Spectral Graph Theory Lecture 3

The Adjacency Matrix and Graph Coloring

Daniel A. Spielman September 9, 2015

Disclaimer

These notes are not necessarily an accurate representation of what happened in class. The notes written before class say what I think I should say. I sometimes edit the notes after class to make them way what I wish I had said.

There may be small mistakes, so I recommend that you check any mathematically precise statement before using it in your own work.

These notes were last revised on September 13, 2015.

3.1 Overview

In this lecture, I will discuss the adjacency matrix of a graph, and the meaning of its largest and smallest eigenvalues. Note that the largest eigenvalue of the adjacency matrix corresponds to the smallest eigenvalue of the Laplacian.

I introduce the Perron-Frobenius theory, which basically says that the largest eigenvalue of the adjacency matrix of a connected graph has multiplicity 1 and that its corresponding eigenvector is uniform in sign.

I will then present bounds on the number of colors needed to color a graph in terms of its extreme adjacency matrix eigenvalues.

The body of the notes includes the material that I intend to cover in class. Proofs that I will skip, but which you should know, appear in the Appendix and Exercises.

3.2 The Adjacency Matrix

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

$$(\boldsymbol{A}\boldsymbol{x})(u) = \sum_{(u,v)\in E} w_{u,v}\boldsymbol{x}(v). \tag{3.1}$$

We will denote the eigenvalues of $\boldsymbol{A}$ 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 $\boldsymbol{D} = \boldsymbol{I}d$, and

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

and so

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

So, we see that 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 $\boldsymbol{A}$ as an operator (3.1): if $\boldsymbol{x}(u) = 1$ for all u, then $(\boldsymbol{Ax})(v) = d$ for all v.

3.3 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. If the graph is weighted, then these are the weighted degrees. In the following, we let $d(u)$ denote the (weighted) degree of vertex u.

Lemma 3.3.1.

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

Proof. While this theorem holds in the weighted case, we just prove it in the unweighted case for simplicity.

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

$$\mu_1 = \max_{\boldsymbol{x}} \frac{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \geq \frac{\boldsymbol{1}^T \boldsymbol{A} \boldsymbol{1}}{\boldsymbol{1}^T \boldsymbol{1}} = \frac{\sum_{(u,v) \in E} \boldsymbol{A}(u,v)}{n} = \frac{\sum_u d(u)}{n}.$$

To prove the upper bound, Let $\boldsymbol{\phi}_1$ be an eigenvector of eigenvalue μ_1. Let v be the vertex on which it takes its maximum value, so $\boldsymbol{\phi}_1(v) \geq \boldsymbol{\phi}_1(u)$ for all u, and assume without loss of generality that $\boldsymbol{\phi}_1(v) \neq 0$. We have

$$\mu_1 = \frac{(A\boldsymbol{\phi}_1)(v)}{\boldsymbol{\phi}_1(v)} = \frac{\sum_{u:(u,v) \in E} \boldsymbol{\phi}_1(u)}{\boldsymbol{\phi}_1(v)} = \sum_{u:(u,v) \in E} \frac{\boldsymbol{\phi}_1(u)}{\boldsymbol{\phi}_1(v)} \leq \sum_{u:(u,v) \in E} 1 \leq d(v) \leq d_{max}. \tag{3.2}$$

□

We can strengthen the lower bound by proving that μ_1 is at least the averge degree of every subgraph of G.

Lemma 3.3.2. *For every $S \subseteq V$, let $d_{ave}(S)$ be the average degree of vertices in the subgraph induced on the vertices in S (that is, having only edges between vertices of S). Then,*

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

An easy way to prove this is to emulate the proof of Lemma 3.3.1, but computing the quadratic form in the characteristic vector of S instead of $\mathbf{1}$. I will instead take this opportunity to state one of the most important properties of the eigenvalues of symmetric matrices, and prove it using that.

