Spectral Graph Theory Lecture 2

The Laplacian

Daniel A. Spielman September 4, 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 3, 2015.

2.1 What I forgot to say last lecture

There are two important things that I forgot to say last lecture. The first is that, while we study many abstractly defined graphs to build our intuition, most of the theorems we prove hold in general.

The second is an extension of the characterization of eigenvalues and eigenvectors as optimization problems. Using the same reasoning, one can prove the following theorem.

Theorem 2.1.1. *Let $\boldsymbol{M}$ be an $n \times n$ symmetric matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ with corresponding eigenvectors $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_n$. Then,*

$$\lambda_i = \min_{\boldsymbol{x} \perp \boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i-1}} \frac{\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}},$$

and the eigenvectors satisfy

$$\boldsymbol{\psi}_i = \arg \min_{\boldsymbol{x} \perp \boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i-1}} \frac{\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

2.2 The Laplacian Matrix

Recall that the Laplacian Matrix of a weighted graph $G = (V, E, w)$, $w : E \to \mathbb{R}^+$, is designed to capture the Laplacian quadratic form:

$$\boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x} = \sum_{(u,v) \in E} w_{u,v} (\boldsymbol{x}(u) - \boldsymbol{x}(v))^2. \tag{2.1}$$

We will now use this quadratic form to derive the structure of the matrix. To begin, consider a graph with just two vertices and one edge. Let's call it $G_{1,2}$. We have

$$\boldsymbol{x}^T \boldsymbol{L}_{G_{1,2}} \boldsymbol{x} = (\boldsymbol{x}(1) - \boldsymbol{x}(2))^2. \tag{2.2}$$

Consider the vector $\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2$, where by $\boldsymbol{\delta}_i$ I mean the elementary unit vector with a 1 in coordinate i. We have

$$\boldsymbol{x}(1) - \boldsymbol{x}(2) = \boldsymbol{\delta}_1^T \boldsymbol{x} - \boldsymbol{\delta}_2^T \boldsymbol{x} = (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2)^T \boldsymbol{x},$$

so

$$(\boldsymbol{x}(1) - \boldsymbol{x}(2))^2 = \left((\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2)^T \boldsymbol{x}\right)^2 = \boldsymbol{x}^T (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2)(\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2)^T \boldsymbol{x} = \boldsymbol{x}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \boldsymbol{x}.$$

So,

$$\boldsymbol{L}_{G_{1,2}} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Now, let $G_{u,v}$ be the graph with just one edge between u and v. It can have as many other vertices as you like. The Laplacian of $G_{u,v}$ can be written in the same way: $\boldsymbol{L}_{G_{u,v}} = (\boldsymbol{\delta}_u - \boldsymbol{\delta}_v)(\boldsymbol{\delta}_u - \boldsymbol{\delta}_v)^T$. This is the matrix that is zero except at the intersection of rows and columns indexed by u and v, where it looks looks like

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Summing the matrices for every edge, we obtain

$$\boldsymbol{L}_G = \sum_{(u,v) \in E} w_{u,v} (\boldsymbol{\delta}_u - \boldsymbol{\delta}_v)(\boldsymbol{\delta}_u - \boldsymbol{\delta}_v)^T = \sum_{(u,v) \in E} w_{u,v} \boldsymbol{L}_{G_{u,v}}.$$

You should check that this agrees with the definition of the Laplacian from last class:

$$\boldsymbol{L}_G = \boldsymbol{D}_G - \boldsymbol{A}_G,$$

where

$$\boldsymbol{D}_G(u,u) = \sum_v w_{u,v}.$$

This formula turns out to be useful when we view the Laplacian as an operator. For every vector $\boldsymbol{x}$ we have

$$(\boldsymbol{L}_G \boldsymbol{x})(u) = d(u)\boldsymbol{x}(u) - \sum_{(u,v) \in E} w_{u,v} \boldsymbol{x}(v) = \sum_{(u,v) \in E} w_{u,v} (\boldsymbol{x}(u) - \boldsymbol{x}(v)). \tag{2.3}$$

2.3 Drawing with Laplacian Eigenvalues

I will now explain the motivation for the pictures of graphs that I drew last lecture using the Laplacian eigenvalues. Well, the real motivation was just to convince you that eigenvectors were cool. The following is the technical motivation. It should come with the caveat that it does not produce nice pictures of all graphs. In fact, it produces bad pictures of most graphs. But, it is still the first thing I always try when I encounter a new graph that I want to understand.

