

Chapter 1

Spectral Statistics

1.1 Introduction

So, what are *spectral statistics*? Do they have to do with rainbows? Sceptres? No, they don't, but they're almost as colorful and regal. The word spectral is borrowed from the spectral-like patterns observed in statistical physics - whether it may be atomic spectra or other quantum mechanical phenomena. The borrowing is loose and not literal, but still somewhat well founded. In fact, the field of Random Matrix Theory was extensively developed in the 1930s by the nuclear physicist Eugene Wigner. He found connections between the deterministic properties of atomic nuclei and their random and stochastic behaviors. The link? Random matrices.

So in the context of this thesis, *spectral statistics* will be an umbrella term for random matrix statistics that somehow involve that matrix's eigenvalues and eigenvectors. That being said, if we fix a *random matrix*, we can study its features by studying its eigenvalues - fundamental numbers that tell us a lot about the matrix. They are quite important for many reasons. For instance in statistical physics, many processes are represented by operators or matrices, and as such, their behaviours could be partially determined by the eigenvalues of their corresponding matrices. The study of eigenvalues and eigenvectors primarily falls in the scope of Linear Algebra, but their utility is far-reaching. So, what exactly are *eigenvalues* exactly?

1.1.1 The Quintessential Spectral Statistic: the Eigenvalue

Given any standard square matrix $P \in \mathbb{F}^{N \times N}$, its *eigenvalues* are simply the roots of the characteristic polynomial $\text{char}_P(\lambda) = \det(P - \lambda I)$. By the Fundamental Theorem of Algebra, we know that there is always have as many complex eigenvalues $\lambda \in \mathbb{C}$ as the dimension of the matrix.

That being said, when our random matrix has a specified distribution (say, standard normal), we can see patterns in the eigenvalue distributions. So, an eigenvalue is a **spectral statistic** of a random matrix! To talk about a matrix's eigenvalues in a more formal and concise manner, we motivate what is the eigenvalue spectrum.

Definition 1.1.1 (Spectrum). Suppose $P \in \mathbb{F}^{N \times N}$ is a square matrix of size N over \mathbb{F} . Then, the (eigenvalue) spectrum of P is defined as the multiset of its eigenvalues and it is denoted $\sigma(P) = \{\lambda_i \in \mathbb{C}\}_{i=1}^N$. Note that it is important to specify that a spectrum is a multiset and not just a set; eigenvalues could be repeated due to algebraic multiplicity and we opt to always have N eigenvalues.

1.1.2 Interlude: Ensembles

While the spectrum of a matrix provides a good summary of the matrix, a matrix is only considered a single point/observation in random matrix theory. Additionally, simulating large matrices and computing their eigenvalues becomes harder and more computationally expensive as $N \rightarrow \infty$. As such, to obtain more eigenvalue statistics efficiently, another dimension is introduced by motivating the *random matrix ensemble*.

Definition 1.1.2 (Random Matrix Ensemble). A \mathcal{D} -distributed random matrix ensemble \mathcal{E} over $\mathbb{F}^{N \times N}$ of size K is defined as a set of \mathcal{D} -distributed random matrices $\mathcal{E} = \{P_i \sim \mathcal{D} \mid P_i \in \mathbb{F}^{N \times N}\}_{i=1}^K$. In simple words, it is simply a collection of iterations of some specified class of random matrix.

Now that matrix ensembles are well defined, we can motivate a core object of our study - the spectrum of a random matrix ensemble. From its name, it is indeed what one might expect it to be.

Definition 1.1.3 (Ensemble Spectrum). If we have an ensemble \mathcal{E} , then we can naturally extend the definition of $\sigma(\mathcal{E})$. To take the spectrum of an ensemble, simply take the union of the spectra of each of its matrices. In other words, if $\mathcal{E} = \{P_i \sim \mathcal{D} \mid P_i \in \mathbb{F}^{N \times N}\}_{i=1}^K$, then $\sigma(\mathcal{E}) = \bigcup_{i=1}^K \sigma(P_i)$.

A common theme in this thesis will be that singleton matrices do not provide insightful information on their own. Rather, it is the collective behavior of a \mathcal{D} -distributed ensemble that tells us about how \mathcal{D} impacts our spectral statistics. So in a way, ensemble statistics are the engine of this research.

