Chapter 4

Adjacency matrices, Eigenvalue Interlacing, and the Perron-Frobenius Theorem

In this chapter, we examine the meaning of the smallest and largest eigenvalues of the adjacency matrix of a graph. Note that the largest eigenvalue of the adjacency matrix corresponds to the smallest eigenvalue of the Laplacian. Our focus in this chapter will be on the features that adjacency matrices possess but which Laplacians do not. Where the smallest eigenvector of the Laplacian is a constant vector, the largest eigenvector of an adjacency matrix, called the *Perron vector*, need not be. The Perron-Frobenius theory tells us that the largest eigenvector of an adjacency matrix is non-negative, and that its value is an upper bound on the absolute value of the smallest eigenvalue. These are equal precisely when the graph is bipartite.

We will examine the relation between the largest adjacency eigenvalue and the degrees of vertices in the graph. This is made more meaningful by the fact that we can apply Cauchy's Interlacing Theorem to adjacency matrices. We will use it to prove a theorem of Wilf [Wil67] which says that a graph can be colored using at most $1 + \lfloor \mu_1 \rfloor$ colors. We will learn more about eigenvalues and graph coloring in Chapter 19.

4.1 The Adjacency Matrix

Let $\boldsymbol{M}$ be the adjacency matrix of a (possibly weighted) graph G. As an operator, $\boldsymbol{M}$ acts on a vector $\boldsymbol{x} \in \mathbb{R}^V$ by

$$(\boldsymbol{M}\boldsymbol{x})(a) = \sum_{(a,b) \in E} w_{a,b} \boldsymbol{x}(b). \tag{4.1}$$

We will denote the eigenvalues of $\boldsymbol{M}$ by $\mu_1, \ldots, \mu_n$. But, we order them in the opposite direction than we did for the Laplacian: we assume

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n.$$

The reason for this convention is so that μ_i corresponds to the ith Laplacian eigenvalue, λ_i. If G is a d-regular graph, then $\boldsymbol{D} = \boldsymbol{I}d$,

$$\boldsymbol{L} = \boldsymbol{I}d - \boldsymbol{M},$$

and so

$$\lambda_i = d - \mu_i.$$

Thus the largest adjacency eigenvalue of a d-regular graph is d, and its corresponding eigenvector is the constant vector. We could also prove that the constant vector is an eigenvector of eigenvalue d by considering the action of $\boldsymbol{M}$ as an operator (4.1): if $\boldsymbol{x}(a) = 1$ for all a, then $(\boldsymbol{M}\boldsymbol{x})(b) = d$ for all b.

4.2 The Largest Eigenvalue, μ_1

We now examine μ_1 for graphs which are not necessarily regular. Let G be a graph, let d_{max} be the maximum degree of a vertex in G, and let d_{ave} be the average degree of a vertex in G.

Lemma 4.2.1.

$$d_{ave} \leq \mu_1 \leq d_{max}.$$

Proof. The lower bound follows by considering the Rayleigh quotient with the all-1s vector:

$$\mu_1 = \max_{\boldsymbol{x}} \frac{\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \geq \frac{\boldsymbol{1}^T \boldsymbol{M} \boldsymbol{1}}{\boldsymbol{1}^T \boldsymbol{1}} = \frac{\sum_{a,b} \boldsymbol{M}(a,b)}{n} = \frac{\sum_a \boldsymbol{d}(a)}{n} = d_{ave}.$$

To prove the upper bound, Let $\boldsymbol{\phi}_1$ be an eigenvector of eigenvalue μ_1. Let a be the vertex on which $\boldsymbol{\phi}_1$ takes its maximum value, so $\boldsymbol{\phi}_1(a) \geq \boldsymbol{\phi}_1(b)$ for all b, and we may assume without loss of generality that $\boldsymbol{\phi}_1(a) > 0$ (use $-\boldsymbol{\phi}_1$ if $\boldsymbol{\phi}_1$ is strictly negative). We have

$$\mu_1 = \frac{(\boldsymbol{M}\boldsymbol{\phi}_1)(a)}{\boldsymbol{\phi}_1(a)} = \frac{\sum_{b:b\sim a} \boldsymbol{\phi}_1(b)}{\boldsymbol{\phi}_1(a)} = \sum_{b:b\sim a} \frac{\boldsymbol{\phi}_1(b)}{\boldsymbol{\phi}_1(a)} \leq \sum_{b:b\sim a} 1 = d(a) \leq d_{max}. \tag{4.2}$$