This approch to using eigenvectors to draw graphs was suggested by Hall [Hal70] in 1970.

To explain Hall's approach, I'll begin by describing the problem of drawing a graph on a line. That is, mapping each vertex to a real number. It isn't easy to see what a graph looks like when you do this, as all of the edges sit on top of one another. One can fix this either by drawing the edges of the graph as curves, or by wrapping the line around a circle.

Let $\boldsymbol{x} \in \mathbb{R}^V$ be the vector that describes the assignment of a real number to each vertex. We would like most pairs of vertices that are neighbors to be close to one another. So, Hall suggested that we choose an $\boldsymbol{x}$ minimizing (2.1). Unless we place restrictions on $\boldsymbol{x}$, the solution will be degenerate. For example, all of the vertices could map to 0. To avoid this, and to fix the scale of the embedding overall, we require

$$\sum_{u \in V} \boldsymbol{x}(u)^2 = \|\boldsymbol{x}\|^2 = 1. \tag{2.4}$$

Even with this restriction, another degenerate solution is possible: it could be that every vertex maps to $1/\sqrt{n}$. To prevent this from happening, we add the additional restriction that

$$\sum_{u} \boldsymbol{x}(u) = \mathbf{1}^T \boldsymbol{x} = 0. \tag{2.5}$$

On its own, this restriction fixes the shift of the embedding along the line. When combined with (2.4), it guarantees that we get something interesting.

As $\mathbf{1}$ is the eigenvector of the 0 eigenvalue of the Laplacian, the nonzero vectors that minimize (2.1) subject to (2.5) are the eigenvectors of the Laplacian of eigenvalue λ_2. When we impose the additional restriction (2.4), we eliminate the zero vectors, and obtain an eigenvector of norm 1.

Of course, we really want to draw a graph in two dimensions. So, we will assign two coordinates to each vertex given by $\boldsymbol{x}$ and $\boldsymbol{y}$. As opposed to minimizing (2.1), we will minimize

$$\sum_{(u,v) \in E} \left\| \begin{pmatrix} \boldsymbol{x}(u) \\ \boldsymbol{y}(u) \end{pmatrix} - \begin{pmatrix} \boldsymbol{x}(v) \\ \boldsymbol{y}(v) \end{pmatrix} \right\|^2 .$$

This turns out not to be so different from minimizing (2.1), as it equals

$$\sum_{(u,v) \in E} (\boldsymbol{x}(u) - \boldsymbol{x}(v))^2 + (\boldsymbol{y}(u) - \boldsymbol{y}(v))^2 = \boldsymbol{x}^T \boldsymbol{L} \boldsymbol{x} + \boldsymbol{y}^T \boldsymbol{L} \boldsymbol{y}.$$

As before, we impose the scale conditions

$$\|\boldsymbol{x}\|^2 = 1 \quad \text{and} \quad \|\boldsymbol{y}\|^2 = 1,$$

and the centering constraints

$$\mathbf{1}^T \boldsymbol{x} = 0 \quad \text{and} \quad \mathbf{1}^T \boldsymbol{y} = 0.$$

However, this still leaves us with the degnerate solution $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{\psi}_2$. To ensure that the two coordinates are different, Hall introduced the restriction that $\boldsymbol{x}$ be orthogonal to $\boldsymbol{y}$. One can use the characterization of eigenvalues that we derived last lecture to prove that the solution is then given by setting $\boldsymbol{x} = \boldsymbol{\psi}_2$ and $\boldsymbol{y} = \boldsymbol{\psi}_3$, or by taking a rotation of this solution (this is a problem on the first problem set).

2.4 Isoperimetry and λ_2

Computer Scientists are often interested in cutting, partitioning, and clustering graphs. Their motivations range from algorithm design to data analysis. We will see that the second-smallest eigenvalue of the Laplacian is intimately related to the problem of dividing a graph into two pieces without cutting too many edges.