For a matrix $\boldsymbol{A}$ with rows and columns indexed by $\{1, \ldots, n\}$ and for $S \subseteq \{1, \ldots, n\}$, let $\boldsymbol{A}(S)$ denote the sub-matrix of $\boldsymbol{A}$ with rows and columns indexed by S. Throughout this lecture, we will let $\lambda_{max}(\boldsymbol{A})$ denote the largest eigenvalue of the matrix $\boldsymbol{A}$, and we will let $\lambda_{\min}(\boldsymbol{A})$ denote the smallest, even if we are calling the eigenvalues μ_n and μ_1.

Lemma 3.3.3. *Let $\boldsymbol{A}$ be a symmetric matrix and let S be a subset of its row and column indices. Then*

$$\lambda_{max}(\boldsymbol{A}) \geq \lambda_{max}(\boldsymbol{A}(S)) \geq \lambda_{min}(\boldsymbol{A}(S)) \geq \lambda_{min}(\boldsymbol{A}).$$

Proof. It suffices to the lemma in the case that $S = \{1, \ldots, n-1\}$. So, let $S = \{1, \ldots, n-1\}$ and let $\boldsymbol{B} = \boldsymbol{A}(S)$.

For any vector $\boldsymbol{y} \in \mathbb{R}^{n-1}$, we have

$$\boldsymbol{y}^T \boldsymbol{B} \boldsymbol{y} = \begin{pmatrix} \boldsymbol{y} \\ 0 \end{pmatrix}^T \boldsymbol{A} \begin{pmatrix} \boldsymbol{y} \\ 0 \end{pmatrix}.$$

So, for $\boldsymbol{y}$ an eigenvector of $\boldsymbol{B}$ of eigenvalue $\lambda_{max}(\boldsymbol{B})$,

$$\lambda_{max}(\boldsymbol{B}) = \frac{\boldsymbol{y}^T \boldsymbol{B} \boldsymbol{y}}{\boldsymbol{y}^T \boldsymbol{y}} = \frac{\begin{pmatrix} \boldsymbol{y} \\ 0 \end{pmatrix}^T \boldsymbol{A} \begin{pmatrix} \boldsymbol{y} \\ 0 \end{pmatrix}}{\begin{pmatrix} \boldsymbol{y} \\ 0 \end{pmatrix}^T \begin{pmatrix} \boldsymbol{y} \\ 0 \end{pmatrix}} \leq \max_{\boldsymbol{x} \in \mathbb{R}^n} \frac{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \lambda_{max}(\boldsymbol{A}).$$

The same argument works for the smallest eigenvalues. □

Proof of Lemma 3.3.2. Let $S \subseteq V$, and let $G(S)$ denote the subgraph induced on the vertices in S. If $\boldsymbol{A}$ is the adjacency matrix of G, then $\boldsymbol{A}(S)$ is the adjacency matrix of $G(S)$. Lemma 3.3.1 says that $d_{ave}(S)$ is at most the largest eigenvalue of the adjacency matrix of $G(S)$, and Lemma 3.3.3 says that this is at most μ_1. □

Lemma 3.3.4. *If G is connected and $\mu_1 = d_{max}$, then G is d_{max}-regular.*

Proof. If we have equality in (3.2), then it must be the case that $d(v) = d_{max}$ and $\boldsymbol{\phi}_1(u) = \boldsymbol{\phi}_1(v)$ for all $(u, v) \in E$. Thus, we may apply the same argument to every neighbor of v. 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 $\boldsymbol{\phi}_1(z) = \boldsymbol{\phi}_1(v)$ and $d(z) = d_{max}$ for all $z \in V$. □

3.4 The Corresponding Eigenvector

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. 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! If you'd like to learn about the general theory, look at the third lecture from my notes from 2009.

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.