□

Lemma 4.2.2. *If G is connected and $\mu_1 = d_{max}$, then G is d_{max}-regular.*

Proof. If we have equality in (4.2), then it must be the case that $d(a) = d_{max}$ and $\boldsymbol{\phi}_1(b) = \boldsymbol{\phi}_1(a)$ for all $(a,b) \in E$. Thus, we may apply the same argument to every neighbor of a. As the graph is connected, we may keep applying this argument to neighbors of vertices to which it has already been applied to show that $\boldsymbol{\phi}_1(c) = \boldsymbol{\phi}_1(a)$ and $d(c) = d_{max}$ for all $c \in V$. □

The technique used in these last two proofs will appear many times in this Chapter.

4.3 Eigenvalue Interlacing

We can strengthen the lower bound in Lemma 4.2.1 by proving that μ_1 is at least the average degree of every subgraph of G. We will prove this by applying Cauchy's Interlacing Theorem.

For a graph $G = (V, E)$ and $S \subset V$, we define the subgraph induced by S, written $G(S)$, to be the graph with vertex set S and all edges in E connecting vertices in S:

$$\{(a, b) \in E : a \in S \text{ and } b \in S\}.$$

For a symmetric matrix $\boldsymbol{M}$ whose rows and columns are indexed by a set V, and a $S \subset V$, we write $\boldsymbol{M}(S)$ for the symmetric submatrix with rows and columns in S.

Theorem 4.3.1 (Cauchy's Interlacing Theorem)**.** *Let $\boldsymbol{A}$ be an n-by-n symmetric matrix and let $\boldsymbol{B}$ be a principal submatrix of $\boldsymbol{A}$ of dimension $n-1$ (that is, $\boldsymbol{B}$ is obtained by deleting the same row and column from $\boldsymbol{A}$). Then,*

$$\alpha_1 \geq \beta_1 \geq \alpha_2 \geq \beta_2 \geq \cdots \geq \alpha_{n-1} \geq \beta_{n-1} \geq \alpha_n,$$

where $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ and $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_{n-1}$ are the eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively.

Proof. Without loss of generality we will assume that $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by removing its first row and column. We now apply the Courant-Fischer Theorem, which tells us that

$$\alpha_k = \max_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S)=k}} \min_{\boldsymbol{x} \in S} \frac{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

Applying this to $\boldsymbol{B}$ gives

$$\beta_k = \max_{\substack{S \subseteq \mathbb{R}^{n-1} \\ \dim(S)=k}} \min_{\boldsymbol{x} \in S} \frac{\boldsymbol{x}^T \boldsymbol{B} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \max_{\substack{S \subseteq \mathbb{R}^{n-1} \\ \dim(S)=k}} \min_{\boldsymbol{x} \in S} \frac{\begin{pmatrix} 0 \\ \boldsymbol{x} \end{pmatrix}^T \boldsymbol{A} \begin{pmatrix} 0 \\ \boldsymbol{x} \end{pmatrix}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

We see that the right-hand expression is taking a maximum over a special family of subspaces of dimension k: all the vectors in the family must have first coordinate 0. As the maximum over all subspaces of dimension k can only be larger, we immediately have

$$\alpha_k \geq \beta_k.$$

We may prove the inequalities in the other direction, such as $\beta_k \geq \alpha_{k+1}$, by replacing $\boldsymbol{A}$ and $\boldsymbol{B}$ with $-\boldsymbol{A}$ and $-\boldsymbol{B}$. □

Lemma 4.3.2. *For every $S \subseteq V$, let $d_{ave}(S)$ be the average degree of $G(S)$. Then,*

$$d_{ave}(S) \leq \mu_1.$$

Proof. If $\boldsymbol{M}$ is the adjacency matrix of G, then $\boldsymbol{M}(S)$ is the adjacency matrix of $G(S)$. Lemma 4.2.1 says that $d_{ave}(S)$ is at most the largest eigenvalue of the adjacency matrix of $G(S)$, and Theorem 4.3.1 says that this is at most μ_1. □

If we remove the vertex of smallest degree from a graph, the average degree can increase. On the other hand, Cauchy's Interlacing Theorem says that μ_1 can only decrease when we remove a vertex.