Let S be a subset of the vertices of a graph. One way of measuring how well S can be separated from the graph is to count the number of edges connecting S to the rest of the graph. These edges are called the *boundary* of S, which we formally define by

$$\partial(S) \stackrel{\text{def}}{=} \{(u,v) \in E : u \in S, v \notin S\}.$$

We are less interested in the total number of edges on the boundary than in the ratio of this number to the size of S itself. For now, we will measure this in the most natural way–by the number of vertices in S. We will call this ratio the *isoperimetric ratio* of S, and define it by

$$\theta(S) \stackrel{\text{def}}{=} \frac{|\partial(S)|}{|S|}.$$

The *isoperimetric number* of a graph is the minimum isoperimetric number over all sets of at most half the vertices:

$$\theta_G \stackrel{\text{def}}{=} \min_{|S| \leq n/2} \theta(S).$$

We will now derive a lower bound on θ_G in terms of λ_2. We will present an upper bound, known as Cheeger's Inequality, in a later lecture.

Theorem 2.4.1. *For every $S \subset V$*

$$\theta(S) \geq \lambda_2(1-s),$$

where $s = |S| / |V|$. In particular,

$$\theta_G \geq \lambda_2/2.$$

Proof. As

$$\lambda_2 = \min_{\boldsymbol{x}:\boldsymbol{x}^T\mathbf{1}=0} \frac{\boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}},$$

for every non-zero $\boldsymbol{x}$ orthogonal to $\mathbf{1}$ we know that

$$\boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x} \geq \lambda_2 \boldsymbol{x}^T \boldsymbol{x}.$$

To exploit this inequality, we need a vector related to the set S. A natural choice is $\boldsymbol{\chi}_S$, the characteristic vector of S,

$$\boldsymbol{\chi}_S(u) = \begin{cases} 1 & \text{if } u \in S \\ 0 & \text{otherwise.} \end{cases}$$

We find

$$\boldsymbol{\chi}_S^T \boldsymbol{L}_G \boldsymbol{\chi}_S = \sum_{(u,v)\in E} (\boldsymbol{\chi}_S(u) - \boldsymbol{\chi}_S(v))^2 = |\partial(S)| \,.$$

However, $\boldsymbol{\chi}_S$ is not orthogonal to $\mathbf{1}$. To fix this, use

$$\boldsymbol{x} = \boldsymbol{\chi}_S - s\mathbf{1},$$

so

$$\boldsymbol{x}(u) = \begin{cases} 1-s & \text{for } u \in S, \text{ and} \\ -s & \text{otherwise.} \end{cases}$$

We have $\boldsymbol{x}^T \mathbf{1} = 0$, and

$$\boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x} = \sum_{(u,v)\in E} ((\boldsymbol{\chi}_S(u) - s) - (\boldsymbol{\chi}_S(v) - s))^2 = |\partial(S)| \,.$$

To finish the proof, we compute

$$\boldsymbol{x}^T \boldsymbol{x} = |S| \, (1-s)^2 + (|V| - |S|)s^2 = |S| \, (1 - 2s + s^2) + |S| \, s - |S| \, s^2 = |S| \, (1-s).$$

□

This theorem says that if λ_2 is big, then G is very well connected: the boundary of every small set of vertices is at least λ_2 times something just slightly smaller than the number of vertices in the set.

We will use the computation in the last line of that proof often, so we will make it a claim.

Claim 2.4.2. *Let $S \subseteq V$ have size $s \, |V|$. Then*

$$\| \chi_S - s\mathbf{1} \|^2 = s(1-s) \, |V| \,.$$

2.5 The Animals in the Zoo

We now examine the eigenvalues and eigenvectors of the Laplacians of some fundamental graphs. It is important to see many examples like these. They will help you develop your intuition for how eigenvalues behave. As you encounter new graphs, you will compare them to the graphs that you already know and hope that they behave similarly.

Today we will examine

- The complete graph on n vertices, K_n, which has edge set $\{(u,v) : u \neq v\}$.
- The star graph on n vertices, S_n, which has edge set $\{(1,u) : 2 \leq u \leq n\}$.
- The hypercube, which we defined last lecture.

As all these graphs are connected, they all have eigenvalue zero with multiplicity one.