Theorem 3.4.1. *[Perron-Frobenius, Symmetric Case] Let G be a connected weighted graph, let $\boldsymbol{A}$ be its adjacency matrix, and let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$ be its eigenvalues. Then*

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

b. $\mu_1 > \mu_2$,

c. The eigenvalue μ_1 has a strictly positive eigenvector.

Before proving Theorem 3.4.1, we will state a lemma that will be useful in the proof and a few other places today. It says that non-negative eigenvectors of non-negative adjacency matrices of connected graphs must be strictly positive.

Lemma 3.4.2. *Let G be a connected weighted graph (with non-negative edge weights), let $\boldsymbol{A}$ be its adjacency matrix, and assume that some non-negative vector $\boldsymbol{\phi}$ is an eigenvector of $\boldsymbol{A}$. Then, $\boldsymbol{\phi}$ is strictly positive.*

You should solve this as an exercise. The proof is similar to that of Lemma 3.3.4.

Proof of Theorem 3.4.1. Let $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_n$ be the eigenvectors corresponding to $\mu_1, \ldots, \mu_n$.

We will just prove part c for now. Recall that

$$\mu_1 = \max_{\boldsymbol{x}} \frac{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

Let $\boldsymbol{\phi}_1$ be an eigenvector of μ_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,

$$\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} = \sum_{u,v} \boldsymbol{A}(u,v) |\boldsymbol{\phi}_1(u)| |\boldsymbol{\phi}_1(v)| \geq \sum_{u,v} \boldsymbol{A}(u,v) \boldsymbol{\phi}_1(u) \boldsymbol{\phi}_1(v) = \boldsymbol{\phi}_1^T \boldsymbol{A} \boldsymbol{\phi}_1 = \mu_1.$$

So, the Rayleigh quotient of $\boldsymbol{x}$ is at least μ_1. As μ_1 is the maximum possible Rayleigh quotient, the Rayleigh quotient of $\boldsymbol{x}$ must be μ_1 and $\boldsymbol{x}$ must be an eigenvector of μ_1.

So, we now know that $\boldsymbol{A}$ has an eigenvector $\boldsymbol{x}$ that is non-negative. We can then apply Lemma 3.4.2 to show that $\boldsymbol{x}$ is strictly positive.

The rest of the proof appears in the Appendix.

□

The following characterization of bipartite graphs follows from similar ideas.

Proposition 3.4.3. *If G is a connected graph, then $\mu_n = -\mu_1$ if and only if G is bipartite.*

The proof appears in the Appendix.

The nth eigenvalue, which is the most negative in the case of the adjacency matrix and is the largest in the case of the Laplacian, corresponds to the highest frequency vibration in a graph. Its corresponding eigenvector tries to assign as different as possible values to neighboring vertices. This is, it tries to assign a coloring. In fact, there are heuristics for finding k colorings by using the $k-1$ largest eigenvectors [AK97].

3.5 Graph Coloring

A coloring of a graph is an assignment of one color to every vertex in a graph so that each edge attaches vertices of different colors. We are interested in coloring graphs while using as few colors as possible. Formally, a k-coloring of a graph is a function $c : V \to \{1, \dots, k\}$ so that for all $(u, v) \in V$, $c(u) \neq c(v)$. A graph is k-colorable if it has a k-coloring. The chromatic number of a graph, written χ_G, is the least k for which G is k-colorable. A graph G is 2-colorable if and only if it is bipartite. Determining whether or not a graph is 3-colorable is an NP-complete problem. The famous 4-Color Theorem [AH77a, AH77b] says that every planar graph is 4-colorable.

3.6 Wilf's Theorem

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 from those assigned to its neighbors. Wilf's theorem will exploit a generalization of this algorithm.

Order the vertices of a graph 1 through n. Let κ be a number so that every vertex has at most κ edges to vertices of lower number. That is,

$$\forall u, |\{v : v < u \text{ and } (u, v) \in E\}| \leq \kappa.$$

Then, the same algorithm will color G with at most $\kappa + 1$ colors. It goes through the vertices in order, and for each vertex assigns it a color that is not assigned to any of its neighbors. As each vertex has at most κ neighbors that were already assigned colors, $\kappa + 1$ colors will suffice.

