Research Statement

Indrajit Jana

As a graduate student, my research was mainly focused on Random Matrix Theory and its applications. In particular, I looked at several problems in random band matrices. Roughly speaking, a random matrix is a matrix whose entries are random variables. In Random Matrix Theory, we study the asymptotic behaviour of eigenvalues, eigenvectors of such matrices.

Background

Random Matrix Theory was developed from several different sources in early 20th century. It is used as an important mathematical tool in various fields namely, Mathematics, Physics, Communication Engineering etc. One of the earliest example of random matrix appeared in the study of sample covariance estimation done by John Wishart [44]. In the early 1950s, Eugene Wigner introduced random matrices to study the energy spectra of heavy atoms undergoing slow nuclear reactions. In 1970s, a connection between Random Matrix Theory and Number Theory was found. The Riemann hypothesis says that the nontrivial zeros of the Riemann zeta function lie on the line $\frac{1}{2} + iE$ with $-\infty < E < \infty$. Assuming the Riemann hypothesis, Montgomery calculated the asymptotic two point correlation of these zeros, which turns out to be same as the two-level correlation function of the unitary random matrix ensemble [22]. Random matrices are also used to model wireless channels. A random matrix model of CDMA networks can be found in [39, 40]. Random matrices can be used in multivariate analysis [20, 23]. A nice description of use of random matrices in statistics can be found here [25]. These are by no means a complete list of applications of random matrices. A more detailed and nicely structured list of applications of random matrices can be found here [38].

A special kind of random matrix ensemble is a random band matrix. In 1955, Wigner studied the matrices $H$ of the form $H = K + V$, where $K$ is a diagonal matrix consisting of integers $\cdots -2, -1, 0, 1, 2, \cdots$, and $V$ is a symmetric sign matrix having non vanishing elements only up to a distance $b_n$ from the main diagonal. He called the matrices $H$ as bordered matrix [41, 42].

A relatively modern treatment of random band matrix was done by G. Casati et al. in the context of Quantum Chaos [7, 6]. They studied $n \times n$ symmetric random band matrices of bandwidth $b_n$, where $b_n$ grows with $n$.

In general, in the study of random matrices, the underlying distributions of the entries of the random matrix are unknown. Many random matrix results are universal, and they do not depend on the underlying distributions.

Early developments

One of the earliest result in this subject is the Wigner Semicircle law [43]. In 1957, Wigner proved that if $W_n$ is an $n \times n$ real symmetric random matrix, then the scaled histogram of the eigenvalues of $\frac{1}{\sqrt{n}} W_n$ converges to a probability density function given by $\frac{1}{2\pi} \sqrt{4 - x^2}$, which is known as the Semicircle law. Later, the technical assumptions were more relaxed, and the theorem was proved in more generality [2].

I am interested in random band matrices. In 1991, the Semicircle law for periodic random band matrices was proved by Bogachev, Casati et al. [5, 4]. Around the same time, Molchanov et al. proved it for a general class of random band matrices, under the assumption that $\frac{b_n}{n} \to 0$ as $n \to \infty$, where $b_n$ is the bandwidth of the band random matrix [4, 21]. Recently a study of the Semicircle law in local scale has been done by Erdős et al. [11]. Around 1993, limiting distribution of the largest eigenvalue of random Hermitian matrices was found [37]. A similar result was found for Hermitian random band matrices with sufficiently large bandwidth (iff $b_n >> n^{5/6}$) [31].
More recent developments

The convergence of empirical distribution of eigenvalues of a random matrix is an analogue of the Strong law large numbers in Probability. After the Strong law of large numbers, we would like to study the fluctuations of the ESD around the Semicircle law. The study of Central limit theorems (CLT) in Probability theory is quite old. Various CLTs were proved for independent, weakly dependent sequence of random variables. But in the case of Random Matrix Theory, the eigenvalues of a random matrix are not independent. In fact, they are highly correlated. Which makes study of CLT for linear eigenvalue statistics more interesting.

