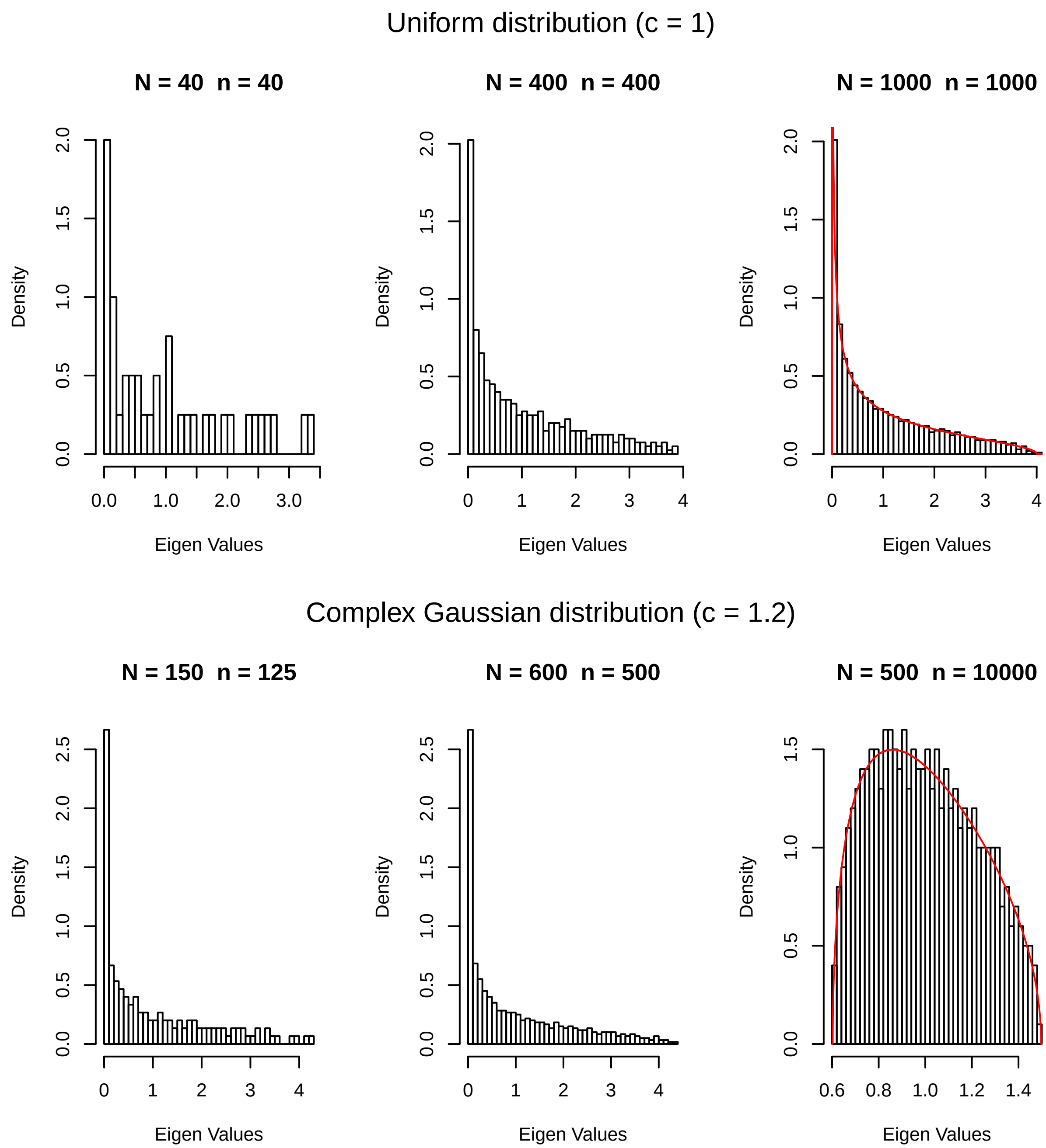
$$\mathcal{L}_N \xrightarrow{\mathcal{D}} \mathbb{P}_{MC},$$

where $\mathbb{P}_{MC}$ is a Marchenko-Pastur distribution with parameters c, σ^2.

Remark 3. *Assume that $c > 1$ then for large enough N, n we have $N > n$ since $rank(Z_N) = min\{N, n\}$ we will have $N - n = n(c-1)$ zero eigenvalues. Since $\mathcal{L}_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}$ we can observe that in the limiting measure there will be a mass of $\frac{n(c-1)}{N} = (1-\frac{1}{c})$ at 0 point.*

Now we want to make a numerical experiments, let Z_N be a matrix which satisfies assumptions of Marchenko-Pastur theorem, we make a simulation of the matrix from a given distribution and compute eigenvalues after we plot a histogram of obtained values and compare them with theoretical results. Red line corresponds to Marchenko-Pastur distribution.





4. Corollary

In this paper we introduced some basic definitions and theorems of random matrix theory, we proceeded a numerical procedure to show the ideas of theorems on simulated data. Our experimental results are in high agreement with the theoretical. For numerical experiments we used R language, particulary **RMTstat** package were used for a Marchenko-Pastur distribution and **cmvnorm** package for simulating complex valued Gaussian variables.