# Introduction to Spectral Graph Theory

Madeleine Udell

March 14, 2011

## 1 Introduction

Up to this point in the class, we have been looking at a graph as a collection of vertices and edges, and most of our algorithms have been very tied to this representation. Today we'll start to look at a very different way of representing the graph, using the eigenvalues and eigenvectors of matrices associated with the graph rather than the vertices and edges themselves.

This lecture is meant as an introduction and overview of some nice ideas from spectral graph theory, so I will occasionally state results without proof. For a more thorough treatment, see the references.

## 2 Matrices associated with a graph

Given a graph $G = (V, E)$, one can define a number of different matrices that one might care to study.

### 2.1 Adjacency Matrix

The adjacency matrix $A$ is defined by

$$A_{i,j} = \begin{cases} 1 & (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

### 2.2 Degree Matrix

The degree matrix $D$ is a diagonal matrix with the degree $d_i$ of each vertex as the diagonal entries. Equivalently, it contains in each diagonal entry $d_{ii}$ the sum of the $i^{th}$ row of the adjacency matrix. Formally,

$$D_{i,j} = \begin{cases} d_i & i = j \\ 0 & \text{otherwise} \end{cases}$$

## 2.3 Laplacian

This lecture will largely be concerned with the Laplacian matrix, which we define by

$$L = D - A = \begin{cases} d_i & i = j \\ -1 & (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

This is a rather nice matrix to work with from a linear algebra perspective because it defines a symmetric, positive semi-definite quadratic form on the vertices. When thinking about the action defined by the Laplacian, it helps to think of the vectors the Laplacian acts on as functions $f$ on the vertices, assigning a real number to each vertex. Although all of the graphs we'll consider today are undirected, for the purposes of this discussion it helps to define the set of directed edges $\vec{E}$ and to consider each undirected edge $(i, j)$ as the two directed edges $(i, j)$ and $(j, i) \in \vec{E}$. Then the action of the Laplacian is defined via

$$\begin{aligned} (Lf)_i &= \sum_{(i,j)\in\vec{E}} (f(i) - f(j)) \\ f^T Lf &= \sum_{i\in V} f(i) \sum_{(i,j)\in\vec{E}} (f(i) - f(j)) \\ &= \sum_{(i,j)\in E} f(i)(f(i) - f(j)) + f(j)(f(j) - f(i)) \\ &= \sum_{(i,j)\in E} (f(i) - f(j))^2 \end{aligned}$$

You can see that the action of the Laplacian on a vector is measuring the smoothness of that vector across the edges in the graph. If the vector $f$ had the same value on every vertex connected by an edge, then the action of the Laplacian on that vector would return 0. This allows us to find the first eigenvector of the Laplacian:

$$L\vec{1} = 0 \cdot \vec{1}$$

where we defined

$$\vec{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Writing the Laplacian in this form also makes it easy to see that the Laplacian is positive semi-definite, since it is a sum of squares: $f^T Lf \geq 0$.

In the following, we will refer to the eigenvalue 0 as the first eigenvalue, with the other eigenvalues obeying

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

We can also use this representation of the Laplacian quadratic form as a sum over edges to write the matrix of the Laplacian as a sum over edges. This representation is more simple than the representation in terms of the vertices, and indicates that we may get more mileage out of considering the fundamental entities to be edges than vertices when we work with the Laplacian. If we define

$$(L_{e=(i,j)})_{k,l} = \begin{cases} 1 & i = k = l \text{ or } j = k = l \\ -1 & i = k, j = l \text{ or } j = k, i = l \\ 0 & \text{otherwise} \end{cases}$$