Lemma 2.5.1. *The Laplacian of K_n has eigenvalue 0 with multiplicity 1 and n with multiplicity $n-1$.*

Proof. To compute the non-zero eigenvalues, let $\boldsymbol{\psi}$ be any non-zero vector orthogonal to the all-1s vector, so

$$\sum_u \boldsymbol{\psi}(u) = 0. \tag{2.6}$$

We now compute the first coordinate of $\boldsymbol{L}_{K_n}\boldsymbol{\psi}$. Using (2.3), we find

$$\left(\boldsymbol{L}_{K_n}\boldsymbol{\psi}\right)(1) = \sum_{v \geq 2}(\boldsymbol{\psi}(1) - \boldsymbol{\psi}(v)) = (n-1)\boldsymbol{\psi}(1) - \sum_{v=2}^{n}\boldsymbol{\psi}(v) = n\boldsymbol{\psi}(1), \quad \text{by (2.6)}.$$

As the choice of coordinate was arbitrary, we have $\boldsymbol{L}\boldsymbol{\psi} = n\boldsymbol{\psi}$. So, every vector orthogonal to the all-1s vector is an eigenvector of eigenvalue n. □

Alternative approach. Observe that $\boldsymbol{L}_{K_n} = n\boldsymbol{I} - \mathbf{1}\mathbf{1}^T$. □

To determine the eigenvalues of S_n, we first observe that each vertex $i \geq 2$ has degree 1, and that each of these degree-one vertices has the same neighbor. Whenever two degree-one vertices share the same neighbor, they provide an eigenvector of eigenvalue 1.

Lemma 2.5.2. *Let $G = (V,E)$ be a graph, and let v and w be vertices of degree one that are both connected to another vertex z. Then, the vector $\boldsymbol{\psi} = \boldsymbol{\delta}_v - \boldsymbol{\delta}_w$ is an eigenvector of $\boldsymbol{L}_G$ of eigenvalue 1.*

Proof. Just multiply $\boldsymbol{L}_G$ by $\boldsymbol{\psi}$, and check vertex-by-vertex that it equals $\boldsymbol{\psi}$. □

As eigenvectors of different eigenvalues are orthogonal, this implies that $\boldsymbol{\psi}(u) = \boldsymbol{\psi}(v)$ for every eigenvector with eigenvalue different from 1.

Lemma 2.5.3. *The graph S_n has eigenvalue 0 with multiplicity 1, eigenvalue 1 with multiplicity $n-2$, and eigenvalue n with multiplicity 1.*

Proof. Applying Lemma 2.5.2 to vertices i and $i+1$ for $2 \leq i < n$, we find $n-2$ linearly independent eigenvectors of the form $\boldsymbol{\delta}_i - \boldsymbol{\delta}_{i+1}$, all with eigenvalue 1. As 0 is also an eigenvalue, only one eigenvalue remains to be determined.

Recall that the trace of a matrix equals both the sum of its diagonal entries and the sum of its eigenvalues. We know that the trace of $\boldsymbol{L}_{S_n}$ is $2n-2$, and we have identified $n-1$ eigenvalues that sum to $n-2$. So, the remaining eigenvalue must be n.

To determine the corresponding eigenvector, recall that it must be orthogonal to the other eigenvectors we have identified. This tells us that it must have the same value at each of the points of the star. Let a be this value, and let b be the value at vertex 1. As the eigenvector is orthogonal to the constant vectors, it must be that

$$(n-1)a + b = 0.$$

It is now a simple exercise to compute the last eigenvector. □

2.6 The Hypercube

The hypercube graph is the graph with vertex set $\{0,1\}^d$, with edges between vertices whose names differ in exactly one bit. The hypercube may also be expressed as the product of the one-edge graph with itself $d-1$ times, with the proper definition of graph product.

Definition 2.6.1. *Let $G = (V, E)$ and $H = (W, F)$ be graphs. Then $G \times H$ is the graph with vertex set $V \times W$ and edge set*

$$\Big((v,w),(\hat{v},w)\Big) \text{ where } (v,\hat{v}) \in E \text{ and}$$
$$\Big((v,w),(v,\hat{w})\Big) \text{ where } (w,\hat{w}) \in F.$$

Let G be the graph with vertex set $\{0,1\}$ and one edge between those vertices. It's Laplacian matrix has eigenvalues 0 and 2. You should check that $H_1 = G$ and that $H_d = H_{d-1} \times G$.

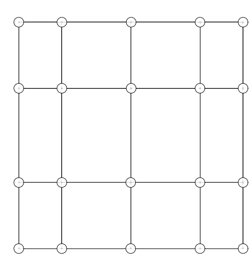

Figure 2.1: An m-by-n grid graph is the product of a path on m vertices with a path on n vertices. This is a drawing of a 5-by-4 grid made using Hall's algorithm.