The following theorem of Wilf [Wil67] uses Lemma 3.3.2 to prove that every graph can be arranged this way with $\kappa \leq \mu_1$.

Theorem 3.6.1.

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

Proof. We prove that one can order the vertices of G so that each vertex has at most μ_1 edges to vertices that come before it in the order. As the number of edges must be an integer, this gives an ordering with $\kappa \leq \lfloor \mu_1 \rfloor$.

We begin by choosing the last vertex in the order. Lemma 3.3.2 tells us that the average degree of G is at most μ_1. This implies that there is some vertex of degree at most μ_1. Call this vertex n, and put it last in the order.

It now remains to order the vertices in the subgraph on the other vertices. Lemma 3.3.3 tells us that the largest eigenvalue of the adjacency matrix of this subgraph is at most μ_1. So, by induction on the number of vertices in the graph, we may assume that this subgraph has an ordering in which every vertex has at most μ_1 neighbors that come before it. Use this order, and put vertex n at the end. □

For an example, consider a path graph with at least 3 vertices. We have $d_{max} = 2$, but $\mu_1 < 2$. So, this theorem tells us that we can color it with two colors. But, you could probably 2-color a path without thinking.

3.7 Hoffman's Bound

Hoffman [Hof70] proved a bound on the chromatic number of a graph in terms of its adjacency matrix eigenvalues that is tight for bipartite graphs. It is also remarkable in that it applies equally well to weighted graphs. Of course, the chromatic number of a graph does not depend on the weights of its edges. But, the adjacency matrix, and therefor its eigenvalues, does.

Theorem 3.7.1.

$$\chi(G) \geq \frac{\mu_1 - \mu_n}{-\mu_n} = 1 + \frac{\mu_1}{-\mu_n}.$$

This theorem follows from a partial converse to Lemma 3.3.3.

Lemma 3.7.2. *Let*

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{1,1} & \boldsymbol{A}_{1,2} & \cdots & \boldsymbol{A}_{1,k} \\ \boldsymbol{A}_{1,2}^T & \boldsymbol{A}_{2,2} & \cdots & \boldsymbol{A}_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{1,k}^T & \boldsymbol{A}_{2,k}^T & \cdots & \boldsymbol{A}_{k,k} \end{bmatrix}$$

be a block-partitioned symmetric matrix with $k \geq 2$. *Then*

$$(k-1)\lambda_{min}(\boldsymbol{A}) + \lambda_{max}(\boldsymbol{A}) \leq \sum_i \lambda_{max}(\boldsymbol{A}_{i,i}).$$

Proof of Theorem 3.7.1. Let G be a k-colorable graph. After possibly re-ordering the vertices, the adjacency matrix of G can be written

$$\begin{bmatrix} \mathbf{0} & \boldsymbol{A}_{1,2} & \cdots & \boldsymbol{A}_{1,k} \\ \boldsymbol{A}_{1,2}^T & \mathbf{0} & \cdots & \boldsymbol{A}_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{1,k}^T & \boldsymbol{A}_{2,k}^T & \cdots & \mathbf{0} \end{bmatrix}.$$

Each block corresponds to a color.

As each diagonal block is all-zero, Lemma 3.7.2 implies

$$(k-1)\lambda_{min}(\boldsymbol{A}) + \lambda_{max}(\boldsymbol{A}) \leq 0.$$

Recalling that $\lambda_{min}(\boldsymbol{A}) = \mu_n < 0$, and $\lambda_{max}(\boldsymbol{A}) = \mu_1$, a little algebra yields

$$1 + \frac{\mu_1}{-\mu_n} \leq k.$$

□

To prove Lemma 3.7.2, we begin with the case of $k = 2$. The general case follows from this one by induction.

