Advanced Algorithms

Lecture 19: Spectral Graph Theory

Lecturer: Nathan Klein

1 Spectral Graph Theory

A beautiful and surprising area of computer science is studying the relationship between the eigenvalues of the adjacency matrix of a graph (or, more often, a very similar object called its *Laplacian*) and properties of the graph itself.

1.1 The Laplacian

The Laplacian of a graph $G = (V, E)$ is a matrix L in $\mathbb{R}^{n \times n}$ (for $|V| = n$) so that $L = D - A$, where D is the diagonal matrix of degrees so that D_{ii} is the degree of the ith vertex and all off-diagonal entries are 0. For example:

$$L = D - A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Fact 1.1. *For any graph G, its Laplacian L is positive semi-definite. Furthermore, we have*

$$x^T L x = \sum_{\{v_i, v_j\} \in E} (x_i - x_j)^2$$

Proof. Let L_e be the Laplacian of the edge e between vertices v_i and v_j, i.e., the Laplacian of the graph $G' = (V, \{e\})$. This is going to be a matrix of the form $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ (with some extra 0s at locations without an index i or j). But this matrix is equal to $b_e b_e^T$, where b_e is the vector with a 1 at position i and a -1 at position j. So, L_e is PSD. Now, notice that $L = \sum_{e \in E} L_e$. So,

$$x^T L x = x^T (\sum_{e \in E} L_e) x = \sum_{e \in E} x^T L_e x \geq 0$$

since each L_e is PSD, showing that L itself is PSD. Even better, for an edge $e = \{v_i, v_j\}$ we can compute

$$x^T L_e x = x_i^2 - 2x_i x_j + x_j^2 = (x_i - x_j)^2$$

so that $x^T L x = \sum_{\{v_i, v_j\} \in E} (x_i - x_j)^2$ as desired. $\square$

From this we know that all the eigenvalues are non-negative. We also get the following corollary which frees us from this annoying $\mathbf{1}_S^T A \mathbf{1}_{V \setminus S}$ notation and is one of the reasons the Laplacian can be more natural to work with than the adjacency matrix.

Corollary 1.2. *Let $S \subseteq V$. Then, $\mathbf{1}_S^T L \mathbf{1}_S = |\delta(S)|$.*

Proof.

$$\mathbf{1}_S^T L \mathbf{1}_S = \sum_{\{v_i, v_j\} \in E} (\mathbb{I}\{v_i \in S\} - \mathbb{I}\{v_j \in S\})^2 = |\delta(S)|$$

since this quantity is 1 if and only if exactly one of v_i, v_j is in S. □

For vectors $x \notin \{0,1\}^n$, this then gives us a kind of "fractional" cut.

Fact 1.3. $\lambda_1 = 0$ *for any Laplacian matrix L.*

Proof. $\mathbf{1}^T L \mathbf{1} = 0$ because the rows sum to 0, so $L\mathbf{1}$ is the 0 vector. So, $\lambda_1 = 0$. □

So, the first eigenvalue is always 0. It turns out the second eigenvalue already tells us something interesting. Before we say this, let's prove a more general form of the Rayleigh quotient.

1.2 More on Eigenvalues

Theorem 1.4. *Given a symmetric matrix $A \in \mathbb{R}^{n \times n}$ with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ with orthonormal eigenvectors $v_1, \ldots, v_n$. Let O_k be the set of non-zero vectors that are orthogonal to the first $k-1$ eigenvectors $v_1, \ldots, v_{k-1}$ (let O_1 be all non-zero vectors). Then for any $1 \leq k \leq n$:*

$$\lambda_k = \min_{x \in O_k} \frac{x^T A x}{x^T x}$$

Proof. The proof of this is similar to the fact that $\lambda_1 = \min_{x \neq 0 \in \mathbb{R}^n} \frac{x^T A x}{x^T x}$, which you showed on your homework. In case you forget how that works, let's show that again. First we apply the spectral theorem so that $A = \sum_{i=1}^n \lambda_i v_i v_i^T$ where $v_1, \ldots, v_n$ are an orthogonal basis of $\mathbb{R}^n$. So, $x = c_1 v_1 + \cdots + c_n v_n$ where $c_i = \langle x, v_i \rangle$. But now:

