

The Adjacency Matrix and The n th Eigenvalue*Daniel A. Spielman*

September 5, 2012

3.1 About these notes

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. The notes written after class say what I wish I said.

Be skeptical of all statements in these notes that can be made mathematically rigorous.

3.2 Overview

In this lecture, I will discuss the adjacency matrix of a graph, and the meaning of its smallest eigenvalue. This corresponds to the largest eigenvalue of the Laplacian, which we will examine as well. We will relate these to bounds on the chromatic numbers of graphs and the sizes of independent sets of vertices in graphs. In particular, we will prove Hoffman's bound, and some generalizations.

Warning: I am going to give an alternative approach to Hoffman's bound on the chromatic number of a graph in which I use the Laplacian instead of the adjacency matrix. I just worked this out last night, so I still don't know if it is a good idea or not. But, I'm going to go with it.

My proof of Hoffman's bound in the regular case will be much simpler than the proof that I gave in 2009.

3.3 The Adjacency Matrix

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

$$(\mathbf{A}\mathbf{x})(u) = \sum_{(u,v) \in E} w(u,v)\mathbf{x}(v). \quad (3.1)$$

We will denote the eigenvalues of \mathbf{A} by μ_1, \dots, μ_n . But, we order them in the opposite direction than we did for the Laplacian: we assume

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_n.$$

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

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

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 \mathbf{A} as an operator (3.1): if $\mathbf{x}(u) = 1$ for all u , then $(\mathbf{A}\mathbf{x})(v) = d$ for all v .

3.4 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 3.4.1.

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

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

$$\mu_1 = \max_{\mathbf{x}} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \geq \frac{\mathbf{1}^T \mathbf{A} \mathbf{1}}{\mathbf{1}^T \mathbf{1}} = \frac{\sum_{i,j} \mathbf{A}(i,j)}{n} = \frac{\sum_i \mathbf{d}(i)}{n}.$$

To prove the upper bound, Let ϕ_1 be an eigenvector of eigenvalue μ_1 . Let v be the vertex on which it takes its maximum value, so $\phi_1(v) \geq \phi_1(u)$ for all u , and assume without loss of generality that $\phi_1(v) \neq 0$. We have

$$\mu_1 = \frac{(\mathbf{A}\phi_1)(v)}{\phi_1(v)} = \frac{\sum_{u \sim v} \phi_1(u)}{\phi_1(v)} = \sum_{u \sim v} \frac{\phi_1(u)}{\phi_1(v)} \leq \sum_{u \sim v} 1 \leq d(v) \leq d_{max}. \quad (3.2)$$

□

Lemma 3.4.2. *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 $\phi_1(u) = \phi_1(v)$ 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 $\phi_1(z) = \phi_1(v)$ and $d(z) = d_{max}$ for all $z \in V$. □

3.5 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! We will defer a discussion of the general theory until we discuss directed graphs, which will happen towards the end of the semester. If you want to see it now, 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 3.5.1. *[Perron-Frobenius, Symmetric Case] Let G be a connected weighted graph, let \mathbf{A} be its adjacency matrix, and let $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$ 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 3.5.1, we will prove 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 3.5.2. *Let G be a connected weighted graph (with non-negative edge weights), let \mathbf{A} be its adjacency matrix, and assume that some non-negative vector ϕ is an eigenvector of \mathbf{A} . Then, ϕ is strictly positive.*

Proof. Assume by way of contradiction that ϕ is not strictly positive. So, there is some vertex u for which $\phi(u) = 0$. Thus, there must be some edge (u, v) for which $\phi(u) = 0$ but $\phi(v) > 0$. We would then

$$(\mathbf{A}\phi)(u) = \sum_{(u,z) \in E} w(u,z)\phi(z) \geq w(u,v)\phi(v) > 0,$$

as all the terms $w(u, z)$ and $\phi(z)$ are non-negative. But, this must also equal $\mu\phi(u) = 0$, where μ is the eigenvalue corresponding to ϕ . This is a contradiction.

So, we conclude that ϕ must be strictly positive. □

Proof of Theorem 3.5.1. Let ϕ_1, \dots, ϕ_n be the eigenvectors corresponding to μ_1, \dots, μ_n .