Lemma 4.3.2 is a good demonstration of Cauchy's Theorem. But, using Cauchy's Theorem to prove it was overkill. An more direct way to prove it is to emulate the proof of Lemma 4.2.1, but computing the quadratic form in the characteristic vector of S instead of $\mathbf{1}$.

4.4 Wilf's Theorem

We now apply Lemma 4.3.2 to obtain an upper bound on the chromatic number of a graph. Recall that a coloring of a graph is an assignment of colors to vertices in which adjacent vertices have distinct colors. A graph is said to be k-colorable if it can be colored with only k colors[1]. The *chromatic number* of a graph, written $\chi(G)$, is the least k for which G is k-colorable. The bipartite graphs are exactly the graph of chromatic number 2.

It is easy to show that every graph is $(d_{max} + 1)$-colorable. Assign colors to the vertices one-by-one. As each vertex has at most d_{max} neighbors, there is always some color one can assign that vertex that is different than those assigned to its neighbors. The following theorem of Wilf [Wil67] improves upon this bound.

Theorem 4.4.1.

$$\chi(G) \leq \lfloor \mu_1 \rfloor + 1.$$

Proof. We prove this by induction on the number of vertices in the graph. To ground the induction, consider the graph with one vertex and no edges. It has chromatic number 1 and largest eigenvalue zero[2]. Now, assume the theorem is true for all graphs on $n-1$ vertices, and let G be a graph on n vertices. By Lemma 4.2.1, G has a vertex of degree at most $\lfloor \mu_1 \rfloor$. Let a be such a vertex and let $S = V \setminus \{a\}$. By Theorem 4.3.1, the largest eigenvalue of $G(S)$ is at most μ_1, and so our induction hypothesis implies that $G(S)$ has a coloring with at most $\lfloor \mu_1 \rfloor + 1$ colors. Let c be any such coloring. We just need to show that we can extend c to a. As a has at most $\lfloor \mu_1 \rfloor$ neighbors, there is some color in $\{1, \ldots, \lfloor \mu_1 \rfloor + 1\}$ that does not appear among its neighbors, and which it may be assigned. Thus, G has a coloring with $\lfloor \mu_1 \rfloor + 1$ colors. □

The simplest example in which this theorem improves over the naive bound of $d_{max} + 1$ is the path graph on 3 vertices: it has $d_{\max} = 2$ but $\mu_1 < 2$. Thus, Wilf's theorem tells us that it can be colored with 2 colors. Star graphs provide more extreme examples. A star graph with n vertices has $d_{max} = n - 1$ but $\mu_1 = \sqrt{n-1}$.

[1] To be precise, we often identify these k colors with the integers 1 through k. A k-coloring is then a function $c : \{1, \ldots, k\} \to V$ such that $c(a) \neq c(b)$ for all $(a, b) \in E$.

[2] If this makes you uncomfortable, you could use both graphs on two vertices

4.5 Perron-Frobenius Theory for symmetric matrices

The eigenvector corresponding to the largest eigenvalue of the adjacency matrix of a graph is usually not a constant vector. However, it is always a positive vector if the graph is connected. This follows from the Perron-Frobenius theory (discovered independently by Perron [Per07] and Frobenius [Fro12]). In fact, the Perron-Frobenius theory says much more, and it can be applied to adjacency matrices of strongly connected directed graphs. Note that these need not even be diagonalizable!

In the symmetric case, the theory is made much easier by both the spectral theory and the characterization of eigenvalues as extreme values of Rayleigh quotients. For a treatment of the general Perron-Frobenius theory, we recommend Seneta [Sen06] or Bapat and Raghavan [BR97].