$$\begin{aligned} x^T A x &= (c_1 v_1 + \cdots + c_n v_n)^T (\sum_{i=1}^n \lambda_i v_i v_i^T)(c_1 v_1 + \cdots + c_n v_n) \\ &= (c_1 v_1 + \cdots + c_n v_n)(\lambda_1 c_1 v_1 + \cdots + \lambda_n c_n v_n) && \text{Since } v_i \text{ are orthonormal} \\ &= \sum_{i=1}^n \lambda_i c_i^2 \end{aligned}$$

Also, we have $x^T x = (c_1 v_1 + \cdots + c_n v_n)^T (c_1 v_1 + \cdots + c_n v_n) = \sum_{i=1}^n c_i^2$. Therefore:

$$\frac{x^T A x}{x^T x} = \frac{\sum_{i=1}^n \lambda_i c_i^2}{\sum_{i=1}^n c_i^2}$$

In other words, $\frac{x^T A x}{x^T x}$ is a *convex combination* of the eigenvalues $\lambda_1, \ldots, \lambda_n$. So, it is certainly at least λ_1, and we can pick the vector v_1 to achieve $c_1 = 1$, $c_i = 0$ for $i > 1$.

The more general proof is not very different. We know that the kth eigenvector is orthogonal to the others. So we have $x \in O_k$, and we must have $c_1, \ldots, c_{k-1} = 0$ since recall $c_i = \langle x, v_i \rangle$. So $x \in O_k$ is a convex combination of $\lambda_k, \ldots, \lambda_n$ so to minimize it we should choose $x = v_k$. □

There is also a cleaner way to view things which is not tied to the eigenvectors.

Theorem 1.5 (Courant-Fischer)**.** *Let A be a symmetric matrix. Then:*

$$\lambda_k = \min_{S \subseteq \mathbb{R}^n : \dim(S) = k} \max_{x \in S} \frac{x^T A x}{x^T x}$$

This is quite similar to what we have already done, so we will leave the proof as an exercise.

1.3 λ_2 and Connectivity

Now that we understand eigenvalues a bit better, we can prove the following:

Fact 1.6. *G is connected if and only if $\lambda_2 > 0$.*

Proof. First suppose G is disconnected. Then, there is a set of vertices $\emptyset \neq S \neq V$ so that $|\delta(S)| = 0$. So, by Corollary 1.2, $\mathbf{1}_S^T L \mathbf{1}_S = 0$, and $\mathbf{1}_S \neq \mathbf{0}$ which also implies $L\mathbf{1}_S = 0 \cdot \mathbf{1}_S$. But then we have two eigenvectors with eigenvalue 0, $\mathbf{1}$ and $\mathbf{1}_S$. Furthermore, they are not linearly dependent. So by Courant-Fischer, we can pick S as the 2-dimensional space spanned by $\mathbf{1}$ and $\mathbf{1}_S$ and λ_1, λ_2 are both 0, since $x^T L x$ will be 0 for any vector in the span of $\mathbf{1}$ and $\mathbf{1}_S$.

Second, suppose G is connected. Suppose that $\lambda_2 = 0$. Then there is a 2-dimensional space S with $x^T L x = 0$ for all $x \in S$. There is therefore some $x \in S$ which is not a multiple of the all 1s vector for which $x^T L x = 0$. Then,

$$\sum_{\{v_i, v_j\} \in E} (x_i - x_j)^2 = x^T L x = 0$$

That implies that $x_i = x_j$ for all i, j. But then x would be a multiple of v_1, contradiction. $\square$

You might say now: who cares? We already know how to figure out if a graph is connected with much simpler methods. The reason we care is this: λ_2 of L is actually a measure of *how connected the graph is*. We'll talk more about this next time.