Then we can write

$$L = \sum_{e \in E} L_e$$

### 2.4 Random Walk Matrix

If we start a random walk at a given vertex, and choose an edge at random to follow to a new vertex $k$ times, where do we end up? The random walk matrix $W = D^{-1}A$ answers this question. If we begin in a probability distribution $\pi^0$, the probability distribution after $k$ steps of the random walk will be

$$\pi^k = \pi^0 W^k$$

### 2.5 Normalized Laplacian

The normalized Laplacian is a symmetric version of the random walk matrix. Its symmetry and invariance under scaling the graph (adding more nodes) makes it a useful theoretical tool. We define the normalized Laplacian

$$L_N = D^{-1/2} A D^{-1/2}$$

### 2.6 Weighted Laplacian

For edges of weights $w_e$, the weighted Laplacian is defined as

$$L = \sum_{e \in E} w_e L_e$$

## 3 Examples

### 3.1 Path Graph

A path graph is a set of $N$ vertices lined up in a row, with edges between each vertex and its neighbors on each side. The two vertices on either end only have one neighbor and hence one edge. Considering the Laplacian on the path graph provides a nice connection with the traditional Laplacian operator $L = -\frac{\partial^2}{\partial x^2}$. The Laplacian matrix for the path graph is

$$L = \begin{bmatrix} 1 & -1 & 0 & 0 & \ddots \\ -1 & 2 & -1 & 0 & \ddots \\ 0 & -1 & 2 & -1 & \ddots \\ \vdots & & & & \\ \ddots & 0 & -1 & 2 & -1 \\ \ddots & 0 & 0 & -1 & 1 \end{bmatrix}$$

In particular, the Laplacian applied to a vector $f$ gives

$$(Lf)_i = -f_{i-1} + 2f_i - f_{i+1}$$

To see the connection with the continuous Laplacian, consider a finite difference approximation to the second derivative of a function discretized on a 1D grid with spacing $h$ between points $\{x_1, \ldots, x_n\}$. The approximations to the first derivative at $x_{i-1/2}$ & $x_{i+1/2}$ are

$$\frac{\partial}{\partial x}\Big|_{x_{i-1/2}} = \frac{f_{x_i} - f_{x_{i-1}}}{h}$$
$$\frac{\partial}{\partial x}\Big|_{x_{i+1/2}} = \frac{f_{x_{i+1}} - f_{x_i}}{h}$$

and we can approximate the second derivative by the central difference

$$\frac{\partial^2}{\partial x^2}\Big|_x = \frac{\frac{\partial}{\partial x}\big|_{x_{i+1/2}} - \frac{\partial}{\partial x}\big|_{x_{i-1/2}}}{h} = \frac{-f_{x_{i+1}} + 2f_{x_i} - f_{x_{i-1}}}{h^2}$$

which is just a scaling of the graph Laplacian on the path graph.

## 3.2 Complete Graph

The complete graph on $N$ vertices has edges between each pair of vertices, so

$$L = \begin{cases} n-1 & i = j \\ -1 & \text{otherwise} \end{cases}$$

It is easy to show that the complete graph has one eigenvalue zero and all other eigenvalues $n$ by using the fact that all eigenvectors with non-zero eigenvalue are orthogonal to the constant vector $\vec{1}$.

### 3.3 Ring Graph

A ring graph is a set of $N$ vertices lined up along a circle, with edges between each vertex and its neighbors on each side. A nice geometric argument shows that the eigenvectors of the graph Laplacian are

$$(x_k)_i = \cos\left(\frac{2\pi k i}{N}\right)$$

and

$$(y_k)_i = \sin\left(\frac{2\pi k i}{N}\right)$$

for $k = 1, \ldots, \frac{N-1}{2}$.

To show this, you can compute the action of $\frac{1}{2}A$ on $x_k$ and $y_k$ *geometrically* by drawing the graph using the eigenvectors $x_k$ and $y_k$ as coordinates. The points $(\frac{1}{2}Ax_k, \frac{1}{2}Ay_k)$ lie on a circle centered at the origin with a smaller radius, since the new location of the vertex is at the midpoint of the line connecting its neighbors on the outer circle. We can compute the eigenvalues by using the cosine or sine addition rule, or more easily by considering the $N^{th}$ entry of $Ax_k$. The neighbors of the $N^{th}$ vertex are vertices 1 and $N-1$:

$$\begin{aligned}
(Ax_k)_n &= \cos\left(\frac{2\pi k(1)}{N}\right) + \cos\left(\frac{2\pi k(N-1)}{N}\right) \\
&= \cos\left(\frac{2\pi k}{N}\right) + \cos\left(\frac{-2\pi k}{N}\right) \\
&= 2\cos\left(\frac{2\pi k}{N}\right)
\end{aligned}$$

