Honors 1: Undergraduate Math Lab*

Peter Sarnak[†], Steven Miller[‡], Alex Barnett[§]

October 9th, 2002

Courant Institute of Mathematical Sciences
New York University
New York, NY

Abstract

Joint Density Function for eigenvalues of real symmetric matrices; spacing of eigenvalues for 2×2 real symmetric matrices; Semi-Circle Rule. Lecture by Steven Miller; notes by Steven Miller and Alex Barnett.

1 Joint density function of eigenvalues of real symmetric matrices ('GOE')

1.1 Dirac Notation

The derivation handed out in lecture used physics notation which should be explained. The matrix is called the 'Hamiltonian' (meaning that it happened to arise in a quantum physics problem). Vectors are often called *states* (referring to quantum states), however they can be thought of as your usual vectors. (Quantum mechanics is just linear algebra, amazingly). A general vector in 2D is written

$$|u\rangle \quad \text{equivalent to} \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{1}$$

*Homepage: `http://www.math.nyu.edu/~millerj/`
†E-mail: `sarnak@math.princeton.edu`
‡E-mail: `millerj@cims.nyu.edu` or `sjmiller@math.princeton.edu`
§E-mail: `barnett@cims.nyu.edu`

the latter being its coordinate representation in some basis. The unit vectors are

$$|1\rangle\,,\ |2\rangle \quad \text{equivalent to} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2}$$

The $|u\rangle$ is a column vector, and $\langle u| \ \equiv \ |u\rangle^T$ is a row vector. Inner product can be written as $\langle v\,|u\rangle = \mathbf{v}^T \cdot \mathbf{u}$. General bilinear product can be written $\langle v\,|M|\,u\rangle = \mathbf{v}^T \cdot M \cdot \mathbf{u}$.

1.2 2×2 Gaussian Orthogonal Ensemble (GOE)

We consider 2×2 real symmetric matrices,

$$A \ \equiv \ \begin{pmatrix} x & y \\ y & z \end{pmatrix}. \tag{3}$$

Understanding this case is *vital* to building intuition about Random Matrix Theory for $N \times N$ matrices.

A can always be diagonalized by an orthogonal matrix Q as follows,

$$Q^T \begin{pmatrix} x & y \\ y & z \end{pmatrix} Q \ = \ \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \ \equiv \ D. \tag{4}$$

In 2×2 case, the characteristic equation $\det(A - \lambda I) = 0$ is quadratic:

$$\lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) \ = \ 0, \tag{5}$$

where

$$\mathrm{Tr}(A) = x + z,\ \det(A) = xz - y^2. \tag{6}$$

Solutions are

$$\lambda_{1,2} \ = \ \frac{x+z}{2} \pm \sqrt{\left(\frac{x-z}{2}\right)^2 + y^2}, \tag{7}$$

where 1 is the $+$ case, 2 the $-$.

If the two eigenvalues are equal, we say the matrix is degenerate. Initially we are in a three-dimensional space (as x, y and z are arbitrary). Degeneracy requires that x, y and z satisfy

$$\left(\frac{x-z}{2}\right)^2 + y^2 = 0, \tag{8}$$

or, equivalently,

$$x - z = 0, \; y = 0. \tag{9}$$

Thus, we lose two degrees of freedom, because there are two equations which must be satisfied. The set of solutions is $\{(x, y, z) = (x, 0, x)\}$.

Exercise 1.1. *Show that $\lambda_1 - \lambda_2$ is twice the distance from the origin in this 2D subspace.*

Corresponding eigenvectors are,

$$\mathbf{v}_1 = \begin{pmatrix} c \\ s \end{pmatrix}, \qquad\qquad \mathbf{v}_2 = \begin{pmatrix} -s \\ c \end{pmatrix}. \tag{10}$$

We use abbreviations $c \equiv \cos\theta$ and $s \equiv \sin\theta$.

Why can we write the eigenvectors as above? We can always normalize the eigenvector attached to a given eigenvalue to have length 1. We have previously shown that, if the eigenvalues are distinct, then the eigenvectors of a real symmetric matrix are perpendicular. This forces the above form for the two eigenvectors, at least when $\lambda_1 \neq \lambda_2$.

