CS 229r Spectral Graph Theory in Computer Science, Lecture 2 Sep 8, 2020

Lecture 2

Instructor: Salil Vadhan *Scribes: Madeleine Barowsky, Matthew Cavorsi*

1 Recap

- We will talk about weighted directed graphs (digraphs), but Spielman only talks about undirected
- We notate adjacency matrix as M, degree matrix as D, and Laplacian as $L = D - M$
  - The degree matrix D is a diagonal matrix with entries $d_{ii} = \sum_{(i,j)\in E} M_{ji} = \sum_{(i,j)\in E} w(i,j)$
  - For D of a directed graph, we can use either indegree or outdegree (both equal if the graph is Eulerian)
- **Spectral Theorem** allows us to write a real symmetric matrix M as $M = \sum_{i=1}^{n} \pi_i v_i v_i^T = V^T \Lambda V$ (where π_i are the diagonal entries of a diagonal matrix Λ and v_i are the columns of an orthogonal matrix V)

Q: For undirected Eulerian G, what can we say about the eigenvalues of L?

A: We know that $\vec{1}$ is an eigenvector of eigenvalue 0. We can say that the other eigenvalues are all non-negative (L is positive semi-definite):

$$L = \sum_{(a,b)\in E} w_{a,b} L_{G_{(a,b)}} = \sum_{(a,b)\in E} w_{a,b}(\delta_a - \delta_b)\cdot(\delta_a - \delta_b)^T,$$

$$x^T L x = \sum_{(a,b)} w_{a,b}(x_a - x_b)^2 \geq 0.$$

where $G_{(a,b)}$ is a graph with one edge, and $\delta_{a,b}$ is a vector with 1 in position {a,b} and zeros elsewhere.

2 Computational Complexity in Linear Algebra

On an $n \times n$ matrix, all of the following can be computed in $\tilde{O}(n^\omega)$ arithmetic operations (where $2 \leq \omega \leq 2.373$):

- matrix multiplication ($\star$ extremely useful!)
- matrix inversion
- determinant
- characteristic polynomial
- solving $Ax = b$
- Singular Value Decomposition (SVD) (*up to arbitrary accuracy)
- eigendecomposition of symmetric matrices (*up to arbitrary accuracy)

Note that counting arithmetic operations ignores bit length of numbers, which is important for real-world implementation. Floating point precision also requires additional work to control errors and numerical stability.

For sparse matrices, we prefer to describe complexity in terms of the number of nonzero entries (since we could potentially notate in some sparse notation with fewer than n^2 entries).

3 Looking at Eigenvalues

Q: Let $G = (V, E)$ be undirected. We know $\vec{1}$ is an eigenvector of L with eigenvalue 0. What can we say about the other eigenvalues?

A: They're all non-negative, i.e. L is positive semi-definite. Why? Observe that

$$L = \sum_{(a,b)\in E} w_{a,b} L_{G_{(a,b)}} = \sum_{(a,b)\in E} w_{(a,b)}(\delta_a - \delta_b)(\delta_a - \delta_b)^T$$

where $L_{G_{(a,b)}}$ is the Laplacian of the single-edge graph $G' = (V, \{(a,b)\})$ and each δ_i is the i-th standard basis vector. So for all $x \in \mathbb{R}^n$, we have

$$x^T L x = \sum_{(a,b)\in E} w_{a,b}(x_a - x_b)^2 \geq 0.$$

This idea of "linear decomposition of matrices" is an important theme in the course because it is similar to looking at local and global properties of graphs.

Theorem 1. *Let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of the Laplacian for a graph G. Then $\lambda_2 = 0 \iff G$ is disconnected.*

Proof of 1: ($\Leftarrow$) Suppose G is disconnected with components a_1 and a_2. Then we can write the Laplacian as

$$L = \begin{pmatrix} L_{a_1} & 0 \\ 0 & L_{a_2} \end{pmatrix}.$$

We see that $\binom{\vec{1}}{\vec{0}}$ is an eigenvector of eigenvalue 0 (since $\vec{1}$ is an eigenvector for L_{a_1} of eigenvalue 0) and similarly $\binom{\vec{0}}{\vec{1}}$ is as well. Thus $\lambda_2 = 0$.

($\Rightarrow$) Now suppose $\lambda_2 = 0$ with eigenvector $\vec{x}$. Assume for the sake of contradiction that G is connected. By definition $L\vec{x} = 0$, so we can write

$$L\vec{x} = \sum_{(a,b)\in E} w_{a,b}(x_a - x_b)^2 = 0,$$

i.e. $x_a = x_b$ for all edges (a,b) (since the weights on edges are strictly positive). By connectedness of G, every pair of vertices is connected by a path, so we can inductively conclude that $x_a = x_b$ for all $a, b \in V$. Thus $\vec{x}$ is a multiple of $\vec{1}$ and the eigenspace for 0 only has dimension 1, contradicting the fact that $\lambda_2 = 0$. So G is disconnected. □

Going forward, we will actually use λ_2 as a measure of how connected the graph is. (Larger λ_2 means G is more "highly connected".)

