Introducing the Laplacian

Prepared by: Steven Butler

October 3, 2005

1 A look ahead to weighted edges

To this point we have looked at random walks where at each vertex we have equal probability of moving to one of its neighbors. However, we might have a preferred neighbor, that is one that we are more likely to want to visit. We can model this by introducing weights on the edges. (Think of the weights as indicating how likely we are to use the edge in a walk.)

To model this we would let w_{uv} denote the weight of the edge going from u to v. Then the transition probability matrix will be defined entrywise by

$$P(u,v) = \frac{w_{uv}}{\sum_s w_{us}}.$$

We also have the stationary distribution defined by

$$\pi(u) = \frac{\sum_s w_{us}}{\operatorname{vol} G}.$$

(Note that underlying this is the idea that $\sum_s w_{us}$ denotes the degree of u.) Assigning different weights on the edges will lead to different weighted matrices.

If we have a directed graph and we want to have a reversible Markov chain, then we need $\pi(u)P(u,v) = \pi(v)P(v,u)$, which inserting the definitions simplifies to $w_{uv} = w_{vu}$. This shows that in order to have a reversible Markov chain not only must we have the flow be able to go in both directions but the weights of the edges in both directions must be the same.

At this point in the course we could start focusing on the weighted case but for now we will continue working with 0-1 matrices. There are two reasons. The first is that if we can deal with 0-1 matrices then we can deal with the more general case, indeed many of our proofs will instantly generalize. The second reason is that when we use 0-1 we will be able to emphasize the discrete structure, i.e., we will not blur the picture of what is going on by putting weights on the edges.

2 A short review

We are working with an undirected graph G and associated with this graph is the transition probability matrix P. In previous lectures we have seen that if $\pi(x) = d_x/\operatorname{vol} G$ then

$\pi P = \pi$, showing that π is the stationary distribution. We also saw that if we started with a probability distribution function f (i.e., a function where $0 \le f(v) \le 1$ for all v and $\sum_v f(v) = 1$) then for well-behaved graphs, $fP^t \to \pi$. More precisely if we let $1 = \rho_0 \ge \rho_1 \ge \cdots \ge \rho_{n-1}$ denote the eigenvalues of P then we saw that

$$\|fP^t - \pi\|_2 \le \max_{i\neq 0} |\rho_i|^t \frac{\max \sqrt{d_x}}{\min \sqrt{d_y}}.$$

So the question of convergence (and how quickly it converges) boils down to examining the quantity

$$\rho = \max_{i\neq 0} |\rho_i| = \max\{|\rho_1|, |\rho_{n-1}|\}.$$

We will (eventually) examine what the quantities ρ_1 and ρ_{n-1} can tell us, in particular we will answer the question of when they will be 1.

3 Introduction to the Laplacian

The quantity ρ is determined by the eigenvalues of $P = D^{-1}A$, in a previous lecture we saw that

$$P = D^{-1/2}\big(\underbrace{D^{-1/2}AD^{-1/2}}_{=M}\big)D^{1/2},$$

so that the eigenvalues of P are the same as the eigenvalues of M, a symmetric matrix. We can use properties of symmetric matrices to help us find and evaluate the eigenvalues.

Our main tool will be the Laplacian which is defined by

$$\mathcal{L} = I - M = I - D^{-1/2}AD^{-1/2} = D^{-1/2}(D - A)D^{-1/2}.$$

Note that $L = D - A$ is referred to as the combinatorial Laplacian while the Laplacian we have introduced here is referred to as the normalized Laplacian. Unless we specify, when we talk about the Laplacian we will be referring to the normalized Laplacian. It is important to note that the combinatorial Laplacian does have some useful properties which we will call on from time to time.