One rotation angle θ defines the orthogonal matrix,

$$Q = Q(\theta) = \begin{pmatrix} \mathbf{v}_1 \; \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}. \tag{11}$$

The structure of the eigenvectors is actually quite rich.

Exercise 1.2. *Find θ in terms of x, y, z. Hint: use trigonometric identities to simplify the resulting form. Hint: solve $(A - \lambda_1 \mathbf{v}_1) = \mathbf{0}$.*

Exercise 1.3. *Show that a general A can be written*

$$A \; = \; \alpha \begin{pmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{pmatrix} \; + \; \gamma \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{12}$$

Exercise 1.4. *Find $\lambda_{1,2}$ in terms of α, β, γ. Show that the eigenvector angle is given by $\theta = \beta/2$. This result is quite deep; for instance notice that taking a complete 2π cycle in β reverses the signs of the eigenvectors! This isn't that relevant for the rest of this lecture.*

We adopt two assumptions about the joint distribution over A, called $p(A) \equiv p(x, y, z)$:

1. Invariance of p under orthogonal transformations (aka 'basis-invariance'), $p(M^TAM) = p(A)$ for all orthogonal M.
2. Independence of distributions of individual matrix elements, $p(x, y, z) = p_x(x)p_y(y)p_z(z)$.

Section 2C-1 of the handout reminds us that these two assumptions taken together demand a unique form of distribution,

$$p(x, y, z) \;\propto\; e^{-C\mathrm{Tr}(A^2)}, \tag{13}$$

depending on only one parameter C; we choose $C = 1$. Note $\propto$ means proportional to; the constant of proportionality is what is needed to make $P9x, y, z)$ a probability distribution (ie, the integral of $\int\int\int p(x, y, z)dxdydz = 1$.

This corresponds to Gaussian distributions of matrix elements,

$$\begin{aligned} p_y(y) &= \sqrt{\frac{2}{\pi}}e^{-2y^2} && \text{off-diag} \\ p_x(x) = p_z(x) &= \frac{1}{\sqrt{\pi}}e^{-x^2} && \text{diag.} \end{aligned} \tag{14}$$

Note that the diag elements have variance $\frac{1}{2}$, the off-diag variance $\frac{1}{4}$. We show how to compute the normalization prefactors later on. This form (for general C) is the so-called GOE. The $n \times n$ case is derived in Miller's handout of 9/25/02.

1.3 Transformation to diagonal representation

The operation of diagonalizing A can be viewed as the transformation from one 3D space to another 3D space,

$$\mathbf{r} \equiv \underbrace{(x, y, z)}_{A} \;\longleftrightarrow\; \mathbf{r}' \equiv \underbrace{(\lambda_1, \lambda_2, \theta)}_{D,Q}. \tag{15}$$

This is 1-to-1 apart from the set of measure zero (ie, a lower dimensional subspace) corresponding to degenerate eigenvalues. Looking at Eq. 4 we can see the transformation is linear in the eigenvalues, nonlinear in θ. We are interested in the *marginal* distribution of the eigenvalues,

$$p'(\lambda_1, \lambda_2) \;\equiv\; \int d\theta\, p'(\lambda_1, \lambda_2, \theta), \tag{16}$$

in other words we don't care what θ is. We use primes to signify distributions over final (Q, D) variables.

We want to know how to transform probability density from $\mathbf{r}$ space to $\mathbf{r}'$ space. In general this must follow the law,

$$p(\mathbf{r})d\mathbf{r} \;=\; p'(\mathbf{r}')d\mathbf{r}', \tag{17}$$

giving

$$p'(\mathbf{r}') \;=\; \det(J)p(\mathbf{r}). \tag{18}$$

The ratio of the volume elements is $|\det J|$ where J is the 3×3 Jacobean matrix of the transformation. J has elements $J_{ij} = \partial r_j / \partial r'_i$.

