Lecture 4: The spectrum of the Laplacian

Prepared by: Daniel Hsu

October 5, 2005
Revised: November 2, 2005

1 Introduction

We continue our study of random walks on undirected graphs, with a present focus on the spectrum of the Laplacian. As usual, for a graph $G = (V, E)$, let A be its adjacency matrix and D be the diagonal matrix with $D(v, v) = d_v$. Then, the random walk on G will be taken according to the transition matrix $P = D^{-1}A$. We also define the stationary distribution π with $\pi(x) = d_x / \operatorname{vol} G$.

Our discussion of random walks on G left off with the result

$$\|fP^t - \pi\|_2 \leq \max_{i \neq 0} |\rho_i|^t \frac{\max_x \sqrt{d_x}}{\min_y \sqrt{d_y}},$$

where f is a probability distribution (i.e. $f \geq 0$ and $\sum_x f(x) = 1$) and $1 = \rho_0 \geq \rho_1 \geq \ldots \geq \rho_{n-1}$ are the eigenvalues of P. This inequality implies that convergence to the stationary distribution π will follow if $\max\{|\rho_1|, |\rho_{n-1}|\} < 1$.

2 The Laplacian and the Rayleigh quotient

The transition probability matrix P is similar to the matrix $M = D^{\frac{1}{2}} P D^{-\frac{1}{2}}$, so P and M have the same eigenvalues. We previously introduced the Laplacian of the graph as $\mathcal{L} = I - M$, so it has eigenvalues $0 = \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_{n-1}$ (where $\lambda_i = 1 - \rho_i$).

The main tool we'll use to study the spectrum of $\mathcal{L}$ is the Rayleigh quotient $R(f)$ of $\mathcal{L}$, defined (for our purposes) as

$$R(f) = \frac{fLf^*}{fDf^*}$$

where $L = D - A$ is the combinatorial Laplacian. This is the same as the usual sense of the Rayleigh quotient $g\mathcal{L}g^*/gg^*$ with the subtitution $f = gD^{-\frac{1}{2}}$. Following this equivalence, if the ϕ_i are the eigenvectors of $\mathcal{L}$, we'll call the $\psi_i = \phi_i D^{-\frac{1}{2}}$ the harmonic eigenvectors of $\mathcal{L}$.

Employing the Rayleigh quotient, we see that the eigenvalue λ_1 can be written as

$$\lambda_1 = \inf_{\substack{f: \\ \sum_x f(x) d_x = 0}} R(f). \tag{1}$$

Since the eigenvector associated with λ_0 is $\phi_0 = \mathbf{1}D^{\frac{1}{2}}$, the condition $\sum_x f(x)d_x = 0$ is an orthogonality condition. Such variational characterizations can also be made for the other eigenvalues:

$$\lambda_{n-1} = \sup_f R(f)$$

and, in general,

$$\lambda_i = \sup_{h_0, h_1, \ldots, h_{i-1}} \inf_{\substack{f: \\ \sum_x f(x) h_j(x) d_x = 0 \\ \forall j \in \{0, \ldots, i-1\}}} R(f).$$

The following characterization of the Rayleigh quotient (demonstrated last time) will be useful later:

$$R(f) = \frac{\sum_{x \sim y} (f(x) - f(y))^2}{\sum_x f^2(x) d_x}. \tag{2}$$

To this point, we have done a lot of linear algebra. We are not here to teach linear algebra; we are here to take linear algebra one step further to understand what is happening in the graph.

3 Properties of the spectrum of $\mathcal{L}$

The graph spectrum reveals many properties of a graph, similarly as the stellar spectrum holds information about a star's makeup. Since direct measurement in both cases often proves difficult or impossible, this shows the importance of spectral analysis. In our study of random walks on graphs, we'll be mostly interested in ρ_1 and ρ_{n-1} (and thus λ_1 and λ_{n-1}), since that they determine the convergence properties of random walks.

Fact 1. *For graphs with no self-loops ($A(v, v) = 0$) and each $d_v \neq 0$, we have $\sum_{i=0}^{n-1} \lambda_i = n$.*

Proof. Follows from the fact that $\sum_{i=0}^{n-1} \lambda_i = \operatorname{trace} \mathcal{L}$ and that $\mathcal{L}$ has ones on its diagonal. □

Fact 2. *For $n \geq 2$, we have $\lambda_1 \leq \frac{n}{n-1}$, with equality if and only if $G = K_n$, the complete graph on n vertices.*

Proof. Since $\lambda_0 = 0$, the inequality follows immediately from Fact 1. We saw before that the spectrum of K_n is $(0, \frac{n}{n-1}, \ldots, \frac{n}{n-1})$. To prove that only K_n has $\lambda_1 = \frac{n}{n-1}$, we'll defer to the stronger Fact 3. □

Fact 3. *If $G \neq K_n$, then $\lambda_1 \leq 1$.*

Proof. Suppose $a \not\sim b$ in G. It suffices to construct an f satisfying $\sum_x f(x) d_x = 0$ (the orthogonality condition) and $R(f) = 1$. Since such an f satisfies $R(f) \geq \lambda_1$ by (1), the claim will follow. The following construction will do:

$$f(x) = \begin{cases} d_a & \text{if } x = b \\ -d_b & \text{if } x = a \\ 0 & \text{otherwise.} \end{cases}$$

The orthogonality condition is clearly satisfied. Using (2),

$$R(f) = \frac{d_b^2 d_a + d_a^2 d_b}{d_b^2 d_a + d_a^2 d_b} = 1,$$

so we're done. □

This shows that the spectrum can dectect whether a graph is complete, even if it is missing only one edge.