Theorem 4.5.1. *[Perron-Frobenius, Symmetric Case] Let G be a connected weighted graph, let $\boldsymbol{M}$ be its adjacency matrix, and let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$ be its eigenvalues. Then*

a. The eigenvalue μ_1 has a strictly positive eigenvector,

b. $\mu_1 \geq -\mu_n$, and

c. $\mu_1 > \mu_2$.

Before proving Theorem 4.5.1, we will prove a lemma that will be used in the proof. It says that non-negative eigenvectors of non-negative adjacency matrices of connected graphs must be strictly positive.

Lemma 4.5.2. *Let G be a connected weighted graph (with non-negative edge weights), let $\boldsymbol{M}$ be its adjacency matrix, and assume that some non-negative vector $\boldsymbol{\phi}$ is an eigenvector of $\boldsymbol{M}$. Then, $\boldsymbol{\phi}$ is strictly positive.*

Proof. If $\boldsymbol{\phi}$ is not strictly positive, there is some vertex a for which $\boldsymbol{\phi}(a) = 0$. As G is connected, there must be some edge (b, c) for which $\boldsymbol{\phi}(b) = 0$ but $\boldsymbol{\phi}(c) > 0$. Let μ be the eigenvalue of $\boldsymbol{\phi}$. As $\boldsymbol{\phi}(b) = 0$, we obtain a contradiction from

$$\mu \boldsymbol{\phi}(b) = (\boldsymbol{M}\boldsymbol{\phi})(b) = \sum_{(b,z) \in E} w_{b,z} \boldsymbol{\phi}(z) \geq w_{b,c} \boldsymbol{\phi}(c) > 0,$$

where the inequalities follow from the fact that the terms $w_{b,z}$ and $\boldsymbol{\phi}(z)$ are non-negative.

So, we conclude that $\boldsymbol{\phi}$ must be strictly positive. □

Proof of Theorem 4.5.1. Let $\boldsymbol{\phi}_1$ be an eigenvector of μ_1 of norm 1, and construct the vector $\boldsymbol{x}$ such that

$$\boldsymbol{x}(u) = |\boldsymbol{\phi}_1(u)|, \text{for all } u.$$

We will show that $\boldsymbol{x}$ is an eigenvector of eigenvalue μ_1.

We have $\boldsymbol{x}^T \boldsymbol{x} = \boldsymbol{\phi}_1^T \boldsymbol{\phi}_1$. Moreover,

$$\mu_1 = \boldsymbol{\phi}_1^T \boldsymbol{M} \boldsymbol{\phi}_1 = \sum_{a,b} \boldsymbol{M}(a,b) \boldsymbol{\phi}_1(a) \boldsymbol{\phi}_1(b) \leq \sum_{a,b} \boldsymbol{M}(a,b) \left|\boldsymbol{\phi}_1(a)\right| \left|\boldsymbol{\phi}_1(b)\right| = \boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x}.$$

So, the Rayleigh quotient of $\boldsymbol{x}$ is at least μ_1. As μ_1 is the maximum possible Rayleigh quotient for a unit vector, the Rayleigh quotient of $\boldsymbol{x}$ must be μ_1 and Theorem 2.2.1 implies that $\boldsymbol{x}$ must be an eigenvector of μ_1. As $\boldsymbol{x}$ is non-negative, Lemma 4.5.2 implies that it is strictly positive.

To prove part b, let $\boldsymbol{\phi}_n$ be the eigenvector of μ_n and let $\boldsymbol{y}$ be the vector for which $\boldsymbol{y}(u) = |\boldsymbol{\phi}_n(u)|$. In the spirit of the previous argument, we can again show that

$$|\mu_n| = |\boldsymbol{\phi}_n \boldsymbol{M} \boldsymbol{\phi}_n| \leq \sum_{a,b} \boldsymbol{M}(a,b) \boldsymbol{y}(a) \boldsymbol{y}(b) \leq \mu_1 \boldsymbol{y}^T \boldsymbol{y} = \mu_1. \tag{4.3}$$