Since the ring graph is a regular graph (all the vertices have the same degree), the eigenvectors of $A$ are also eigenvectors of $L$. So we can compute the eigenvalues of the Laplacian, $\lambda_k$, as

$$\begin{aligned}
\lambda_k &= \frac{(Lx_k)_n}{(x_k)_n} = \frac{((2I-A)x_k)_n}{(x_k)_n} \\
&= \frac{2\cos\left(\frac{2\pi kN}{N}\right) - 2\cos\left(\frac{2\pi k}{N}\right)}{\cos\left(\frac{2\pi kN}{N}\right)} \\
&= 2 - 2\cos\left(\frac{2\pi k}{N}\right)
\end{aligned}$$

A similar computation shows that the eigenvalue for the eigenvector $y_k$ is the same.

Going the other direction, we can see how this might be very useful: on this example, if we just pop the graph into Matlab and calculate the eigenvectors, the eigenvectors of lowest eigenvalue give us a good coordinate system for the graph. A natural question to ask is whether this procedure would also work for more complicated graphs than simple rings. This idea leads to the use of the Laplacian eigenvectors in spectral embeddings.

# 4 Applications

## 4.1 Spectral Embeddings

Considering the Laplacian again as a quadratic form measuring the difference between values of a vector across all edges in a graph, we see that a Laplacian eigenvector of low eigenvalue must have values that are very similar on adjacent vertices. This property makes Laplacian eigenvectors of low eigenvalue useful for generating embeddings of arbitrary graphs into Euclidean space, since a good coordinate system for representing a graph should also have the property that vertices connected by an edge are close by.

This insight is the basis for a number of nonlinear dimensionality reduction techniques. The problem of nonlinear dimensionality reduction is this: given a set of data points in a very high dimensional space, one wants to embed the points in a lower dimensional space such that the distances between points are approximately preserved. But sometimes one doesn't trust the distances between points that are far away, and hence one only wishes to preserve distances between nearby points. Following the intuition of the last example, one might try writing a graph with edges connecting nearby vertices in the original space, and using eigenvectors of low eigenvalue as coordinates in a lower dimensional space.

## 4.2 Graph Coloring Heuristics

On the other side of the spectrum, the example of the Ring graph also had eigenvectors of high eigenvalue that oscillated in value between vertices — in that example, the highest eigenvector alternated in value between $+1$ and $-1$ on adjacent vertices. This leads to a heuristic idea for graph coloring problems: one can look at eigenvectors with high eigenvalue, and color vertices different colors according to their value in that eigenvector. A graphical way to understand this solution is to think once again of using the eigenvectors as coordinates — this time, as coordinates that spread out adjacent vertices. One can then divide up the coordinate system into distinct regions, and color all the vertices in a given region the same color.

## 4.3 Graph Isomorphism

Two graphs $G$ and $H$ are isomorphic if and only if there is a permutation of the labeling of the vertices such that the two graphs have all the same edges. Hence, $H$ and $G$ are isomorphic iff

$$P^T L_G P = L_H$$

for some permutation $P$.

Now, since eigenvalues are invariant under unitary transformations, and permutations are a subset of unitary transformations, it follows that $H$ and $G$ must share all of their eigenvalues if they are isomorphic. The converse is not true, but there are tricks for partical converses. For example, if we can find an eigenvalue whose corresponding eigenvector has all of its entries distinct, then that eigenvector defines a unique permutation of vertices. Some graphs, alas, have no such thing: we've already seen the example of the complete graph, which had no way of distinguishing between vertices by their values in a particular eigenvector (which we should expect, given the symmetries of the graph). Another example of a graph with highly degenerate eigenvectors is the hypercube.

# 5 Connectivity of a Graph

The spectrum of a graph gives us an easy way to tell whether a graph is connected or not.