Theorem 2.6.2. **??** *Let $G = (V, E)$ and $H = (W, F)$ be graphs with Laplacian eigenvalues $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_m$, and eigenvectors $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ and $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m$, respectively. Then, for each $1 \leq i \leq n$ and $1 \leq j \leq m$, $G \times H$ has an eigenvector $\boldsymbol{\gamma}_{i,j}$ of eigenvalue $\lambda_i + \mu_j$ such that*

$$\boldsymbol{\gamma}_{i,j}(v, w) = \boldsymbol{\alpha}_i(v)\boldsymbol{\beta}_j(w).$$

Proof. Let $\boldsymbol{\alpha}$ be an eigenvector of L_G of eigenvalue λ, let $\boldsymbol{\beta}$ be an eigenvector of L_H of eigenvalue μ, and let $\boldsymbol{\gamma}$ be defined as above.

To see that $\boldsymbol{\gamma}$ is an eigenvector of eigenvalue $\lambda + \mu$, we compute

$$\begin{aligned}
(L\boldsymbol{\gamma})(u, v) &= \sum_{(\hat{u},v):(u,\hat{u})\in E} (\boldsymbol{\gamma}(u, v) - \boldsymbol{\gamma}(\hat{u}, v)) + \sum_{(u,\hat{v}):(v,\hat{v})\in F} (\boldsymbol{\gamma}(u, v) - \boldsymbol{\gamma}(u, \hat{v})) \\
&= \sum_{(\hat{u},v):(u,\hat{u})\in E} (\boldsymbol{\alpha}(u)\boldsymbol{\beta}(v) - \boldsymbol{\alpha}(\hat{u})\boldsymbol{\beta}(v)) + \sum_{(u,\hat{v}):(v,\hat{v})\in F} (\boldsymbol{\alpha}(u)\boldsymbol{\beta}(v) - \boldsymbol{\alpha}(u)\boldsymbol{\beta}(\hat{v})) \\
&= \sum_{(\hat{u},v):(u,\hat{u})\in E} \boldsymbol{\beta}(v)\,(\boldsymbol{\alpha}(u) - \boldsymbol{\alpha}(\hat{u})) + \sum_{(u,\hat{v}):(v,\hat{v})\in F} \boldsymbol{\alpha}(u)\,(\boldsymbol{\beta}(v) - \boldsymbol{\beta}(\hat{v})) \\
&= \sum_{(\hat{u},v):(u,\hat{u})\in E} \boldsymbol{\beta}(v)\lambda\boldsymbol{\alpha}(u) + \sum_{(u,\hat{v}):(v,\hat{v})\in F} \boldsymbol{\alpha}(u)\mu\boldsymbol{\beta}(v) \\
&= (\lambda + \mu)(\boldsymbol{\alpha}(u)\boldsymbol{\beta}(v)).
\end{aligned}$$

□

From the fact that the non-zero eigenvalue of $\boldsymbol{L}_G$ is 2, one can conclude that H_d has eigenvalues $2i$ for $i \in \{0, 1, \ldots, d\}$, and that the eigenvalue $2i$ has multiplicity $\binom{d}{k}$. I recommend that you prove this yourself.

Using the fact that the non-zero eigenvector of $\boldsymbol{L}_G$ is $(1, -1)$, one can prove that H_d has a basis of eigenvectors of the form

$$\boldsymbol{\psi}_a(b) = (-1)^{a^T b},$$

where $a \in \{0, 1\}^d$, and we view vertices b as elements of $\{0, 1\}^d$ too. The eigenvalue of which $\boldsymbol{\psi}_a$ is an eigenvector is the number of ones in a.

Using Theorem 2.4.1 and the fact that $\lambda_2(H_d) = 2$, we can immediately prove the following isoperimetric theorem for the hypercube.

Corollary 2.6.3.

$$\theta_{H_d} \geq 1.$$

In particular, for every set of at most half the vertices of the hypercube, the number of edges on the boundary of that set is at least the number of vertices in that set.

This result is tight, as you can see by considering one face of the hypercube, such as all the vertices whose labels begin with 0. It is possible to prove this by more concrete combinatorial means. But, this proof is simpler.

2.7 Exercises

1. The eigenvalues of H_d Prove that $2i$ is an eigenvector of H_d of multiplicity $\binom{d}{i}$.

2. The eigenvectors of H_d

Prove that for each $a \in \{0,1\}^d$,

$$\boldsymbol{\psi}_a(b) = (-1)^{a^T b},$$

is an eigenvector of H_d.

References

[Hal70] K. M. Hall. An r-dimensional quadratic placement algorithm. *Management Science*, 17:219–229, 1970.