ADVANCED ALGORITHMS (V)

CHIHAO ZHANG

1. Basics Spectral Tools

Today we start to introduce *spectral algorithms*. The theory uses eigenvalues and eigenvectors of adjacency matrices to study combinatorial proeprties of graphs.

Recall that we have the spectral decomposition theorem, which is the main tool we are going to use today.

Theorem 1 (Spectral Decomposition Theorem)**.** *An $n \times n$ symmetric matrix A has n real eigenvalues $\lambda_1, \dots, \lambda_n$ with corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ which are orthonomal. Moreover, it holds that*

$$A = V \Lambda V^T,$$

where $V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix}$ *and* $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

Given a graph $G = (V, E)$ with $V = \{v_1, v_2, \dots, v_n\}$, the basic object we study is its adjacency matrix $A = (a_{i,j})_{1 \le i,j \le n}$. To ease the notation, we shift A to a positive semi-definite matrix. Let

$$L = D - A$$

be the *Laplacian* of G where $D = \mathrm{diag}\left(\deg(v_1), \deg(v_2), \dots, \deg(v_n)\right)$.

Proposition 2. *For very $\mathbf{x} \in \mathbb{R}^n$, it holds that*

$$\mathbf{x}^T L \mathbf{x} = \sum_{\{i,j\} \in E} \left(x(i) - x(j)\right)^2.$$

Proof. Direct calculation shows that both sides are equal to $\sum_{i \in V} \deg(i) \cdot x(i)^2 - 2 \sum_{\{i,j\} \in E} x(i)x(j)$. □

Therefore, $L \succeq 0$ and all of its eigenvalues are nonnegative.

Sometimes it is more convenient to work with *normalized Laplacian*, which is defined to be

$$N \triangleq D^{-\frac{1}{2}} L D^{-\frac{1}{2}} = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}.$$

When G is d-regular, we have $N = I - \frac{1}{d} A$.

The *Rayleigh Quotient* of a vector $\mathbf{x} \in \mathbb{R}^n$ with respect to a matrix M is

$$R_M(\mathbf{x}) = \frac{\langle \mathbf{x}, M\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

Note that for the normalizaed Laplacian N of some graph, it holds that

$$R_N(\mathbf{x}) = \frac{\langle \mathbf{x}, (D^{-\frac{1}{2}} L D^{-\frac{1}{2}})\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} = \frac{\langle D^{-\frac{1}{2}}\mathbf{x}, L D^{-\frac{1}{2}}\mathbf{x} \rangle}{\langle D^{-\frac{1}{2}}\mathbf{x}, D D^{-\frac{1}{2}}\mathbf{x} \rangle} = \frac{\langle \mathbf{y}, L\mathbf{y} \rangle}{\langle \mathbf{y}, D\mathbf{y} \rangle}, \tag{1}$$

where $\mathbf{y} \triangleq D^{-\frac{1}{2}} \mathbf{x}$.

The following *variational characterization* of the eigenvalues is our center tool in the development of the theory.

Theorem 3. *Let M ba a symmetrix matrix with real eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$.Then for every $k = 1, \dots, n$,*

$$\lambda_k = \min_{k\text{-dim subspace } X \subseteq \mathbb{R}^n} \max_{\mathbf{x} \in X \setminus \{0\}} R_M(\mathbf{x}).$$

Proof. Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a set of *orthonormal* eigenvectors corresponding to $\lambda_1, \dots, \lambda_n$ respectively.

We first prove

$$\min_{k\text{-dim subspace } X \subseteq \mathbb{R}^n} \max_{\mathbf{x} \in X \setminus \{0\}} R_M(\mathbf{x}) \le \lambda_k.$$