From the definition of the Laplacian we can compute it entrywise.

$$\mathcal{L}(u, v) = \begin{cases} 1 & \text{if } u = v; \\ \dfrac{-1}{\sqrt{d_u d_v}} & \text{if } u \sim v; \\ 0 & \text{otherwise.} \end{cases}$$

Since the Laplacian is symmetric all the eigenvalues are real, further it is easy to show that the Laplacian is positive semi-definite (i.e., all the eigenvalues are non-negative). We let $0 = \lambda_0 \le \lambda_1 \le \cdots \le \lambda_{n-1}$ denote the eigenvalues of the Laplacian.

The Laplacian is named after Pierre-Simon Laplace, the famous French mathematician who studied what has come to be known as the Laplace-Beltrami operator. The Laplacian that we are working with is the discrete analogue of the Laplace-Beltrami operator.

4 Facts about the Laplacian

We will now consider the ratio $g\mathcal{L}g^*/gg^*$ (here g^* denote the conjugate transpose of g). This ratio shows up (among other places) in the Courant-Fischer theorem for computing eigenvalues. So the maximum value that can be achieved by a g in this ratio will be λ_{n-1} while the minimum value will be λ_0.

Let us manipulate the ratio.

$$\frac{g\mathcal{L}g^*}{gg^*} = \frac{gD^{-1/2}LD^{-1/2}g^*}{gg^*} = \frac{fLf^*}{fDf^*},$$

where in the last step we made the substitution $f = gD^{-1/2}$ (or, equivalently, $g = fD^{1/2}$).

Fact 1. $\lambda_0 = 0$ with eigenvector $\phi_0 = \dfrac{\mathbf{1}D^{1/2}}{\sqrt{\operatorname{vol} G}}$.

Fact 2. $(fL)(x) = \big(f(D-A)\big)(x) = f(x)d_x - \sum_{y\sim x} f(y) = \sum_{\substack{y\\ y\sim x}} \big(f(x) - f(y)\big)$, and so

$$fLf^* = \sum_x \sum_{\substack{y\\ y\sim x}} f(x)\big(f(x)-f(y)\big) = \sum_{\substack{x\sim y\\ \text{unordered}\\ \text{pairs}}} \big(f(x)-f(y)\big)^2.$$

The last step is done by combining the term $x{\sim}y$ with the term $y{\sim}x$. The end term is sometimes referred to as a Dirichlet sum.

Fact 3. $\dfrac{g\mathcal{L}g^*}{gg^*} = \dfrac{fLf^*}{fDf^*} = \dfrac{\sum_{x\sim y}\big(f(x)-f(y)\big)^2}{\sum_x \big(f(x)\big)^2 d_x}$.

This ratio is sometimes referred to as the Rayleigh quotient. One of the key reasons that we consider the Laplacian is that it naturally gives us the Rayleigh quotient.

Fact 4 (Courant-Fischer)**.**

$$\begin{aligned}
\lambda_1 &= \inf_{\substack{f\\ \sum_x f(x)d_x = 0}} \frac{\sum_{x\sim y}\big(f(x)-f(y)\big)^2}{\sum_x \big(f(x)\big)^2 d_x} \\
&= \inf_f \sup_c \frac{\sum_{x\sim y}\big(f(x)-f(y)\big)^2}{\sum_x \big(f(x)-c\big)^2 d_x} && (1)\\
&= 2\operatorname{vol} G \inf_f \frac{\sum_{x\sim y}\big(f(x)-f(y)\big)^2}{\sum_{x,y}\big(f(x)-f(y)\big)^2 d_x d_y} && (2)
\end{aligned}$$

Exercise 1. Establish the equality of (1).

Exercise 2. Establish the equality of (2).

5 Examples of spectras

As a general rule, the more symmetric the graph is the cleaner the spectra. So all of our examples we will present below are graphs with high symmetry.