To show that the multiplicity of μ_1 is 1 (that is, $\mu_2 < \mu_1$), consider an eigenvector $\boldsymbol{\phi}_2$. As $\boldsymbol{\phi}_2$ is orthogonal to $\boldsymbol{\phi}_1$, it must contain both positive and negative values. We now construct the vector $\boldsymbol{y}$ such that $\boldsymbol{y}(u) = |\boldsymbol{\phi}_2(u)|$ and repeat the argument that we used for $\boldsymbol{x}$. We find that

$$\mu_2 = \boldsymbol{\phi}_2^T \boldsymbol{M} \boldsymbol{\phi}_2 \leq \boldsymbol{y}^T \boldsymbol{M} \boldsymbol{y} \leq \mu_1.$$

If $\mu_2 = \mu_1$, then $\boldsymbol{y}$ is a nonnegative eigenvector of eigenvalue μ_1, and so Lemma 4.5.2 says that it is strictly positive. Thus, $\boldsymbol{\phi}_2$ does not have any zero entries. As it has both positive and negative entries and the graph is connected, there must be some edge (a,b) for which $\boldsymbol{\phi}_2(a) < 0 < \boldsymbol{\phi}_2(b)$. Then the above inequality must be strict because the edge (a,b) will make a negative contribution to $\boldsymbol{\phi}_2^T \boldsymbol{M} \boldsymbol{\phi}_2$ and a positive contribution to $\boldsymbol{y}^T \boldsymbol{M} \boldsymbol{y}$. This contradicts our assumption that $\mu_2 = \mu_1$. □

Finally, we show that for a connected graph G, $\mu_n = -\mu_1$ if and only if G is bipartite. In fact, if $\mu_n = -\mu_1$, then $\mu_{n-i} = -\mu_{i+1}$ for every i.

Proposition 4.5.3. *If G is a connected graph and $\mu_n = -\mu_1$, then G is bipartite.*

Proof. Consider the conditions necessary to achieve equality in (4.3). First, $\boldsymbol{y}$ must be an eigenvector of eigenvalue μ_1. Thus, $\boldsymbol{y}$ must be strictly positive, $\boldsymbol{\phi}_n$ can not have any zero values, and there must be an edge (a,b) for which $\boldsymbol{\phi}_n(a) < 0 < \boldsymbol{\phi}_n(b)$. It must also be the case that all of the terms in

$$\sum_{(a,b) \in E} \boldsymbol{M}(a,b) \boldsymbol{\phi}_n(a) \boldsymbol{\phi}_n(b)$$

have the same sign, and we have established that this sign must be negative. Thus, for every edge (a,b), $\boldsymbol{\phi}_n(a)$ and $\boldsymbol{\phi}_n(b)$ must have different signs. That is, the signs provide the bipartition of the vertices. □

Proposition 4.5.4. *If G is bipartite then the eigenvalues of its adjacency matrix are symmetric about zero.*

Proof. As G is bipartite, we may divide its vertices into sets S and T so that all edges go between S and T. Let $\boldsymbol{\phi}$ be an eigenvector of $\boldsymbol{M}$ with eigenvalue μ. Define the vector $\boldsymbol{x}$ by

$$\boldsymbol{x}(a) = \begin{cases} \boldsymbol{\phi}(a) & \text{if } a \in S, and \\ -\boldsymbol{\phi}(a) & \text{if } a \in T. \end{cases}$$

To see that $\boldsymbol{x}$ is an eigenvector with eigenvalue $-\mu$, note that for $a \in S$,

$$(\boldsymbol{M}\boldsymbol{x})(a) = \sum_{(a,b) \in E} \boldsymbol{M}(a,b)\boldsymbol{x}(b) = \sum_{(a,b) \in E} \boldsymbol{M}(a,b)(-\boldsymbol{\phi}(b)) = -\mu\boldsymbol{\phi}(a) = -\mu\boldsymbol{x}(a).$$

We may similarly show that $(\boldsymbol{M}\boldsymbol{x})(a) = -\mu\boldsymbol{x}(a)$ for $a \in T$. $\square$