Lemma 3.7.3. *Let*

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{C}^T & \boldsymbol{D} \end{bmatrix}$$

be a symmetric matrix. Then

$$\lambda_{min}(\boldsymbol{A}) + \lambda_{max}(\boldsymbol{A}) \leq \lambda_{max}(\boldsymbol{B}) + \lambda_{max}(\boldsymbol{D}).$$

Proof. Let $\boldsymbol{x}$ be an eigenvector of $\boldsymbol{A}$ of eigenvalue $\lambda_{max}(\boldsymbol{A})$. To simplify formulae, let's also assume that $\boldsymbol{x}$ is a unit vector. Write $\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}$, using the same partition as we did for $\boldsymbol{A}$.

We first consider the case in which neither $\boldsymbol{x}_1$ nor $\boldsymbol{x}_2$ is an all-zero vector. In this case, we set

$$\boldsymbol{y} = \begin{pmatrix} \frac{\|\boldsymbol{x}_2\|}{\|\boldsymbol{x}_1\|}\boldsymbol{x}_1 \\ -\frac{\|\boldsymbol{x}_1\|}{\|\boldsymbol{x}_2\|}\boldsymbol{x}_2 \end{pmatrix}.$$

The reader may verify that $\boldsymbol{y}$ is also a unit vector, so

$$\boldsymbol{y}^T A \boldsymbol{y} \geq \lambda_{min}(\boldsymbol{A}).$$

We have

$$\begin{aligned}
\lambda_{max}(\boldsymbol{A}) + \lambda_{min}(\boldsymbol{A}) &\leq \boldsymbol{x}^T A \boldsymbol{x} + \boldsymbol{y}^T A \boldsymbol{y} \\
&= \boldsymbol{x}_1^T \boldsymbol{B} \boldsymbol{x}_1 + \boldsymbol{x}_1^T \boldsymbol{C} \boldsymbol{x}_2 + \boldsymbol{x}_2^T \boldsymbol{C}^T \boldsymbol{x}_1 + \boldsymbol{x}_2^T \boldsymbol{D} \boldsymbol{x}_2 + \\
&\quad + \frac{\|\boldsymbol{x}_2\|^2}{\|\boldsymbol{x}_1\|^2} \boldsymbol{x}_1^T \boldsymbol{B} \boldsymbol{x}_1 - \boldsymbol{x}_1^T \boldsymbol{C} \boldsymbol{x}_2 - \boldsymbol{x}_2^T \boldsymbol{C}^T \boldsymbol{x}_1 + \frac{\|\boldsymbol{x}_1\|^2}{\|\boldsymbol{x}_2\|^2} \boldsymbol{x}_2^T \boldsymbol{D} \boldsymbol{x}_2 \\
&= \boldsymbol{x}_1^T \boldsymbol{B} \boldsymbol{x}_1 + \boldsymbol{x}_2^T \boldsymbol{D} \boldsymbol{x}_2 + \frac{\|\boldsymbol{x}_2\|^2}{\|\boldsymbol{x}_1\|^2} \boldsymbol{x}_1^T \boldsymbol{B} \boldsymbol{x}_1 + \frac{\|\boldsymbol{x}_1\|^2}{\|\boldsymbol{x}_2\|^2} \boldsymbol{x}_2^T \boldsymbol{D} \boldsymbol{x}_2 \\
&\leq \left(1 + \frac{\|\boldsymbol{x}_2\|^2}{\|\boldsymbol{x}_1\|^2}\right) \boldsymbol{x}_1^T \boldsymbol{B} \boldsymbol{x}_1 + \left(1 + \frac{\|\boldsymbol{x}_1\|^2}{\|\boldsymbol{x}_2\|^2}\right) \boldsymbol{x}_2^T \boldsymbol{D} \boldsymbol{x}_2 \\
&\leq \lambda_{max}(\boldsymbol{B}) \left(\|\boldsymbol{x}_1\|^2 + \|\boldsymbol{x}_2\|^2\right) + \lambda_{max}(\boldsymbol{D}) \left(\|\boldsymbol{x}_1\|^2 + \|\boldsymbol{x}_2\|^2\right) \\
&= \lambda_{max}(\boldsymbol{B}) + \lambda_{max}(\boldsymbol{D}),
\end{aligned}$$

as $\boldsymbol{x}$ is a unit vector.