Example 1. For the complete graph K_n the eigenvalues of the Laplacian are 0 (with multiplicity 1) and $n/(n-1)$ (with multiplicity of $n-1$). To see this note that

$$\mathcal{L} = \begin{pmatrix} 1 & \frac{-1}{n-1} & \cdots & \frac{-1}{n-1} \\ \frac{-1}{n-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{-1}{n-1} \\ \frac{-1}{n-1} & \cdots & \frac{-1}{n-1} & 1 \end{pmatrix}.$$

Now let $\vartheta = e^{2\pi ik/n}$ for some $k = 0, 1, \ldots, n-1$ be an nth root of unity. Then

$$\begin{pmatrix} 1 & \frac{-1}{n-1} & \cdots & \frac{-1}{n-1} \\ \frac{-1}{n-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{-1}{n-1} \\ \frac{-1}{n-1} & \cdots & \frac{-1}{n-1} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vartheta \\ \vdots \\ \vartheta^{n-1} \end{pmatrix} = \big(1 - \frac{1}{n-1}(\vartheta + \cdots + \vartheta^{n-1})\big) \begin{pmatrix} 1 \\ \vartheta \\ \vdots \\ \vartheta^{n-1} \end{pmatrix}.$$

When $k = 0$ the expression $\big(1 - \frac{1}{n-1}(\vartheta + \cdots + \vartheta^{n-1})\big)$ simplifies to 0, otherwise (using the property that the sum of all nth roots is 0) we get $n/(n-1)$.

Exercise 3. Show that the eigenvalues of the Laplacian for the complete bipartite graph $K_{m,n}$ are 0 (with multiplicity 1), 1 (with multiplicity $m+n-2$), and 2 (with multiplicity 1).

Example 2. For the cycle on n vertices (denoted C_n) the eigenvalues of the Laplacian are $1 - \cos(2\pi k/n)$ for $k = 0, 1, \ldots, n-1$. To see this note that

$$\mathcal{L} = \begin{pmatrix} 1 & -\frac{1}{2} & O & -\frac{1}{2} \\ -\frac{1}{2} & \ddots & \ddots & O \\ O & \ddots & \ddots & -\frac{1}{2} \\ -\frac{1}{2} & O & -\frac{1}{2} & 1 \end{pmatrix}.$$

Again let $\vartheta = e^{2\pi ik/n}$ for some $k = 0, 1, \ldots, n-1$ be an nth root of unity. Then

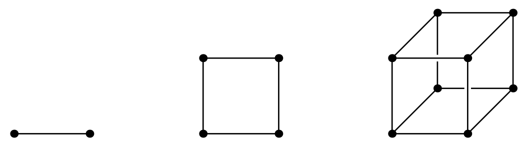
$$\begin{pmatrix} 1 & -\frac{1}{2} & O & -\frac{1}{2} \\ -\frac{1}{2} & \ddots & \ddots & O \\ O & \ddots & \ddots & -\frac{1}{2} \\ -\frac{1}{2} & O & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vartheta \\ \vdots \\ \vartheta^{n-1} \end{pmatrix} = \big(1 - \frac{\vartheta + \vartheta^{-1}}{2}\big) \begin{pmatrix} 1 \\ \vartheta \\ \vdots \\ \vartheta^{n-1} \end{pmatrix}.$$

Since $1 - (\vartheta + \vartheta^{-1})/2 = 1 - \cos(2\pi k/n)$ the result follows.

Note that the spectrum of the cycle is important and shows up, for example, in the Fast Fourier Transform. A key reason that the cycle shows up frequently in applications is that locally it looks like a line (even more than the path because the path has the two end vertices) and so can be used as a discrete model for $\mathbb{R}$.

Exercise 4. Show that the eigenvalues of P_n (the path on n vertices) are $1-\cos(\pi k/(n-1))$ for $k = 0, 1, \ldots, n-1$.

The last graph we will look at is the n-cube Q_n. This is the graph where vertices are indexed by Boolean words of length n and two vertices are adjacent if they differ in exactly one letter. Such graphs have 2^n vertices, $n2^{n-1}$ edges, and beautiful structure.



Q_1 Q_2 Q_3

Exercise 5. Show that the eigenvalues of the n-cube Q_n are $2k/n$ (with multiplicity $\binom{n}{k}$) for $k = 0, 1, \ldots, n$.