We start with part c. Recall that

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

Let ϕ_1 be an eigenvector of μ_1 , and construct the vector \mathbf{x} such that

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

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

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

$$\phi_1^T \mathbf{A} \phi_1 = \sum_{u,v} \mathbf{A}(u,v) \phi(u) \phi(v) \leq \sum_{u,v} \mathbf{A}(u,v) |\phi(u)| |\phi(v)| = \mathbf{x}^T \mathbf{A} \mathbf{x}.$$

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

So, we now know that \mathbf{A} has an eigenvector \mathbf{x} that is non-negative. We can then apply Lemma 3.5.2 to show that \mathbf{x} is strictly positive.

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

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

To show that the multiplicity of μ_1 is 1 (that is, $\mu_2 < \mu_1$), consider an eigenvector ϕ_2 . As ϕ_2 is orthogonal to ϕ_1 , it must contain both positive and negative values. We now construct the vector \mathbf{y} such that $\mathbf{y}(u) = |\phi_2(u)|$ and repeat the argument that we used for \mathbf{x} . We find that

$$\mu_2 = \frac{\phi_2^T \mathbf{A} \phi_2}{\phi_2 \phi_2} \leq \frac{\mathbf{y}^T \mathbf{A} \mathbf{y}}{\mathbf{y}^T \mathbf{y}} \leq \mu_1.$$

From here, we divide the proof into two cases. First, consider the case in which \mathbf{y} is never zero. In this case, there must be some edge (u, v) for which $\phi_2(u) < 0 < \phi_2(v)$. Then the above inequality must be strict because the edge (u, v) will make a negative contribution to $\phi_2^T \mathbf{A} \phi_2$ and a positive contribution to $\mathbf{y}^T \mathbf{A} \mathbf{y}$.

We will argue by contradiction in the case that \mathbf{y} has a zero value. In this case, if $\mu_2 = \mu_1$ then \mathbf{y} will be an eigenvector of eigenvalue μ_1 . This is a contradiction, as Lemma 3.5.2 says that a non-negative eigenvector cannot have a zero value. So, if \mathbf{y} has a zero value then $\mathbf{y}^T \mathbf{A} \mathbf{y} < \mu_1$ and $\mu_2 < \mu_1$ as well. \square

The following characterization of bipartite graphs follows from similar ideas.

Proposition 3.5.3. *If G is a connected graph, then $\mu_n = -\mu_1$ if and only if G is bipartite.*

Proof. 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 ϕ_1 be the eigenvector of μ_1 . Define

$$\mathbf{x}(u) = \begin{cases} \phi_1(u) & \text{if } u \in U, \text{ and} \\ -\phi_1(u) & \text{if } u \in W. \end{cases}$$

For $u \in U$, we have

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

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

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

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

For this to be an equality, it must be the case that \mathbf{y} is an eigenvalue of μ_1 , and so $\mathbf{y} = \phi_1$. For the first inequality above to be an equality, it must also be the case that all the terms $\phi_n(u)\phi_n(v)$ have the same sign. In this case that sign must be negative. So, every edge goes between a vertex for which $\phi_n(u)$ is positive and a vertex for which $\phi_n(v)$ is negative. Thus, the signs of ϕ_n give the bi-partition. \square

The n th 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].

3.6 Graph Coloring and Independent Sets

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 $c : V \rightarrow \{1, \dots, k\}$ so that for all $(u, v) \in E$, $c(u) \neq c(v)$. 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.

A set of vertices S is *independent* if there are no edges between vertices in S . In particular, each color class in a coloring is an independent set. The problem of finding large independent sets in a graph is NP-Complete, and it is very difficult to even approximate the size of the largest independent set in a graph.

However, for some carefully chosen graphs one can obtain very good bounds on the sizes of independent sets by using spectral graph theory. We may later see some uses of this theory in the analysis of error-correcting codes and sphere packings.

3.7 Hoffman's Bound

Hoffman proved the following upper bound on the size of an independent set in a graph G .

Theorem 3.7.1. *Let $G = (V, E)$ be a d -regular graph. Then*

$$\alpha(G) \leq n \frac{-\mu_n}{d - \mu_n}.$$

I gave an overly-complicated proof of this theorem in 2009. Part of the complication was that I wrote my proof using the adjacency matrix. I will now give a very simple proof using the Laplacian matrix. In fact, I will prove a slightly stronger statement that does not require the graph to be regular.