We consider the linear eigenvalue statistics \( \sum_{i=1}^{n} \phi(\lambda_i) \), and then study its fluctuations around its mean. This is a relatively new topic. In 1982, Jonsson proved the CLT for the linear eigenvalue statistics of sample covariance matrices with polynomial test functions \[16\]. Khorunzhy et. al. proved it for the resolvent of a

This is a relatively new topic. In 1982, Jonsson proved the CLT for the linear eigenvalue statistics of sample covariance matrices with polynomial test functions \[16\]. Khorunzhy et. al. proved it for the resolvent of a Wigner matrix \[17\]. In 1998, Sinai and Soshnikov proved the CLT for traces of random symmetric matrices \[30\]. They proved that \( \text{Tr}(W_n^p) - E[\text{Tr}(W_n^p)] \xrightarrow{d} N(0, \sigma) \), where \( W_n \) is \( n \times n \) symmetric random matrix and \( p \) increases as \( n \) increases. Later, CLT was proved for the linear eigenvalue statistics \( \sum_{i=1}^{n} \phi(\lambda_i) \) for random Hermitian \[19\], random sample covariance \[27\], and sparse matrices \[29\], where \( \lambda_i \) are the eigenvalues of the random matrix and \( \phi : \mathbb{R} \to \mathbb{R} \) is a sufficiently smooth function.

Now we focus on the similar results for random band matrices. Proving the CLT for linear eigenvalue statistics of a random band matrix is more difficult, as it lacks the rotational invariance. Recently Soshnikov, and Li \[18\] proved CLT for random band matrices. The theorem is stated below.

**Theorem 0.1.** Let \( M_n = \frac{1}{n} W_n \) be a real symmetric random band matrix of bandwidth \( b_n \), where \( \{b_n\} \) satisfies \( \sqrt{n} << b_n << n \). Assume that

(i) Diagonal and non-zero off diagonal entries of \( W_n \) are two sets of i.i.d. random variables.

(ii) The marginal probability distribution of \( W_{ij} \) satisfies the Poincaré inequality with some uniform constant.

(iii) \( E[W_{ij}] = 0, E[W_{ij}^2] = (1 + \delta_{ij})\sigma^2 \), and the fourth moment of the nonzero off diagonal entries does not depend on \( n \).

Then we have

\[
S_n^b := \sqrt{\frac{b_n}{n}} \left[ \sum_{i=1}^{n} \phi(\lambda_i) - E \left( \sum_{i=1}^{n} \phi(\lambda_i) \right) \right] \xrightarrow{d} N(0, \sigma^2_\phi),
\]

where \( \phi : \mathbb{R} \to \mathbb{R} \) is a function with continuous bounded derivatives, and \( \sigma^2_\phi \) is given by

\[
\sigma^2_\phi = \int_{-2\sqrt{2}\sigma}^{2\sqrt{2}\sigma} \int_{-2\sqrt{2}\sigma}^{2\sqrt{2}\sigma} \frac{\phi(x) - \phi(\lambda)\phi'(y)\sqrt{8\sigma^2-x^2}\sqrt{8\sigma^2-y^2}}{4\pi^4(x-\lambda)^2} F_\sigma(x, y) \, dx \, dy \, d\lambda + \frac{\kappa_4}{16\pi^4\sigma^8} \left( \int_{-2\sqrt{2}\sigma}^{2\sqrt{2}\sigma} \frac{\phi(\lambda)(4\sigma^2-\lambda^2)}{\sqrt{8\sigma^2-\lambda^2}} \, d\lambda \right)^2,
\]

and for \( x \neq y \)

\[
F_\sigma(x, y) := \int_{-\infty}^{\infty} \frac{(s^3 \sin s - s \sin^3 s)}{2\sigma^2(s^2 - \sin^2 s)^2 - (s^3 \sin s + s \sin^3 s)xy + s^2 \sin^2 s(x^2 + y^2)},
\]

\( \kappa_4 \) is the fourth cumulant of off-diagonal entries, i.e., \( \kappa_4 = E[W_{12}^4] - 3\sigma^4 \).