Inverting Eq. 4 we can write $A(\mathbf{r}')$ as

$$\begin{aligned} \begin{pmatrix} x & y \\ y & z \end{pmatrix} = QDQ^T \;&=\; \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \\ &=\; \begin{pmatrix} \lambda_1 c^2 + \lambda_2 s^2 & (\lambda_1 - \lambda_2)sc \\ (\lambda_1 - \lambda_2)sc & \lambda_1 s^2 + \lambda_2 c^2 \end{pmatrix} \end{aligned} \tag{19}$$

We evaluate J for this case,

$$J \;\equiv\; \begin{pmatrix} \frac{\partial x}{\partial \lambda_1} & \frac{\partial y}{\partial \lambda_1} & \frac{\partial z}{\partial \lambda_1} \\ \frac{\partial x}{\partial \lambda_2} & \ddots & \\ \frac{\partial x}{\partial \theta} & & \end{pmatrix}. \tag{20}$$

We see λ's only appear in the bottom three entries, and furthermore they only appear as factors $(\lambda_1 - \lambda_2)$ in each entry.

Exercise 1.5. *Evaluate the bottom row of J to prove the above.*

Therefore this factor of a row of J can be brought out in evaluating the determinant:

$$\det(J) = \left| \text{messy } \theta\text{-dep } 3 \times 3 \text{ matrix} \right| \cdot (\lambda_1 - \lambda_2) \;=\; g(\theta)(\lambda_1 - \lambda_2). \tag{21}$$

Warning! The Jacobian is the absolute value of the determinant. Thus, we need $|\lambda_1 - \lambda_2|$ above, or we need to adopt the convention that we label the eigenvalues so that $\lambda_1 \geq \lambda_2$.

The only dependence on the λ's is given by the second factor. Plugging into Eq. 18 and marginalizing over θ gives,

$$\begin{aligned} p'(\lambda_1, \lambda_2) &= \int d\theta\, g(\theta)\, (\lambda_1 - \lambda_2) e^{-(\lambda_1^2+\lambda_2^2)} \\ &\propto (\lambda_1 - \lambda_2) e^{-(\lambda_1^2+\lambda_2^2)}. \end{aligned} \tag{22}$$

Note that we do not need the absolute value sign around $(\lambda_1 - \lambda_2)$ because we chose $\lambda_1 > \lambda_2$. This is the joint density of the eigenvalues in 2×2 GOE.

1.4 Generalization to $n \times n$ case

The above generalizes quite easily, with the dimension of the two spaces being $N = \frac{1}{2}n(n+1)$. The $\frac{1}{2}n(n+1)$ degrees of freedom in A equal n degrees of freedom in D (namely the eigenvalues $\{\lambda_i\}$) plus $\frac{1}{2}n(n-1)$ degrees of freedom in Q (namely the generalized angles Ω). We sort the eigenvalues such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

Theorem 1.6. *If $\lambda_i = \lambda_j$ for some $1 \leq i < j \leq n$, then the Jacobean of the transformation $A \leftrightarrow (D, Q)$ vanishes, that is* $\det(J) = 0$.

The proof relies on realising that the two eigenvectors $\mathbf{v}_i$ and $\mathbf{v}_j$ span a 2D subspace, invariant under A. (Recall here we refer to a subspace of the n-dim vector space upon which A operates by multiplication). The invariance means that the choices of directions of the eigenvectors is arbitrary in this 2D plane. Therefore there is one angle degree of freedom in Ω which in not constrained by A, that is, it is independent of the elements of A. Now think of the inverse transformation from $(D, Q) \to A$. An infinitesimal volume element is transformed as

$$d\mathbf{r} = \det(J) d\mathbf{r}'. \tag{23}$$

Changes of eigenvector angle within the 2D subspace have no effect on A, so the volume element $d\mathbf{r}$ is collapsed to zero. (Another way of putting this is that J acquires a null-space of dimension 1). Therefore $\det(J) = 0$. $\Box$.

This vanishing of the Jacobean at degeneracies renders the non-uniqueness of the forward map $A \to (D, Q)$ at these points harmless in the following.

The upper n rows of J are messy functions of angles Ω, and the bottom $\frac{1}{2}n(n-1)$ rows contain entries each which is *linear* in the eigenvalues. Therefore $\det(J)$ is a polynomial of degree $\frac{1}{2}n(n-1)$ in the eigenvalues λ_i. Further, $\det(J) = 0$ if any two eigenvalues are equal.