To this end, we only need to construct a k-dimensional space X such that every nonzero vector $\mathbf{x} \in X \setminus \{0\}$ satisfies $R_M(\mathbf{x}) \le \lambda_k$. Let $X = \mathrm{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$, then for every $\mathbf{x} = \sum_{i=1}^k a_i \mathbf{v}_i \in X \setminus \{0\}$, we have

$$R_M(\mathbf{x}) = \frac{\langle \sum_{i=1}^k a_i \mathbf{v}_i, M\left(\sum_{i=1}^k a_i \mathbf{v}_i\right)\rangle}{\langle \sum_{i=1}^k a_i \mathbf{v}_i, \sum_{i=1}^k a_i \mathbf{v}_i\rangle} = \frac{\langle \sum_{i=1}^k a_i \mathbf{v}_i, \sum_{i=1}^k \lambda_i a_i \mathbf{v}_i\rangle}{\langle \sum_{i=1}^k a_i \mathbf{v}_i, \sum_{i=1}^k a_i \mathbf{v}_i\rangle} = \frac{\sum_{i=1}^k \lambda_i a_i^2}{\sum_{i=1}^k a_i^2} \le \lambda_k.$$

We then prove the other direction, i.e.,

$$\min_{k\text{-dim subspace } X \subseteq \mathbb{R}^n} \max_{\mathbf{x} \in X \setminus \{0\}} R_M(\mathbf{x}) \ge \lambda_k.$$

It is sufficient to show that, for every k-dimensional subspace $X \subseteq \mathbb{R}^N$, there exists some $\mathbf{x} \in X \setminus \{0\}$ satisfying $R_x(\mathbf{x}) \ge \lambda_k$. Let us fix such a k-dimensional subspace, and let $\mathbf{x}$ be an arbitrary nonzero vector in

$$X \cap \mathrm{span}(\mathbf{v}_k, \dots, \mathbf{v}_{n-k+1}).$$

Such a vector $\mathbf{x}$ must exist since the sum of the dimensions of the two spaces is $n+1$. Assuming $\mathbf{x} = \sum_{i=k}^n b_i \mathbf{v}_i$, then

$$R_M(\mathbf{x}) = \frac{\langle \sum_{i=k}^n b_i \mathbf{v}_i, M\left(\sum_{i=k}^n b_i \mathbf{v}_i\right)\rangle}{\langle \sum_{i=k}^n b_i \mathbf{v}_i, \sum_{i=k}^n b_i \mathbf{v}_i\rangle} = \frac{\langle \sum_{i=k}^n b_i \mathbf{v}_i, \sum_{i=k}^n \lambda_i b_i \mathbf{v}_i\rangle}{\langle \sum_{i=k}^n b_i \mathbf{v}_i, \sum_{i=k}^n b_i \mathbf{v}_i\rangle} = \frac{\sum_{i=k}^n \lambda_i b_i^2}{\sum_{i=k}^n b_i^2} \ge \lambda_k.$$

□

An immediate corollary of above is

Corollary 4.

$$\lambda_1 = \min_{\mathbf{x} \in \mathbb{R}^n \setminus \{0\}} R_M(\mathbf{x}), \quad \lambda_n = \max_{\mathbf{x} \in \mathbb{R}^n \setminus \{0\}} R_M(\mathbf{x}).$$

Besides, we can similarly use the spectral decomposition of vectors to prove another useful characterization of eigenvectors:

Theorem 5. *Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a set of* orthonormal *(with respect to $\langle\cdot,\cdot\rangle$) eigenvectors corresponding to $\lambda_1, \dots, \lambda_n$ respectively. Then for every $k = 1, \dots, n$,*

$$\lambda_k = \min_{\mathbf{x} \perp \mathrm{span}(\mathbf{v}_1, \dots, \mathbf{v}_{k-1})} R_M(\mathbf{x}).$$

The following proposition sheds some light on the relation between eigenvalues and the structure of the graph.

Proposition 6. *Let $G = (V, E)$ be an undirected graph and N be its normalized Laplacian. Assuming $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$ are its eigenvalues, then*

(1) *$\lambda_1 = 0$;*
(2) *$\lambda_n \le 2$ and $\lambda_n = 2$ if and only if one of components of G is bipartite;*
(3) *$\lambda_k = 0$ if and only if G has at least k components.*

Proof. It follows from (1) and Corollary (4) that

$$\lambda_1 = \min_{\mathbf{x} \neq 0} R_N(\mathbf{x}) = \min_{\mathbf{y} \neq 0} \frac{\langle \mathbf{y}, L\mathbf{y}\rangle}{\langle \mathbf{y}, D\mathbf{y}\rangle} = \min_{\mathbf{y} \neq 0} \frac{\sum_{\{i,j\} \in E} (y_i - y_j)^2}{\sum_{i \in V} \deg(i) \cdot y_i^2} \ge 0,$$

where the second equality is because the mapping $\mathbf{x} \to D^{-\frac{1}{2}}\mathbf{x} =: \mathbf{y}$ is linear and bijective. The equality is achieved when $\mathbf{y} = \mathbf{1}$. This proves (1).