Q: How do we characterize λ_2 in terms of the Rayleigh quotient?

A:

$$\lambda_2 = \min_{x \perp \vec{1}} \frac{x^T L x}{x^T x}$$
$$= \min_{\substack{S \subseteq \mathbb{R}^n \\ S \text{ 2-dim subspace}}} \max_{x \in S \setminus \{0\}} \frac{x^T L x}{x^T x} \qquad \text{[Courant-Fischer Theorem].}$$

3.1 Exercise

Compute λ_2 for the complete graph K_n (with self-loops).

We have

$$M = \begin{bmatrix} 1 & 1 & 1 & \dots \\ 1 & 1 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \dots \end{bmatrix} = \mathbf{1}\mathbf{1}^T$$

with eigenvalues $\mu_1 = n$ and $\mu_2 = 0$. Then

$$L = D - M = dI - M,$$
$$\lambda_2 = d - \mu_2 = d = n.$$

4 Eigenvalues for Visual Properties of Graphs

4.1 Hall's Graph Drawing Method

Goal: Find $f : V \to \mathbb{R}^2$ (drawing of points in two dimensions) minimizing

$$\sum_{(a,b) \in E} w_{a,b} \|f(a) - f(b)\|^2 \qquad \text{(sum of weighted lengths of edges)}$$
$$= \sum_{(a,b) \in E} w_{a,b} \left\| \begin{pmatrix} x(a) \\ y(a) \end{pmatrix} - \begin{pmatrix} x(b) \\ y(b) \end{pmatrix} \right\|^2$$

such that $\|x\| = \|y\| = 1, x \perp \vec{1}, y \perp \vec{1}$, and $x \perp y$.

Optimal Solution: This minimization equals

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

The vectors x and y are restricted to be orthogonal because otherwise the optimal solution to this would be $x = y = v_2$ since we know v_2 corresponds to $\lambda_2 = \min_{x \perp \vec{1}} \frac{x^T L x}{x^T x}$. Since x and y must be orthogonal, we can say that the optimal solution would be $x = v_2$ and then $y = v_3$ which is the eigenvector of the next smallest corresponding eigenvalue that is perpendicular to both $\vec{1}$ and $x = v_2$.

4.2 Wilf's Theorem

Theorem 2 (Wilf's Theorem)**.** *Let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$ be the eigenvalues of M. Then G can be properly colored with $\lfloor \mu_1 \rfloor + 1$ colors.*

Note that this is an improvement of the solution using $d_{\max}+1$ colors from PS0. The star graph represents a graph which Wilf's Theorem improves on the d_{max} bound asymptotically, but note that the star graph is also an example that even Wilf's Theorem is not tight. It is a fact that $\mu_1 \leq d_{\max}$ and $\mu_1 = d_{\max} \iff G$ is regular.

Sketch of Proof of Wilf's Theorem: By induction on n.

(1) Find a vertex v with degree $\leq \mu_1$

(2) Color the rest of the graph with $\lfloor \mu_1 \rfloor + 1$ colors

(3) Color vertex v with any color not used by a neighbor

Proof of Steps of Sketch:
(1) *Lemma*: In every graph, we have $d_{\text{avg}} \leq \mu_1$.

Proof. We know $\mu_1 = \max_x \frac{x^T M x}{x^T x} \geq \frac{1^T M 1}{1^T 1} = d_{\text{avg}}$. $\square$

(2) By the inductive hypothesis and
Lemma: $G \setminus \{v\}$ has largest eigenvalue $\leq \mu_1$.

Proof. Write $\mathbb{R}^V$ for the $\|V\|$-dimensional space indexed by the vertices in V. Let $v \in V$. Without loss of generality suppose v is indexed as the last component.

$$\begin{aligned}
\mu_1 = \max_{x \in \mathbb{R}^n} \frac{x^T M x}{x^T x} &\geq \max_{y \in \mathbb{R}^{V-\{v\}}} \frac{\binom{y}{0}^T M \binom{y}{0}}{\binom{y}{0}^T \binom{y}{0}} \\
&= \max_{y \in \mathbb{R}^{V-\{v\}}} \frac{y^T M_{G-\{v\}} y}{y^T y}.
\end{aligned}$$

$\square$

(3) TBD $\square$

4.3 Cauchy's Interlacing Theorem

Theorem 3. *If A is a symmetric $n \times n$ matrix with eigenvalues $\alpha_1 \geq \cdots \geq \alpha_n$ and B is a principal $(n-1) \times (n-1)$ submatrix of A (B is obtained by deleting the same row and column from A) with eigenvalues $\beta_1 \geq \cdots \geq \beta_{n-1}$, then $\alpha_1 \geq \beta_1 \geq \alpha_2 \geq \beta_2 \geq \cdots \geq \beta_{n-1} \geq \alpha_n$.*

Cauchy's Interlacing Theorem is a generalization of the proof of step (2) from above.