Later on, we have found a shorter proof of the same result, using martingale difference technique \[14\]. Further developments had been done by M. Shcherbina \[28\].
Some tools which I used

Most common tools used in the Random Matrix Theory are the moment method, and the Stieltjes transform method. Both tools are widely used to prove the Semicircle law in various cases.

Generally, we use the method of characteristic functions to prove the CLT or convergence in distribution in general. The method works as follows.

Let $S_n$ be a sequence of random variables, and $Z_n(t) = E[\exp\{itS_n\}]$, $Z(t) = \exp[-t^2\sigma^2/2]$. By Lévy’s continuity theorem we know that if the sequence of characteristic functions $Z_n(t)$ converges to $Z(t)$ pointwise as $n \to \infty$ then $S_n \xrightarrow{d} N(0, \sigma^2)$ as $n \to \infty$. Now, in particular, if we take

$$S_n^* = \sqrt{\frac{b_n}{n}} \left[ \sum_{i=1}^{n} \phi(\lambda_i) - E\left( \sum_{i=1}^{n} \phi(\lambda_i) \right) \right]$$

and if we can prove the above argument then we will have the result described in [1]. However, the main difficulty is to compute the characteristic function $Z_n(t)$ for the linear eigenvalue statistics, and then to show that it actually converges to $Z(t)$. To avoid that computation, we use the Stein’s Method, which was originally introduced by Stein [32, 33]. I will describe a modified version of the Stein’s method.

First of all, we notice that $Z(t) = \exp[-t^2\sigma^2/2]$ is the unique bounded continuous solution of the following integral equation.

$$X(t) = 1 - \sigma^2 \int_0^t sX(s) \, ds. \quad (2)$$

We also observe that at $t = 0$ the value of the characteristic functions $Z_n(0) = 0$. Therefore to prove the CLT, it is enough to show that the characteristic functions $Z_n(t)$ satisfies the same equation (2) as $n \to \infty$. In other words, it is enough to show that any converging subsequence $\{Z_{n_j}(t)\}$ and $\{Z_{n_j}'(t)\}$ satisfy

$$\lim_{n_j \to \infty} Z_{n_j}'(t) = \lim_{n_j \to \infty} tZ_{n_j}'(t)$$

To prove the above identity, we use several tools, the decoupling formula, martingale method etc. Use of those tools is briefly described below.

The decoupling formula can be found in [19, 17]. Let $\xi$ be a real-valued random variable such that $E[|\xi|^{p+2}] < \infty$, and $f : \mathbb{R} \to \mathbb{C}$ be a function with $p + 1$ continuous and bounded derivatives. Then the decoupling formula is given by

$$E[\xi f(\xi)] = \sum_{i=0}^{p} \frac{\kappa_i + 1}{i!} E[f^{(i)}(\xi)] + \epsilon_p,$$

where $\kappa_j$ are the cumulants of $\xi$ and $\epsilon_p \leq C_p \sup_t |f^{(p+1)}(t)|E[|\xi|^{p+2}]$, $C_p \leq \frac{1+(3+2p)^{p+2}}{(p+1)!}$. In particular, if $\xi$ is a Gaussian random variable with zero mean then we have $E[\xi f(\xi)] = E[\xi^2]E[f'(\xi)]$. One example of the use of this method is given below.

Let the function $\hat{\phi} : \mathbb{R} \to \mathbb{R}$ in theorem 0.1 be sufficiently smooth so that we can apply the Fourier inversion formula $\phi(\lambda) = \frac{1}{2\pi} \int_\mathbb{R} e^{it\lambda} \hat{\phi}(t) \, dt$. Then we can write the linear eigenvalue statistics $\sum_{k=1}^{n} \phi(\lambda_k)$ as