We now return to the case in which $\|\boldsymbol{x}_2\| = 0$ (or $\|\boldsymbol{x}_1\| = 0$, which is really the same case). Lemma 3.3.3 tells us that $\lambda_{max}(\boldsymbol{B}) \leq \lambda_{max}(\boldsymbol{A})$. So, it must be the case that $\boldsymbol{x}_1$ is an eigenvector of eigenvalue $\lambda_{max}(\boldsymbol{A})$ of $\boldsymbol{B}$, and thus $\lambda_{max}(\boldsymbol{B}) = \lambda_{max}(\boldsymbol{A})$. To finish the proof, also observe that Lemma 3.3.3 implies

$$\lambda_{max}(\boldsymbol{D}) \geq \lambda_{min}(\boldsymbol{D}) \geq \lambda_{min}(\boldsymbol{A}).$$

□

Proof of Lemma 3.7.2. For $k = 2$, this is exactly Lemma 3.7.3. For $k > 2$, we apply induction. Let

$$B = \begin{bmatrix} \boldsymbol{A}_{1,1} & \boldsymbol{A}_{1,2} & \cdots & \boldsymbol{A}_{1,k-1} \\ \boldsymbol{A}_{1,2}^T & \boldsymbol{A}_{2,2} & \cdots & \boldsymbol{A}_{2,k-1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{1,k-1}^T & \boldsymbol{A}_{2,k-1}^T & \cdots & \boldsymbol{A}_{k-1,k-1} \end{bmatrix}.$$

Lemma 3.3.3 now implies.

$$\lambda_{min}(\boldsymbol{B}) \geq \lambda_{min}(\boldsymbol{A}).$$

Applying Lemma 3.7.3 to B and the kth row and column of $\boldsymbol{A}$, we find

$$\begin{aligned}
\lambda_{min}(\boldsymbol{A}) + \lambda_{max}(\boldsymbol{A}) &\leq \lambda_{max}(\boldsymbol{B}) + \lambda_{max}(\boldsymbol{A}_{k,k}) \\
&\leq -(k-2)\lambda_{min}(\boldsymbol{B}) + \sum_{i=1}^{k-1} \lambda_{max}(\boldsymbol{A}_{i,i}) + \lambda_{max}(\boldsymbol{A}_{k,k}) \qquad \text{(by induction)} \\
&\leq -(k-1)\lambda_{min}(\boldsymbol{A}) + \sum_{i=1}^{k} \lambda_{max}(\boldsymbol{A}_{i,i}),
\end{aligned}$$

which proves the lemma. □

A Some Proofs

The rest of the proof of Theorem 3.4.1 . 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

$$|\mu_n| = |\boldsymbol{\phi}_n \boldsymbol{A} \boldsymbol{\phi}_n| \leq \sum_{u,v} \boldsymbol{A}(u,v)\boldsymbol{y}(u)\boldsymbol{y}(v) \leq \mu_1 \boldsymbol{y}^T \boldsymbol{y} = \mu_1.$$

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

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

From here, we divide the proof into two cases. First, consider the case in which $\boldsymbol{y}$ is never zero. In this case, there must be some edge (u,v) for which $\boldsymbol{\phi}_2(u) < 0 < \boldsymbol{\phi}_2(v)$. Then the above inequality must be strict because the edge (u,v) will make a negative contribution to $\boldsymbol{\phi}_2^T \boldsymbol{A} \boldsymbol{\phi}_2$ and a positive contribution to $\boldsymbol{y}^T \boldsymbol{A} \boldsymbol{y}$.