Consider the polynomial $\prod_{1 \leq i < j \leq n}(\lambda_i - \lambda_j)$. First, note that this polynomial vanishes whenever two eigenvalues are the same. We claim it is a polynomial of degree $\frac{1}{2}n(n-1)$ in the eigenvalues. For each j, there are $j-1$ choices for i. Thus, the degree is

$$\sum_{j=2}^{n} j - 1 = \sum_{k=1}^{n-1} k = \frac{(n-1)(n-1+1)}{2} = \frac{n(n-1)}{2}. \tag{24}$$

Thus, $\det(J)$ and $\prod_{1 \leq i < j \leq n}(\lambda_i - \lambda_j)$ both vanish whenever two eigenvalues are equal, and they have the same degree. Therefore, they must be scalar multiples of each other.

So,

$$\det(J) \propto \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j). \tag{25}$$

Combining with the GOE form of $p(A)$ gives, after marginalizing over Ω as before,

$$p(\{\lambda_i\}) = \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j) \cdot e^{-\sum_{i=1}^{n} \lambda_i^2}. \tag{26}$$

The vanishing of this probability density as any two eigenvalues come close is called *level repulsion.*

2 Eigenvalue spacing distribution in 2×2 real symmetric matrices

2.1 Reminder: Integral of the Gaussian

We want

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx. \tag{27}$$

Square it and rearrange the summation over area by using polar coordinates:

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-x^2} dx \cdot \int_{-\infty}^{\infty} e^{-y^2} dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2} dx\, dy \\ &= \int_{0}^{2\pi} d\theta \int_{0}^{\infty} r\, dr\, e^{-r^2} = 2\pi \cdot \left[-\frac{1}{2} e^{-r^2} \right]_0^{\infty} \\ &= \pi. \end{aligned} \tag{28}$$

Introduction of the radius factor r produced re^{-r^2}, a known differential. So,

$$I \;=\; \sqrt{\pi}. \tag{29}$$

Changing the variable in the above, and rearranging, gives

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx \;=\; 1. \tag{30}$$

This is therefore the correct normalization for a 1D Gaussian probability density, of variance σ^2.

2.2 Spacing distribution

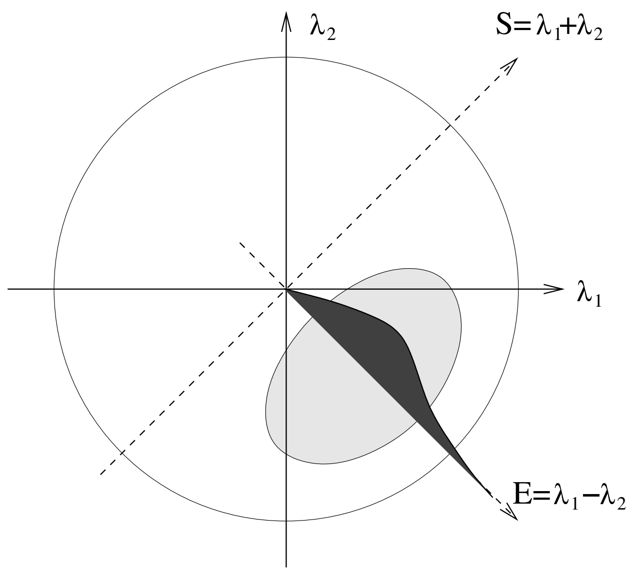

Figure 1: Change of coordinates to get spacing distribution in E. Light shading suggests form of density across the 2D plane. Dark shading shows a graph of its projection onto the E axis.

Here for convenience we present a slightly simpler derivation than in lecture. Given the 2D density $p'(\lambda_1, \lambda_2)$ we want the 1D density of the difference $E \equiv \lambda_1 - \lambda_2$. This will require marginalizing again, since there is a reduction in dimensionality. We define $S \equiv \lambda_1 + \lambda_2$. The linear transformation

$(\lambda_1, \lambda_2) \to (E, S)$ has fixed Jacobean (it is a rotation by -45^o and a compression by $\sqrt{2}$ in each axis). See Fig. 1.