$$\sum_{k=1}^{n} \phi(\lambda_k) = \frac{1}{2\pi} \int_\mathbb{R} \sum_{k=1}^{n} e^{it\lambda_k} \hat{\phi}(t) \, dt = \frac{1}{2\pi} \int_\mathbb{R} \left[ \text{Tr} e^{itM_n} \right] \hat{\phi}(t) \, dt = \frac{1}{2\pi} \int_\mathbb{R} u_n(t) \hat{\phi}(t) \, dt,$$
where \( u_n(t) = \text{Tr}[\exp(itM_n)] \). Then we can write
\[
S_n^o = \sqrt{\frac{b_n}{n}} \left[ \sum_{i=1}^{n} \phi(\lambda_i) - \mathbb{E} \left( \sum_{i=1}^{n} \phi(\lambda_i) \right) \right]
\]
\[
= \frac{1}{2\pi} \sqrt{\frac{b_n}{n}} \int_{\mathbb{R}} u_n^o(t) \phi(t) \, dt,
\]
where \( u_n^o(t) = u_n(t) - \mathbb{E}[u_n(t)] \). Now we notice that \( e^{itM_n} = I + \int_0^t M_n e^{isM_n} \, ds \) and therefore
\[
u_n(t) = \text{Tr} \left[ e^{itM_n} \right] = n + i \int_0^t \sum_{ij} (M_n)_{ij} (e^{isM_n})_{ij} \, ds.
\]
Now we can estimate \( \mathbb{E}[(M_n)_{ij}(\exp(isM_n))_{ij}] \) using the decoupling formula. There are some technical limitations of this method. For example, it does not work in the regime \( b_n << \sqrt{n} \). Soshnikov, and Li demonstrated this technique in their paper \[27\].

Another interesting method is the martingale difference technique. We know that if \( \{X_n\}_{n \geq 0} \) is a martingale
\[
\mathbb{E}[(X_{n+1} - X_n)(X_n - X_{n-1})] = 0
\]
for all \( n \geq 1 \). In other words the fluctuations are uncorrelated. Although this statement looks quite simple, but can be very useful in some cases. For example, if we know the second moments of the fluctuations then we can compute the variance of \( X_n \) by the formula \( \text{Var}(X_n) = \sum_{k=1}^{n} \mathbb{E}[(X_k - X_{k-1})^2] \), which can be obtained as a simple consequence of the facts \[3\] and \( X_n - \mathbb{E}[X_n] = \sum_{k=1}^{n} (X_k - X_{k-1}) \). In the context of Random Matrix Theory, this tool can be used to estimate a lot of things. For example, whenever we use Stieltjes Matrix Theory we transform have to come across the resolvent of the random matrix \( M_n \), which is defined by \( G_n(z) = (z - M_n)^{-1} \). Let us denote the trace of the resolvent by \( \gamma_n(z) := \text{Tr}(G_n(z)) \). It is very difficult to handle the variance of \( \gamma_n(z) \) directly. But we notice that \( \{\mathbb{E}[\gamma_n|F_k]\}_k \) is a martingale, where \( F_k \) is the sigma algebra generated by the main bottom \( k \times k \) submatrix of \( M \). Now we can estimate the variance of \( \gamma_n(z) \) by writing
\[
\text{Var}(\gamma_n(z)) = \sum_{k=1}^{n} \mathbb{E} \left[ |\mathbb{E}[\gamma_n|F_k] - \mathbb{E}[\gamma_n|F_{k-1}]|^2 \right].
\]
Since \( \mathbb{E}[\gamma_n|F_k] - \mathbb{E}[\gamma_n|F_{k-1}] \) is a quantity related to a rank one matrix, it is relatively easy to estimate this. M. Shcherbina had demonstrated this technique in her paper \[27\].