1.2 Eigenvalue Spectra

1.2.1 Ordered Spectra

When we motivate the idea of matrix dispersion in the next section, we will consider order statistics of that matrix's eigenvalues in tandem with its dispersion. However, to do so presupposes that we have a sense of what *ordered* eigenvalues means. Take a matrix P and its *unordered* spectrum $\sigma(P) = \{\lambda_j\}$. It is paramount to know what ordering scheme $\sigma(P)$ is using, because otherwise, the eigenvalue indices are meaningless! So, to eliminate confusion, we add an index to σ that indicates how the spectrum is ordered. Often, the ordering context will be clear and the indexing will be omitted. Consider the two following *ordering schema*:

Standard definitions of an ordered spectrum follows the standard ordering in the reals; denote this as the ordering by the **sign scheme**. Note that because total-ordering is only well-defined on the reals, we can only use this scheme when on a spectrum with real entries. So, we write the *sign-ordered spectrum* as follows:

$$\sigma_S(P) = \{\lambda_j : \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N\}_{j=1}^N$$

Alternatively, we can motivate a different scheme that properly handles complex eigenvalues. We could sort the spectrum by the norm of its entries; denote this as ordering using the *norm scheme*. This way, all the eigenvalues are mapped to a real value, in which we could use the sign-scheme of ordering. Without further ado, we write the *norm-ordered spectrum* as follows:

$$\sigma_N(P) = \{\lambda_j : |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_N|\}_{j=1}^N$$

Note that when we take the norms of the eigenvalues, we essentially ignore "rotational" features of the eigenvalues. Signs of eigenvalues indicate reflection or rotation, so when we take the norm, we essentially become more concerned with scaling.

1.2.2 Order Statistics

With eigenvalue ordering unambiguous and well-defined, we may proceed to start talking about their order statistics. In short, given a random sample of fixed size, order statistics are random variables defined as the value of an element conditioning on its rank within the sample. (See A.2.2)

For example, the maximum of a sample is an order statistic concerned with the highest ranked element. In our case, this could correspond the largest eigenvalue of a spectrum. After all, a spectrum is a random sample of fixed size, so this statistic is well-defined. To obtain the distribution of the largest eigenvalue for some distribution \mathcal{D} , we would simulate an ensemble \mathcal{E} and observe λ_1 for each of its matrices.

In general, order statistics are quite useful and tell us a lot about how the eigenvalues distribute given a distribution. They tell us how the eigenvalues space themselves and give us useful upper and lower bounds.

In addition to these simple order statistics, we introduce a new variant statistic called the **ranking difference class**. Instead of observing a single eigenvalue at a given rank, we will now observe a pair of eigenvalues at a time. To standardize the process, we introduce a new equivalence class called the *ranking difference*. As suggested, it is precisely the integer difference of the eigenvalue ranks.

Definition 1.2.1 (Ranking Difference). *The ranking difference is a function $\delta : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ which takes the index of two eigenvalues (from an ordered spectrum) and returns their difference. In other words, $\delta : (\lambda_i, \lambda_j) \mapsto (i - j)$.*

With the function δ , we may take the set of unique eigenvalue pairs $(i > j)$ and partition it into equivalence classes. To do so, we define the equivalence relation \sim_δ which says $(\lambda_a, \lambda_b) \sim_\delta (\lambda_c, \lambda_d) \iff (a - b) = (c - d)$. These equivalence classes then

naturally correspond to pairs a set distance $\rho = i - j$ apart. So, for an $N \times N$ matrix, δ assumes a range $\rho \in \{1, \dots, N - 1\}$.

In summary, \sim_δ takes the set $\{(\lambda_i, \lambda_j) \mid \lambda_i, \lambda_j \in \sigma(P) \text{ and } i > j\}$ and surjectively partitions it onto the equivalence classes $[(\lambda_i, \lambda_j)]_\rho$ for $\rho \in \{1, \dots, N - 1\}$. For example, if we consider $\rho = 1$, then we are considering all pairs of eigenvalue neighbors.

Note that the sizes of each equivalence class are **never equal**. With this partition in mind, we can consider various statistics conditioning on the value of ρ . Conditioning on ρ will be especially useful in the cases where we are considering matrices like the Hermite- β matrices; the eigenvalues of those matrices tend to *repel*, so to speak, and we can observe these patterns using ρ .

1.2.3 Symmetric and Hermitian Matrices

A very important class of matrices in Linear Algebra is that of Symmetric or Hermitian matrices (See A.1.2). Simply put, those are matrices which are equal to their conjugate transpose.

Note: Since real numbers are their own conjugate transpose, every Symmetric matrix is Hermitian. However, we will still delineate the two terms to avoid confusion.

In any case, one critical result in Linear Algebra that will be extensively wielded in this thesis is the fact that a matrix is Symmetric or Hermitian if and only if it has real eigenvalues. In other words:

$$P = \overline{P^T} \iff \sigma(P) = \{\lambda_i \mid \lambda_i \in \mathbb{R}\}$$

Having a complete set of real eigenvalues yields many great properties. For instance, if all eigenvalues are real, we have the option of observing either the sign-ordered spectrum or the norm-ordered spectrum. This way, we can preserve negative signs and we would not lose the rotational aspect of the eigenvalue when we study its statistics. That is just one reason out of many more why having real eigenvalues is quite nice.

1.3 Eigenvalue Dispersions

In this section, we define the final spectral statistic studied in this chapter: eigenvalue dispersion. As the name suggests, these statistics concern the distribution of the spacings between the eigenvalues. Funnily enough, this is almost as literal as it gets when we use the word "spectral". In physics and chemistry, atomic spectra are essentially differences between energy levels or quanta, so the translation is close.

