

Chapter 1

Spectra

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.*

While our definition for spectrum is nice and clean, there are a few caveats that we must take care of. First, how are spectra statistics? So we can even call them statistics, so there needs to be some formalization as to why we consider eigenvalues statistics. In this dialogue, we will call back on the same sort of formalization dialogue in the Random Matrices section. Before beginning, the reader is encouraged to review what a statistic is in the review appendix (A.x).

Recall that when we defined and motivated the \mathcal{D} -distribution framework of simulating random matrices, we always had one thing - a vector representation of random variables. Using this framework, the formalization is trivial.

Remark 1.1.1 (Formalization). *Suppose $P \sim \mathcal{D}$ is an $N \times N$ random matrix. Then, P has a representation as a sequence of N^2 random variables, denote it $\vec{X} = \{X_i \mid i = 1, 2, \dots, N^2 - 1, N^2\}$. Then, the spectrum of the matrix P is simply a function of the vector \vec{X} . We can denote this $\sigma(\vec{X})$, where the operator σ is overloaded to mean the spectrum of a matrix **with respect to the vector representation**. The actual process for σ is not necessary to explicitly write out, since characterizing it will be sufficient for now. Essentially, σ is a function that would parse the random variables into the array form by index hacking. Then, it must compute the determinant of the matrix $P - \lambda I$, and solve for its roots. To summarize this, consider the flow chart below.*

To simplify the process, here is how we formalize the spectrum as a statistic. Suppose we sample $P \sim \mathcal{D}$. Then, we take its spectrum formally using σ as such:

$$P_{\square} \rightarrow \vec{P} \rightarrow \sigma(\vec{P}) \rightarrow \text{index magic} \rightarrow \det(P - \lambda I) \rightarrow \sigma(P_{\square})$$

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 was here]

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 1.1.2 (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 $\sigma(\mathcal{E}) = \bigcup_{i=1}^K \sigma(P_i)$.*

[Plot showing the difference between matrix ensemble and matrix spectrum]

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 Ordered Spectra

1.2.1 Eigenvalue Ordering

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 Singular Values

An alternative to using the norm ordering scheme is using the singular values of the matrix. If a matrix is symmetric, the singular values are simply the norm of the eigenvalues. We ignore rotational features and focus solely on scale when we do so.

Suppose P is a random matrix. Then, we can take its singular values as such.

Definition 1.2.1 (Singular Values). *The singular values of a matrix P are given by the square root of the eigenvalues of the corresponding product of that matrix and its transpose. That is, $\sigma_+(P) = \sqrt{\sigma(P \cdot P^T)}$.*

Common Spectrum Schema

Table of Spectrum Schema			
Scheme	Matrix	Notation	Ordering
Sign-Ordered	P	$\sigma_{\pm}(P)$	$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$
Norm-Ordered	P	$\sigma_{ \cdot }(P)$	$ \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N $
Singular	$P \cdot P^T$	$\sigma_{\sqrt{\cdot}}(P)$	$\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_N}$

1.2.3 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.x)

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.

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.

Example 1.2.1 (The Largest Eigenvalue). *Suppose we have seek the largest eigenvalue distribution for a ensemble distribution \mathcal{D} , we would simulate an ensemble \mathcal{E} and observe λ_1 for each of its matrices. Then, we can set the distribution of the largest eigenvalue for \mathcal{D} by observing the distribution of λ_1 .*

So, we fill consider the conditional order statistics $\mathbb{E}(\lambda_i | i)$ and $\text{Var}(\lambda_i | i)$.

1.2.4 Case Study: Tracy-Widom Distribution

Distribution of the normalized largest eigenvalue. Let λ_1 denote the largest eigenvalue. Then, let $\langle \lambda_j - \lambda_i \rangle$ be the mean eigenvalue spacing (of consecutive eigenvalues). The Tracy-Widom distribution is the distribution of $\frac{(\lambda_1 - 2\sqrt{N})n^{1/6}}{\langle \lambda_j - \lambda_i \rangle}$.

[Plot]

1.3 Symmetric and Hermitian Matrices

1.3.1 Introduction

A very important class of matrices in Linear Algebra is that of Symmetric or Hermitian matrices (See A.x). 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.

1.3.2 Case Study: Wigner's Semicircle Distribution

1.4 Findings