Therefore, substituting in $\lambda_1 = (S+E)/2$ and $\lambda_1 = (S-E)/2$ into Eq. 22 gives

$$\begin{aligned} p'(E,S) &\propto p'(\lambda_1(E,S), \lambda_2(E,S)) \;=\; Ee^{-\frac{1}{4}[(S+E)^2+(S-E)^2]} \\ &= Ee^{-E^2/2} \cdot e^{-S^2/2}, \end{aligned} \tag{31}$$

which is separable. Therefore integrating over S gives an E-independent number, and

$$p'(E) \;\propto\; Ee^{-E^2/2}. \tag{32}$$

This is the so-called 'Wigner Surmise' for the eigenvalue spacing density. Remarkably, in the $n \times n$ case, even for large n, this density is very close to the true spacing distribution of adjacent eigenvalues. The limiting powerlaw $\lim_{E\to 0} p'(E) \propto E^\beta$ with $\beta = 1$ is intimately related to the matrix symmetry class GOE that we close. It is also possible to achieve $\beta = 2$ and $\beta = 4$ by choosing different symmetry classes.

Finally, let's say you couldn't be bothered to construct your second variable S. Instead you could derive the above using the Dirac delta-function (see below) to marginalize:

$$p'(E) \;=\; \int_{-\infty}^{\infty} d\lambda_1 \int_{-\infty}^{\lambda_1} d\lambda_2 \, p'(\lambda_1, \lambda_2) \, \delta(E - (\lambda_1 - \lambda_2)). \tag{33}$$

Apart from the unusual limits (due to ordering of eigenvalues), this is the standard procedure to extract a marginal density.

Exercise 2.1. *Simplify the above to arrive at* $p'(E)$.

3 Delta Function(al)

Let $f(x)$ be a nice function; for example, let $f(x)$ be an infinitely differentiable function whose Taylor Series converges to $f(x)$:

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots \tag{34}$$

Let

$$\delta_n(x) = \begin{cases} n & \text{if } x \in \left[-\frac{1}{2n}, \frac{1}{2n}\right] \\ 0 & \text{otherwise} \end{cases} \tag{35}$$

Exercise 3.1. *Show that*

$$\int_{-\infty}^{\infty} f(x)\delta_n(x)dx = f(0) + O\Big(\frac{1}{n}\Big). \tag{36}$$

Let δ be the limit as $n \to \infty$ of δ_n. We find / define

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(x)\delta_n(x)dx \;=\; \int_{-\infty}^{\infty} f(x)\delta(x)dx \;=\; f(0). \tag{37}$$

Exercise 3.2. *Show that*

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx \;=\; f(a). \tag{38}$$

A good analogy for the δ functional is a point mass. A point mass has no extension (no length, width or height) but finite mass. Therefore, a point mass has infinite density.

A probability density must integrate to one. This corresponds to $\int 1 \cdot \delta(x)dx = 1$. We often refer to $\delta(x)$ as a point mass at the origin, and $\delta(x-a)$ as a point mass at a.

4 Definition of the Semi-Circle Density

Consider

$$P(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2} & \text{if } |x| \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{39}$$

Exercise 4.1. *Show that $P(x)$ is a probability density. IE, show that it is non-negative and integrates to 1. Graph $P(x)$.*

We call $P(x)$ the semi-circle density.

5 Semi-Circle Rule: Preliminaries

Let λ_j be the eigenvalues of a real, symmetric $N \times N$ matrix A. We normalize the eigenvalues of A by dividing by $2\sqrt{N}$.

Define

$$\mu_{A,N}(x) = \frac{1}{N} \sum_{j=1}^{N} \delta\Big(x - \frac{\lambda_j(A)}{2\sqrt{N}}\Big). \tag{40}$$

$\delta\Big(x - \frac{\lambda_j(A)}{2\sqrt{N}}\Big)$ is a point mass at $\frac{\lambda_j(A)}{2\sqrt{N}}$. By summing these point masses and dividing by N, we have a probability distribution. For example,

$$\int_{-\infty}^{\infty} f(x)\mu_{A,N}(x)dx = \sum_{j=1}^{N} f\Big(\frac{\lambda_j(A)}{2\sqrt{N}}\Big). \tag{41}$$

We will show that, as $N \to \infty$, the above converges to the integral of f against the semi-circle density:

$$\int_{-\infty}^{\infty} f(x)P(x)dx. \tag{42}$$

What does this mean?