**Current research**

The limiting spectral distribution of a symmetric random matrices follow the Semicircle law. However if all the matrix entries are iid random variables then the limiting spectral distribution follows the Circular law. In other words, if \( M_n \) is an \( n \times n \) random matrix with iid random variables with mean zero and unit variance, then the scaled histogram of the eigenvalues of \( \frac{1}{\sqrt{n}} M_n \) converges the uniform distribution over the unit circle \( \{z \in \mathbb{C} : |z| \leq 1\} \). This was achieved by a series of work by Ginibre, Mehta, Girko, Bai, Tao et al. \[12\,3\,33\,36\]. More precisely speaking, if \( M_n \) is an \( n \times n \) random matrix with iid entries with mean zero and unit variance, and \( \lambda_1, \ldots, \lambda_n \) are the eigenvalues of \( \frac{1}{\sqrt{n}} M_n \), then the empirical spectral distribution \( \mu_n \) converges to the uniform distribution on the unit circle on the complex plane, where
\[
\mu_n(x, y) := \frac{1}{n} \#\{\lambda_i, 1 \leq i \leq n : \Re(\lambda_i) \leq x, \Im(\lambda_i) \leq y\}.
\]
The main idea, which was pioneered by Girko, is that
\[
\int \int e^{i(ux+vy)} \mu_n(dx,dy) = \frac{u^2 + v^2}{4\pi n} \int \int \frac{\partial}{\partial s} \left[ \int_0^\infty \log x \nu_n(dx,z) \right] e^{i(ux+vt)} \ dt ds,
\]
where \(\nu_n(\cdot, z)\) is the ESD of \(\left( \frac{1}{\sqrt{n}} M_n - zI \right) \left( \frac{1}{\sqrt{n}} M_n - zI \right)^*\), and \(z = s + it\). If we can analyse \(\mu_n\) via \(\nu_n\).

The main difficulty was to control the smallest singular value of \(\frac{1}{\sqrt{n}} M_n\). I expected that the circular law should hold for random band matrices too. So I started digging on the problem. Recently, I have found that the limiting spectral distribution of \(\left( \frac{1}{\sqrt{n}} M_n + R \right) \left( \frac{1}{\sqrt{n}} X_n + R \right)^*\) type of matrices, where \(M_n\) is a random band matrix and \(R\) is a deterministic matrix. The limiting spectral distribution follows an integral which is same as described in [8]. In particular, limiting spectral distribution of \(\frac{1}{n} M_n M_n^*\) follows Marchenko-Pastur law [15]. The main difficulty to handle band matrices is that they lack several symmetry properties unlike the full matrices. Now, I am looking at the smallest singular value of random band matrices.

A generalization of the Circular law is the Elliptic law [24]. If the \(ij\)th and \(ji\)th the entries of the random matrix are correlated, then the limiting spectral distribution of such matrices converges to an ellipse which is determined by the correlation coefficient of \(ij\)th and \(ji\)th entries. Currently, I am working on the extension of this result to random band matrices.

During my Masters program, I was working on matching problems. Take \(n\) red points and \(n\) blue points chosen independently and uniformly in \([0, 1]^2\). There are \(n!\) many possible matchings between these red and blue points. We investigate the optimal average matched edge length and the minimum of the maximum matched edge lengths, where the minimum is over all possible \(n!\) matchings. There we see that the optimal average matched edge length is like \(\sqrt{\log n / n}\) and the min-max matched edge length is like \((\log n)^{3/4} / \sqrt{n}\). A series of work have been done by Ajtai, Holroyd, Talagrand et. al. [1, 13, 34]. Now I am looking at the same problem from a Random Matrix Theoretic viewpoint. More precisely, I am looking at the matrix obtained by the mutual distance of \(n\) red and \(n\) blue points. The entries of this matrix is not independent. I am interested in the spectral properties of this matrix. And what can I infer about the matched distances from the spectral properties.

In 1990, Casati et. al. conjectured that the eigenvectors of a Hermitian random band matrix is localized if the bandwidth is less than \(\sqrt{n}\) and delocalized if the bandwidth is greater than \(\sqrt{n}\) [7]. Here ‘localized eigenvector’ means that only a few entries of the eigenvector are non-zero. A series of work has been done on this topic, but the full conjecture has not been proved yet [26, 9, 10]. I consider this problem as one of my long time research goals. I am also interested in interdisciplinary research where I can apply the tools which I have learned in Random Matrix Theory.

Indrajit Jana
Davis, CA.

References