We will argue by contradiction in the case that $\boldsymbol{y}$ has a zero value. In this case, if $\mu_2 = \mu_1$ then $\boldsymbol{y}$ will be an eigenvector of eigenvalue μ_1. This is a contradiction, as Lemma 3.4.2 says that a non-negative eigenvector cannot have a zero value. So, if $\boldsymbol{y}$ has a zero value then $\boldsymbol{y}^T \boldsymbol{A} \boldsymbol{y} < \mu_1$ and $\mu_2 < \mu_1$ as well.

□

Proof of Proposition 3.4.3. First, assume that G is bipartite. That is, we have a decomposition of V into sets U and W such that all edges go between U and W. Let $\boldsymbol{\phi}_1$ be the eigenvector of μ_1. Define

$$\boldsymbol{x}(u) = \begin{cases} \boldsymbol{\phi}_1(u) & \text{if } u \in U, \textit{and} \\ -\boldsymbol{\phi}_1(u) & \text{if } u \in W. \end{cases}$$

For $u \in U$, we have

$$(\boldsymbol{A}\boldsymbol{x})(u) = \sum_{(u,v) \in E} \boldsymbol{x}(v) = -\sum_{(u,v) \in E} \boldsymbol{\phi}(v) = -\mu_1 \boldsymbol{\phi}(u) = -\mu_1 \boldsymbol{x}(u).$$

Using a similar argument for $u \notin U$, we can show that $\boldsymbol{x}$ is an eigenvector of eigenvalue $-\mu_1$.

To go the other direction, assume that $\mu_n = -\mu_1$. We then construct $\boldsymbol{y}$ as in the previous proof, and again observe

$$|\mu_n| = |\boldsymbol{\phi}_n \boldsymbol{A} \boldsymbol{\phi}_n| = \left| \sum_{u,v} \boldsymbol{A}(u,v)\boldsymbol{\phi}_n(u)\boldsymbol{\phi}_n(v) \right| \leq \sum_{u,v} \boldsymbol{A}(u,v)\boldsymbol{y}(u)\boldsymbol{y}(v) \leq \mu_1 \boldsymbol{y}^T \boldsymbol{y} = \mu_1.$$

For this to be an equality, it must be the case that $\boldsymbol{y}$ is an eigenvalue of μ_1, and so $\boldsymbol{y} =$ ‘$\boldsymbol{\phi}_1$. For the first inequality above to be an equality, it must also be the case that all the terms $\boldsymbol{\phi}_n(u)\boldsymbol{\phi}_n(v)$

have the same sign. In this case that sign must be negative. So, we every edge goes between a vertex for which $\boldsymbol{\phi}_n(u)$ is positive and a vertex for which $\boldsymbol{\phi}_n(v)$ is negative. Thus, the signs of $\boldsymbol{\phi}_n$ give the bi-partition. □

B Exercises

1. Prove Lemma 3.4.2.

Note that the proof is similar to that of Lemma 3.3.4.

References

[AH77a] Kenneth Appel and Wolfgang Haken. Every planar map is four colorable part i. discharging. *Illinois Journal of Mathematics*, 21:429–490, 1977.

[AH77b] Kenneth Appel and Wolfgang Haken. Every planar map is four colorable part ii. reducibility. *Illinois Journal of Mathematics*, 21:491–567, 1977.

[AK97] Alon and Kahale. A spectral technique for coloring random 3-colorable graphs. *SICOMP: SIAM Journal on Computing*, 26, 1997.

[Hof70] A. J. Hoffman. On eigenvalues and colorings of graphs. In *Graph Theory and its Applications*, pages 79–92. Academic Press, New York, 1970.

[Wil67] Herbert S. Wilf. The eigenvalues of a graph and its chromatic number. *J. London math. Soc.*, 42:330–332, 1967.