$$\sum_{j=1}^{N} f\Big(\frac{\lambda_j(A)}{2\sqrt{N}}\Big) \tag{43}$$

looks like a Riemann Sum. The statement that, for nice $f(x)$,

$$\sum_{j=1}^{N} f\Big(\frac{\lambda_j(A)}{2\sqrt{N}}\Big) \to \int_{-\infty}^{\infty} f(x)P(x)dx \tag{44}$$

means that as $N \to \infty$, the number of eigenvalues of a random A in $[a, b]$ equals

$$\int_{a}^{b} P(x)dx. \tag{45}$$

Theorem 5.1. *Choose the entries a_{ij} of a real, symmetric matrix independently from a fixed probability distribution p with mean zero, variance one, and finite higher moments. For each A, form the probability measure $\mu_{A,N}$.*

As $N \to \infty$, with probability one the measures $\mu_{A,n}(x)dx$ converge to the semi-circle probability $P(x)dx$.

This is not the most general version; however, it is rich enough for our purposes.

6 Sketch of Proof of the Semi-Circle Law

6.1 Calculation of Moments via Trace Formula

We will show that the expected value of the moments of the $\mu_{A,N}(x)$ equal the moments of the semi-circle.

Definition 6.1. *$M_{A,N}(k)$ is the k^{th} moment of the probability measure attached to $\mu_{A,N}(x)dx$:*

$$M_{A,N}(k) \;=\; \int x^k \mu_{A,N}(x)dx \;=\; \frac{1}{N}\sum_{j=1}^{N}\Big(\frac{\lambda_j(A)}{2\sqrt{N}}\Big)^k. \tag{46}$$

Note that $\sum \lambda_j(A)^k = \text{Trace}(A^k)$. Thus, we have

$$M_{A,N}(k) \;=\; \frac{1}{2^k N^{1+\frac{k}{2}}}\text{Trace}(A^k). \tag{47}$$

We now calculate the expected values of the first few moments ($k = 0, 1$, 2 and 3).

Lemma 6.2. *The expected value of $M_{A,N}(0) = 1$.*

Proof:

$$E\Big[M_{A,N}(0)\Big] \;=\; \frac{1}{N}E\Big[\text{Trace}(I)\Big] \;=\; 1. \tag{48}$$

Note that summing the eigenvalues to the zeroth power is the same as taking the trace of the identity matrix. □

Lemma 6.3. *The expected value of $M_{A,N}(1) = 0$.*

Proof:

$$\begin{aligned} E\Big[M_{A,N}(1)\Big] &= \frac{1}{2N^{3/2}}E\Big[\mathrm{Trace}(A)\Big] \\ &= \frac{1}{2N^{3/2}}E\Big[\sum_i a_{ii}\Big] \\ &= \frac{1}{2N^{3/2}}\sum_i E[a_{ii}] \;=\; 0, \end{aligned} \tag{49}$$

because we have assumed that each a_{ij} is drawn from a probability distribution with mean zero. □

Lemma 6.4. *The expected value of* $M_{A,N}(2) = \frac{1}{4}$.

Proof: Note that

$$\mathrm{Trace}(A^2) = \sum_i \sum_j a_{ij}a_{ji}. \tag{50}$$

As our matrix is symmetric, $a_{ij} = a_{ji}$. Thus, the trace is $\sum_i \sum_j a_{ij}^2$. Now

$$\begin{aligned} E\Big[M_{A,N}(2)\Big] &= \frac{1}{4N^2}E\Big[\mathrm{Trace}(A^2)\Big] \\ &= \frac{1}{4N^2}E\Big[\sum_i \sum_j a_{ij}^2\Big] \\ &= \frac{1}{4N^2}\sum_i \sum_j E[a_{ij}^2] \;=\; \frac{1}{4}, \end{aligned} \tag{51}$$

where the last line follows from each a_{ij} has variance 1. As their means are zero, the variance $E[a_{ij}^2] - E[a_{ij}]^2 = 1$ implies $E[a_{ij}] = 1$. There are N^2 pairs (i, j). Thus, we have $\frac{1}{4N^2} \cdot (N \cdot 1) = \frac{1}{4}$. □

Lemma 6.5. *The expected value of* $M_{A,N}(3) = 0$ *as* $N \to \infty$.