**Theorem 1** *The multiplicity of the eigenvalue 0 in the spectrum of the graph Laplacian is the number of connected components of the graph*

**Proof:** If the graph $G$ consists of $k$ connected components $G_1, \ldots, G_k$ which are not connected to each other, then each of the characteristic vectors

$$(\mathbf{1}_i)_j = \begin{cases} 1 & j \in G_i \\ 0 & \text{otherwise} \end{cases}$$

is an eigenvector of eigenvalue 0. And all such eigenvectors are orthogonal to each other, since their support is disjoint. For the other direction, given any eigenvector $v$ of eigenvalue 0,

$$v^T L_G v = \sum_{e=(i,j)} (x(i) - x(j))^2$$

so the value of $x$ must be the same across every edge, and hence the same on each connected component. Hence we can always decompose the space spanned by $k$ such vectors in terms of characteristic vectors of the connected components, of which there must be at least $k$ to span the same eigenspace. ■

This theorem gives a qualitative understanding of the meaning of the second eigenvalue: the graph is disconnected if and only if the second smallest eigenvalue is zero. We can generalize this to a quantitative statement by defining $\lambda_2$ to be the *algebraic connectivity* of a graph $G$. This value is also known as the *Fiedler value* and the corresponding eigenvector $v_2$ as the *Fiedler vector*.

This value does indeed give good bounds on the connectedness of a graph. For example, we can define the *conductance* of set of vertices $S \subset V$ to be

$$\phi(S) = d(V) \frac{E(S, \bar{S}}{d(S) d(\bar{S})}$$

where $d(S)$ is the sum of the degrees of the vertices in $S$, and $\bar{S} = V \backslash S$, and define the conductance of a graph to be

$$\phi(G) = \min_{S \subset V} \phi(S)$$

Then one can prove

**Theorem 2** *Cheeger Inequality*

$$\frac{\phi^2}{2d} \leq \lambda_2 \leq \phi$$

Where $d$ is the maximum degree of any vertex in the graph. This says nice things whether $\lambda_2$ is big or small. If $\lambda_2$ is small, then there exists a cut disconnecting the graph that cuts very few edges. If $\lambda_2$ is large, then *every* cut disconnecting the graph cuts a large number of edges.

# 6 Exercises

In this problem, we use the *reflection principle* from the study of partial differential equations to find the eigenvalues and eigenvectors of the path graph, i.e. two vertices have degree one and all other vertices have degree two.

1. Write the Laplacian matrix for the path graph on $N$ vertices, and for the ring graph (a simple cycle) on $2N$ vertices.

2. Show that if $v \in \mathbb{R}^N$ is an eigenvector for the path graph, then

$$w = (v_1, v_2, \ldots, v_n, v_n, v_{n-1}, \ldots, v_1)^\top$$

is an eigenvector for the ring graph on $2N$ vertices. Likewise, show that if $w$ is an eigenvector of the ring graph obeying $w_i = w_{2N+1-i}$ for $i = 1, \ldots, N$, then

$$v = (w_1, \ldots, w_n)^\top$$

is an eigenvector for the path graph.

3. Recall the geometric argument from class showing that the eigenvectors of the ring graph on $N$ vertices are

$$(x_k)_i = \cos\left(\frac{2\pi k i}{N}\right)$$

and

$$(y_k)_i = \sin\left(\frac{2\pi k i}{N}\right)$$

Use this to find the eigenvalues and eigenvectors of the path graph on $N$ vertices.

**Hint:** Both $x_k$ and $y_k$ have the same eigenvalue. So any vector in the space spanned by these vectors is an eigenvector with the same eigenvalue.

4. This trick is called the reflection principle. If a neighbor in a mirror is treated exactly as a real neighbor, show how to set up mirrors next to a path graph in order to make the connection between the ring and path graphs geometrically obvious.

# 7 References

1. Daniel Spielman's Notes on Spectral Graph Theory, Fall 2009.
   http://www.cs.yale.edu/homes/spielman/561/