In any case, we motivate a few definitions and formalisms in this section, then motivate the definition of a matrix's eigenvalue dispersion. To start off, we define an object useful for pairing our eigenvalues together, the pairing scheme.

1.3.1 Pairing Schema

Next up, we introduce a new notation for a pairing scheme denoted Π . What are pairing schemes and why do they matter? Recall that our goal is to study the spacings between eigenvalues. If we are studying spacing, then a priori, we are concerned with pairs of eigenvalues! Spacing, after all, is a binary relationship.

So, with the definitions of spectra well motivated, a natural definition of pairing schema follows. Essentially, Π is just a subset of the Cartesian product of a spectrum with itself. In other words, if we denote $\mathbb{S} := \sigma(P)$, then we say that a pairing scheme is simply a subset $\Pi \subseteq \mathbb{S}^2$. Here are some pairing schema that we will consider:

1. Let $\Pi_{>}$ be the set of unique combinations of ordered eigenvalues. This will be the standard ordered pair scheme used in lieu of our dispersion metric argument orders (more later).

$$\Pi_{>} = \{\pi_{ij} = (\lambda_i, \lambda_j) \mid i > j\}_{i=1}^{N-1}$$

2. Let Π_1 be the largest pair of eigenvalues of a spectrum. Nice and simple.

$$\Pi_1 = \{(\lambda_2, \lambda_1)\}$$

3. Let Π_C be the consecutive pairs of eigenvalues in a spectrum. This pairing scheme gives us the minimal information needed to express important bounds and spacings in terms of its elements.

$$\Pi_C = \{\pi_{ij} = (\lambda_i, \lambda_j) \mid i = j + 1\}_{i=1}^{N-1}$$

1.3.2 Dispersion Metrics

When we define the dispersion of a matrix, we will see that there is a free argument d called the dispersion metric. This function d is a general function whose domain is always two eigenvalues. In set notation, this is the set $\mathbb{C} \times \mathbb{C}$ - a pair of two complex numbers. Its range will often be the positive reals \mathbb{R}^+ ; this is because the dispersion metric often will be substitutable with distance metric. Sometimes, the range will be \mathbb{C} . So, the dispersion metric will take the following form:

$$d : \mathbb{C} \times \mathbb{C} \rightarrow \{\mathbb{R}^+, \mathbb{C}\}$$

Consider the following dispersion metrics below. Out of those 4 dispersion metrics, only the first one has a range of \mathbb{C} . The rest have a range of \mathbb{R}^+ . Additionally, the second and third metrics are *symmetric* operations while the rest are not. The β -norm is only a symmetric operation when β is even.

1. The identity difference: $d_{id}(z, z') = z' - z$
2. The standard norm: $d_n(z, z') = |z' - z|$

3. The β -norm: $d_\beta(z, z') = |z' - z|^\beta$
4. The difference of absolutes: $d_{ad}(z, z') = |z'| - |z|$

Finally, we are able to motivate the definition of a matrix dispersion! Suppose we have a \mathcal{D} -distributed random matrix $P \in \mathbb{F}^{N \times N}$ or a random matrix ensemble $\mathcal{E} = \{P_i \sim \mathcal{D} \mid P_i \in \mathbb{F}^{N \times N}\}$. Then we define their dispersion as follows.

1.3.3 Matrix and Ensemble Dispersions

Definition 1.3.1 (Dispersion). *The dispersion of a matrix $P \in \mathbb{F}^{N \times N}$ with respect to a dispersion metric $d : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{F}$ and pairing scheme Π , call it $\Delta_d(P, \Pi)$, is defined as follows. Suppose $\sigma(P) := \mathbb{S}$ is the ordered spectrum of P where $\sigma(P) = \{\lambda_1, \dots, \lambda_N\}$. Then, let $\Pi = \{\pi_{ij} = (\lambda_i, \lambda_j)\} \subseteq \mathbb{S}^2$ be a subset of eigenvalue ordered pairs. Then, the dispersion of P with respect to d is simply the set $\Delta_d(P, \Pi) = \{\delta_{ij} = d(\pi_{ij}) \mid \pi_{ij} = (\lambda_i, \lambda_j) \in \Pi\}$.*

As we usually do, we define the dispersion of an ensemble in a similar fashion.

Definition 1.3.2 (Ensemble Dispersion). *If we have an ensemble \mathcal{E} , then we can naturally extend the definition of $\Delta_d(\mathcal{E}, \Pi)$. To take the dispersion of an ensemble, simply take the union of the dispersions of each of its matrices. In other words, if $\mathcal{E} = \{P_i \sim \mathcal{D} \mid P_i \in \mathbb{F}^{N \times N}\}_{i=1}^K$, then $\Delta_d(\mathcal{E}, \Pi) = \bigcup_{i=1}^K \Delta_d(P_i, \Pi)$.*

With our spectral statistics defined, we are prepared to discuss prominent results in Random Matrix Theory alongside our new findings from the simulations.