Spectral Graph Theory Lecture 2

The Laplacian

Daniel A. Spielman August 31, 2012

2.1 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 way what I wish I said.

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

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

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

so

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

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: it is the matrix that in the intersection of rows and columns indexed by u and v looks like

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

and is zero elsewhere. So,

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

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

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

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. Hall first considered the problem of assigning a real number $\boldsymbol{x}(u)$ to each vertex u so that $(\boldsymbol{x}(u) - \boldsymbol{x}(v))^2$ is small for most edges (u,v). This led him to consider the problem of minimizing (2.1). So as to avoid the degenerate solutions in which every vertex is mapped to one value, he introduced the restriction that $\boldsymbol{x}$ be orthogonal to $\mathbf{1}$. As the utility of the embedding does not really depend upon its scale, he suggested the normalization $\|\boldsymbol{x}\| = 1$. As we saw last class, the solution to the resulting optimization problem is precisely an eigenvector of the second-smallest eigenvalue of the Laplacian.

But, what if we want to assign the vertices to points in $\mathbb{R}^2$? The obvious approach is to solve for $\boldsymbol{x}$ and $\boldsymbol{y}$ minimizing

$$\min_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^V} \sum_{(u,v)\in E} \|(\boldsymbol{x}(u),\boldsymbol{y}(u)) - (\boldsymbol{x}(v),\boldsymbol{y}(v))\|^2$$

such that

$$\sum_u (\boldsymbol{x}(u),\boldsymbol{y}(u)) = (0,0).$$

However, doing so typically results in the degenerate solution $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{\psi}_2$, as

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

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 (I will probably make this 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 $s = |S| / |V|$.

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

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

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

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

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

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

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

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

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.

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

$$\sum_u \boldsymbol{\psi}(u) = 0. \tag{2.3}$$

We now compute the first coordinate of $\boldsymbol{L}_{K_n}\boldsymbol{\psi}$. We find

$$(\boldsymbol{L}_{K_n}\boldsymbol{\psi})(1) = \sum_{v \geq 2} (\boldsymbol{\psi}(1) - \boldsymbol{\psi}(v)) = (n-1)\boldsymbol{\psi}(1) - \sum_{v=2}^{n} \boldsymbol{\psi}(v) = n\boldsymbol{\psi}(1), \quad \text{by (2.3).}$$

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

Alternative approach. Observe that $L_{K_n} = nI - \mathbf{1}\mathbf{1}^T$. ☐

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 $\boldsymbol{\psi}$ given by*

$$\boldsymbol{\psi}(u) = \begin{cases} 1 & u = v \\ -1 & u = w \\ 0 & \text{otherwise,} \end{cases}$$

is an eigenvector of the Laplacian of G of eigenvalue 1.

Proof. Verify this for the path graph with three vertices, and then check that it holds in general. ☐

The existence of this eigenvector implies that $\boldsymbol{\psi}(i) = \boldsymbol{\psi}(j)$ for every eigenvector $\boldsymbol{\psi}$ of a different eigenvalue.

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 eigenvalue 1. To determine the last eigenvalue, 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. Knowing this, and the fact that the corresponding eigenvector must be constant across vertices 2 through n, make it an easy 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 $G = (\{0,1\}, \{(0,1)\})$, and let H_d be the d-dimensional hypercube graph. You should check that $H_1 = G$ and that $H_d = H_{d-1} \times G$.

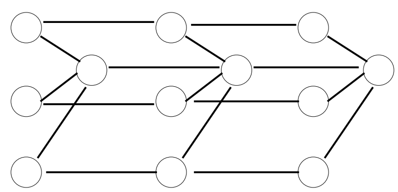

Figure 2.1: The product of a star graph on 4 vertices with a path on 3.

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 $1 \le i \le n$ and $1 \le j \le m$, $G \times H$ 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

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

□

As the non-zero eigenvector of G is $(1,-1)$ and has eigenvalue 2, we see that H_d has eigenvalue $2k$ with multiplicity $\binom{d}{k}$, for $0 \leq k \leq d$. Using the above theorem, you should also confirm that the eigenvectors of H_d are given by the functions

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

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

Using Theorem 2.4.1 and our knowledge of the eigenvalues of the hypercube, we can immediately prove the following isoperimetric theorem for the hypercube.

Corollary 2.6.3. *Let S be a subset of $\{0,1\}^d$ of size at most 2^{d-1}. Then,*

$$|\delta(S)| \geq |S|.$$

It is possible to prove this by more concrete combinatorial means. But, this proof is simpler.

References

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