Similarly, we have

$$\lambda_n = \max_{\mathbf{x} \neq 0} R_N(\mathbf{x}) = \max_{\mathbf{y} \neq 0} \frac{\langle \mathbf{y}, L\mathbf{y}\rangle}{\langle \mathbf{y}, D\mathbf{y}\rangle} = 2 - \min_{\mathbf{y} \neq 0} \frac{\sum_{\{i,j\} \in E} (y_i + y_j)^2}{\sum_{i \in V} \deg(i) \cdot y_i^2} \le 2.$$

To achieve the equality, we need $y_i = -y_j$ whenever $\{i, j\} \in E$. It is not hard to verify that, such a nonzero $\mathbf{y}$ exists if and only if one of components of G is bipartite.

To show (3), we use

$$\lambda_k = \min_{k\text{-dim subspace } X \subseteq \mathbb{R}^n} \max_{\mathbf{x} \in X} R_N(\mathbf{x}) = \min_{k\text{-dim subspace } Y \subseteq \mathbb{R}^n} \max_{\mathbf{y} \in Y} \frac{\sum_{\{i,j\} \in E} (y_i - y_j)^2}{\sum_{i \in V} \deg(i) \cdot y_i^2}.$$

First assume $\lambda_k = 0$, then we know for some k-dim subspace Y, every vector $\mathbf{y} \in Y$ satisfies $y_i = y_j$ whenever $\{i, j\} \in E$. This implies that for every $\mathbf{y} \in Y$, $y_i = y_j$ whenever i and j are in the same component. Let $\mathcal{C}$ denote the set of components and for every $C \in \mathcal{C}$, we use $\mathbf{1}_C$ to denote the vector such that $\mathbf{1}_C(i) = \begin{cases} 1, & i \in C; \\ 0, & i \notin C \end{cases}$. Then $\mathrm{span}(\{\mathbf{1}_C\}_{C\in\mathcal{C}})$ must contain the whole space Y and therefore $|\mathcal{C}| \geq k$. Conversely, if $|\mathcal{C}| \geq k$, the space $\mathrm{span}(\{\mathbf{1}_C\}_{C\in\mathcal{C}})$ witnesses $\lambda_k = 0$. □

2. Edge Expansion and Cheeger's Inequality

We want to obtai n some *quantitative* version of Proposition 6. To this end, we introduce the notion of edge expansion to measure the connectivity of a graph.

Let $G = (V, E)$ be a graph, for every $S \subseteq V$, we define

$$\phi(S) = \frac{\left|E(S, \bar{S})\right|}{\sum_{i\in S} \deg(i)},$$

where $E(S, \bar{S})$ is the set of edges crossing S and $\bar{S}$. When the graph is d-regular, $\phi(S)$ becomes to $\frac{E(S,\bar{S})}{d|S|}$. It is clear that the quantity is in fact the fraction of edges *going out of* S among those edges incident to S. The (edge) expansion of the graph G is the minimum of $\phi(S)$, over all $S \subseteq V$ with $|S| \leq \frac{V}{2}$:

$$\phi(G) \triangleq \min_{S\subseteq V, |S|\leq \frac{V}{2}} \phi(S).$$

The edge expansion is an important object that appears in many areas of computer science. For example, we will show in later lectures that it is closely related to the random walk on the graph. Today, we will introduce an imporatnt result that relating $\phi(G)$ with eigenvalues of the normalized Laplacian.

Theorem 7 (Cheeger's Inequality). *Let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of the normalized Laplacian of G, then*

$$\frac{\lambda_2}{2} \leq \phi(G) \leq \sqrt{2\lambda_2}.$$