Theorem 3.7.2. *Let S be an independent set in G , and let $d_{ave}(S)$ be the average degree of a vertex in S . Then,*

$$|S| \leq n \left(1 - \frac{d_{ave}(S)}{\lambda_n} \right).$$

To compare these two, observe that in the d -regular case $d_{ave} = d$ and $\lambda_n = d - \mu_n$. So, we have

$$1 - \frac{d_{ave}(S)}{\lambda_n} = \frac{\lambda_n - d}{\lambda_n} = \frac{-\mu_n}{d - \mu_n}.$$

Proof. Let S be an independent set of vertices and let $d(S)$ be the sum of the degrees of vertices in S .

Recall that

$$\lambda_n = \max_{\mathbf{x}} \frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

We also know that the vector \mathbf{x} that maximizes this quantity is ϕ_n , and that ϕ_n is orthogonal to ϕ_1 . So, we can refine this expression to

$$\lambda_n = \max_{\mathbf{x} \perp \mathbf{1}} \frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

As we did last class, we will consider the vector

$$\mathbf{x} = \chi_S - s\mathbf{1},$$

where $s = |S|/n$. As S is independent, we have

$$\mathbf{x}^T \mathbf{L} \mathbf{x} = |\partial(S)| = d(S) = d_{ave}(S) |S|.$$

We also computed the square of the norm of \mathbf{x} last class, and it comes out to

$$\mathbf{x}^T \mathbf{x} = n(s - s^2).$$

So, we have

$$\lambda_n \geq \frac{d_{ave}(S) |S|}{n(s - s^2)} = \frac{d_{ave}(S) sn}{n(s - s^2)} = \frac{d_{ave}(S)}{1 - s}.$$

Re-arranging terms, this gives

$$1 - \frac{d_{ave}(S)}{\lambda_n} \geq s,$$

which is equivalent to the claim of the theorem. \square

Hoffman's bound using the adjacency matrix eigenvalues does not necessarily hold for irregular graphs. However, the bound that one would expect to get from it on the chromatic number does. As a k -colorable graph must have an independent set of size at least n/k , the expected theorem follows.

Theorem 3.7.3.

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

Using Theorem 3.7.2, we can prove (and you will do so on the first problem set)

$$\chi_G \geq \frac{\lambda_n}{\lambda_n - d_{ave}}.$$

These are the same in the regular case. I'm not sure which is better, or if they are even comparable, in general.

3.8 Wilf's Theorem

We did not get to the following material in today's lecture, but I assume that you will read it.

While we may think of μ_1 as being related to the average degree, it does behave differently. In particular, if we remove the vertex of smallest degree from a graph, the average degree can increase. On the other hand, μ_1 can only decrease when we remove a vertex. Let's prove that now.

Lemma 3.8.1. *Let A be a symmetric matrix with largest eigenvalue α_1 . Let B be the matrix obtained by removing the last row and column from A , and let β_1 be the largest eigenvalue of B . Then,*

$$\alpha_1 \geq \beta_1.$$

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

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

So, for \mathbf{y} an eigenvector of B of eigenvalue β_1 ,

$$\beta_1 = \frac{\mathbf{y}^T B \mathbf{y}}{\mathbf{y}^T \mathbf{y}} = \frac{\begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix}^T A \begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix}}{\begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix}^T \begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix}} \leq \max_{\mathbf{x} \in \mathbb{R}^n} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

□

Of course, this holds regardless of which row and column we remove, as long as they are the same row and column.

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 improves upon this bound.

Theorem 3.8.2.

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

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 zero¹. Now, assume the theorem is true for all graphs on $n - 1$ vertices, and let G be a graph on n vertices. By Lemma 3.4.1, G has a vertex of degree at most $\lfloor \mu_1 \rfloor$. Let v be such a vertex and let $G - \{v\}$ be the graph obtained by removing this vertex. By Lemma 3.8.1 and our induction hypothesis, $G - \{v\}$ 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 v . As v has at most $\lfloor \mu_1 \rfloor$ neighbors, there is some color in $\{1, \dots, \lfloor \mu_1 \rfloor + 1\}$ 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. □

For an example, consider a path graph with at least 3 vertices. We have $d_{max} = 2$, but $\alpha_1 < 2$.

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.

¹If this makes you uncomfortable, you could use both graphs on two vertices