We need

$$\mathrm{Trace}(A^3) \;=\; \sum_i \sum_j \sum_k a_{ij}a_{jk}a_{ki}. \tag{52}$$

We find

$$\begin{aligned}
E\Big[M_{A,N}(3)\Big] &= \frac{1}{8N^{2.5}}E\Big[\text{Trace}(A^3)\Big] \\
&= \frac{1}{8N^{2.5}}E\Big[\sum_i\sum_j\sum_k a_{ij}a_{jk}a_{ki}\Big] \\
&= \frac{1}{8N^{2.5}}\sum_i\sum_j\sum_k E[a_{ij}a_{jk}a_{ki}]. \qquad (53)
\end{aligned}$$

There are three cases. If the subscripts i, j and k are all distinct, then a_{ij}, a_{jk}, and a_{ki} are three independent variables. Hence

$$E[a_{ij}a_{jk}a_{ki}] \;=\; E[a_{ij}]\cdot E[a_{jk}]\cdot E[a_{ki}] \;=\; 0. \qquad (54)$$

If two of the subscripts are the same (say $i = j$) and the third is distinct, we have

$$E[a_{ii}a_{ik}a_{ki}] \;=\; E[a_{ii}]\cdot E[a_{ik}^2] \;=\; 0\cdot 1 \;=\; 0. \qquad (55)$$

If all three subscripts are the same, we have

$$E[a_{ii}^3] \qquad (56)$$

This is the third moment of a_{ii}. It is the same for all variables a_{ii}, and is finite by assumption. There are N triples where $i = j = k$.

Thus,

$$E\Big[M_{A,N}(3)\Big] \;=\; \frac{1}{8N^{2.5}}\cdot NE[a_{11}^3] \;=\; \frac{E[a_{11}^3]}{8}\cdot\frac{1}{N^{1.5}}. \qquad (57)$$

Thus, as $N \to \infty$, the expected value of the third moment is zero. $\Box$

To calculate the higher moments requires significantly more delicate combinatorial arguments.

6.2 Calculation of Moments from the Semi-Circle

We now calculate the moments of the semi-circle. For $k \leq 3$, the k^{th} moment of the semi-circle $C(k)$ equals the expected k^{th} moment of $\mu_{A,N}(x)$ as $N \to \infty$.

$$C(k) \;=\; \int_{-\infty}^{\infty} x^k P(x)dx \;=\; \frac{2}{\pi}\int_{-1}^{1} x^k\sqrt{1-x^2}dx. \tag{58}$$

We note that, by symmetry, $C(k) = 0$ for k odd, and $C(0) = 1$ as $P(x)$ is a probability density.

For $k = 2m$ even, we change variables $x = \sin\theta$.

$$C(2m) \;=\; \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2m}\theta\cdot\cos^2\theta d\theta. \tag{59}$$

Using $\sin^2\theta = 1 - \cos^2\theta$ gives

$$C(2m) \;=\; \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2m}\theta d\theta \;-\; \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2m+2}\theta d\theta. \tag{60}$$

The above integrals can be evaluated exactly. We constantly use

$$\begin{aligned} \cos^2(\phi) &= \frac{1}{2} + \frac{1}{2}\cos(2\phi) \\ \sin^2(\phi) &= \frac{1}{2} - \frac{1}{2}\cos(2\phi). \end{aligned} \tag{61}$$

Repeated applications of the above allow us to write $\sin^{2m}(\theta)$ as a linear combination of $1, \cos(2\theta), \ldots, \cos(2m\theta)$.

Let

$$n!! = \begin{cases} n\cdot(n-2)\cdots 2 & \text{if } n \text{ is even} \\ n\cdot(n-2)\cdots 1 & \text{if } n \text{ is odd} \end{cases} \tag{62}$$

We find (either prove directly or by induction) that

$$\frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2m}\theta d\theta \;=\; 2\frac{(2m-1)!!}{(2m)!!}. \tag{63}$$

Exercise 6.6. *Show the above gives*

$$C(2m) \;=\; 2\frac{(k-1)!!}{(k+2)!!}. \tag{64}$$

Also, show $C(2)$ agrees with our earlier calculation.