Fact 4. *The graph G is connected $\Leftrightarrow$ $\lambda_1 > 0$.*

This is a connection between a topological invariant (connectivity) and an analytic invariant (λ_1). These types of connections give strength to spectral methods.

Proof. Let f be a harmonic eigenvector associated with the eigenvalue $0 = \lambda_0$. Looking at (2), we see that $f(x) = f(y)$ for $x \sim y$. Thus, if G is connected, $f(x) = c$ for some constant c, i.e. $f = c\mathbf{1}$, which is unique up to scaling. So the zero eigenvalue has multiplicity one, giving $\lambda_1 > 0$.

Now, suppose G is not connected. Then G is the disjoint union of two graphs G_1 and G_2. Since the spectrum of G is the union of the spectra of G_1 and G_2, the eigenvalue 0 appears twice in G's spectrum. So $\lambda_1 = 0$. □

Fact 5. *The graph G has $i+1$ components $\Leftrightarrow \lambda_i = 0$ and $\lambda_{i+1} \neq 0$.*

Proof. Follows from Fact 4. □

Fact 6. *For all i, $0 \leq \lambda_i \leq 2$.*

Proof. From (2),

$$\begin{aligned} R(f) &= \frac{\sum_{x\sim y}(f(x)-f(y))^2}{\sum_x f^2(x)d_x} \\ &\leq \frac{\sum_{x\sim y} 2(f^2(x)+f^2(y))}{\sum_x f^2(x)d_x} \\ &= \frac{2\sum_x f^2(x)d_x}{\sum_x f^2(x)d_x} \\ &= 2. \end{aligned}$$

□

Fact 7. *The graph G is bipartite $\Leftrightarrow \lambda_{n-1} = 2$.*

Proof. Let f be the harmonic eigenvector associated with the eigenvalue 2. Then equality holds throughout in the proof of Fact 6. Because all of the terms under summation are nonnegative, we have $f(x) = -f(y)$ for $x \sim y$. So G is two-colorable and thus bipartite.

To prove the other implication, we start with G bipartite with edges only between the disjoint vertex sets V_1 and V_2. We construct f with $f(x) = 1$ if $x \in V_1$ and $f(x) = -1$ if $x \in V_2$. It is easy to check that this satisifes the orthogonality condition, and

$$R(f) = \frac{\sum_{x\sim y} 2^2}{\sum_x d_x} = \frac{4|E|}{2|E|} = 2.$$

□

Fact 6 was really a consequence of the normalization in $R(f)$, which comes from using $\mathcal{L}$ as opposed to using L directly. In general, the spectrum of the combinatorial Laplacian is heavily influenced by the graph's degree distribution; the degree of one vertex is a "local" property. For non-regular graphs, a few vertices of high degree hamper the spectrum's usefulness as an indicator of "global" properties. The situation is fine for regular graphs, as their degree distributions are constant; and in this case, normalization only scales. The combinatorial Laplacian does have its uses; for example, it can be used to count the number of spanning trees in a graph. A second reason for normalizing is to allow us to compare the spectra of two graphs more naturally.

4 Random walks and Markov chains

Returning to our discussion about random walks on graphs, we see that Facts 4 and 7 reveal the conditions under which $\lambda_1 > 0$ and $\lambda_{n-1} < 2$. In light of this, our previous discussion can now be concluded with the following theorem.

Theorem 1. *If G is connected (so $|1 - \lambda_1| < 1$) and not bipartite (so $|1 - \lambda_{n-1}| < 1$), then the random walk converges to its unique stationary distribution.*

Now, we draw a connection between random walks on graphs and Markov chains. A Markov chain is defined by its state transition matrix P. The matrix P has $\sum_v P(u,v) = 1$, where $P(u,v)$ is the probability of going from state u to state v. We can think of the states of a Markov chain as vertices in a graph. For a reversible Markov chain, there is a (stationary) probability distribution π that satisfies

$$\pi(u)P(u,v) = \pi(v)P(v,u).$$

Thus, transitions between two states are taken in each direction with equal probability. This is exactly the case for random walks on graphs with edge weights $w_{u,v} \propto \pi(u)P(u,v)$.

The property of Markov chains that governs the limiting behavior of fP^t is ergodicity. We say a Markov chain is ergodic if there is a unique stationary distribution π to which fP^t converges. Another notion of ergodic (as defined in some places) is that a system is ergodic if it mixes well. The necessary and sufficient conditions for ergodicity are that the Markov chain be

1. irreducible: can always move from one state to any other state in finite time, and
2. aperiodic: the chain doesn't "oscillate"; that is, the greatest common divisor of all closed circuits is 1.

The definitions given for these conditions are left intentionally vague, but the corresponding conditions for random walks on graphs are very concrete: the graph must be

1. connected, and
2. non-bipartite.

A word of caution: the discussion thus far has only been true for undirected graphs. The situation will be quite different when we encounter directed graphs!

Exercise 1. Determine the spectrum of the complete tripartite graph $K_{l,m,n}$. One will find that it satisfies the conditions for the random walk to converge.

It is good to look at the spectra of different graphs to build up a dictionary of spectras to be able to refer to.