[CS1961: Lecture 13] Spectral Graph Theory

Instructor: Chihao Zhang;
Scribed by Zidong Yang, Tianyu Zhang, Yuchen He

1 Graph Adjacency Matrix and Its Spectrum

Given an undirected graph $G = (V, E)$ where $V = [n]$, let $A_G = (a_{ij})_{i,j\in[n]}$ be the adjacent matrix of G. That is, A_G is a binary matrix with $a_{ij} = 1$ iff $(i, j) \in E$. Note that A_G is symmetric. Therefore the n eigenvalues of A_G, $\lambda_1, \lambda_2, \ldots, \lambda_n$, are all real numbers. W.l.o.g, we assume $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. This is called the spectrum of G. For example, when G is the complete graph K_n,

$$A_{K_n} = \begin{bmatrix} 0 & 1 & \cdots & \\ 1 & 0 & & \\ \vdots & & \ddots & 1 \\ & & 1 & 0 \end{bmatrix}.$$

The n eigenvalues of A_{K_n} is $\lambda_1 = n - 1$ and $\lambda_2 = \cdots = \lambda_n = -1$. The corresponding eigenvectors are

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \mathbf{v}_n = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ -1 \end{bmatrix}.$$

When the graph G is given, its spectrum is also determined. However, the same spectrum may corresponds to different graphs. The properties of the spectrum usually reflect properties of the graph and have been extensively studied.

Proposition 1 *If the maximium degree of a graph G is Δ, then $\lambda_1 \leqslant \Delta$. In particular, if G is Δ-regular, then $\lambda_1 = \Delta$.*

Proof. Let $\delta = \max_{i\in[n]} \sum_{j=1}^{n} |a_{ij}|$ be the maximum absolute row sum of A. We claim that $\|A\|_\infty = \delta$.[1] Here A is not necessarily a binary matrix.

Choosing $\mathbf{x} = \mathbf{v}_1$, we have

$$|\lambda_1|\|\mathbf{v}_1\|_\infty = \|A_G\mathbf{v}_1\|_\infty \leqslant \|A_G\|_\infty\|\mathbf{v}_1\|_\infty$$

by definition. If the claim holds, we can further yield that $|\lambda_1| \leqslant \|A_G\|_\infty = \delta = \Delta$. When G is Δ-regular, it is easy to verify that $\mathbf{1}$ is an eigenvector corresponding to the eigenvalue Δ. Therefore, we have $\lambda_1 = \Delta$.

It remains to prove the claim. We write A as $A = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n]$. Then $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$. W.l.o.g, assume that $\arg\max_i \sum_i |a_{ij}| = 1$,

[1] The p-norm of a vector $\mathbf{x}$ is defined as $\|\mathbf{x}\|_p = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}}$. Specifically, when $p = \infty$, $\|\mathbf{x}\|_\infty = \max_i |x_i|$. The p-norm of a matrix A is defined as $\|A\|_p = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p}$, which measures the size of the operator A.

i.e., the first row of A has the maximum absolute row sum. Note that $\frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty}$ reaches the peak when $\mathbf{x} \in \{-1, 1\}^n$ and each $x_i = \mathrm{sgn}(a_{1i})$. This naturally yields that $\|A\|_\infty = \delta$. $\square$

The result shows that λ_1 is related to the degree of the graph. We will see how the other eigenvalues reflect graph properties.

2 Rayleigh Quotient

Given $A \in \mathbb{R}^{n\times n}$ and $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, the Rayleigh quotient is defined as $R_A(\mathbf{x}) \triangleq \frac{\langle \mathbf{x}, A\mathbf{x}\rangle}{\langle \mathbf{x},\mathbf{x}\rangle}$. By the spectral decomposition theorem, A can be written as $\sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^T$ where $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a group of orthonormal eigenvectors of A corresponding to eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ respectively.

Unless otherwise stated, we assume A is symmetric.

We can write $\mathbf{x}$ as $\sum_{i=1}^n a_i \mathbf{v}_i$ for some constants $a_1, a_2 \dots, a_n$. Then

$$\langle \mathbf{x}, \mathbf{x}\rangle = \langle \sum_{i=1}^n a_i \mathbf{v}_i, \sum_{j=1}^n a_j \mathbf{v}_j\rangle = \sum_{i,j\in[n]} a_i a_j \langle \mathbf{v}_i, \mathbf{v}_j\rangle = \sum_{i=1}^n a_i^2$$

and

$$A\mathbf{x} = \left(\sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^T\right)\left(\sum_{j=1}^n a_j \mathbf{v}_j\right) = \sum_{i,j\in[n]} \lambda_i a_j \mathbf{v}_i \langle \mathbf{v}_i, \mathbf{v}_j\rangle = \sum_{i=1}^n \lambda_i a_i \mathbf{v}_i.$$

Similarly,

$$\langle \mathbf{x}, A\mathbf{x}\rangle = \langle \sum_{j=1}^n a_j \mathbf{v}_j, \sum_{i=1}^n \lambda_i a_i \mathbf{v}_i\rangle = \sum_{i=1}^n \lambda_i a_i^2.$$

Therefore, $R_A(\mathbf{x}) = \frac{\langle \mathbf{x}, A\mathbf{x}\rangle}{\langle \mathbf{x},\mathbf{x}\rangle} = \frac{\sum_{i=1}^n \lambda_i a_i^2}{\sum_{i=1}^n a_i^2}$. With this form of Rayleigh quotient, we can introduce the Courant-Ficher theorem, which gives a variational characterization of the eigenvalues.

Claim 2 $\lambda_1 = \max_{\mathbf{x}\neq\mathbf{0}} R_A(\mathbf{x})$

Proof. Since $R_A(\mathbf{x}) = \frac{\sum_{i=1}^n \lambda_i a_i^2}{\sum_{i=1}^n a_i^2} = \sum_{i=1}^n \frac{a_i^2}{\sum_{j=1}^n a_j^2}\lambda_i$ achieves the maximum when the weight concentrates on λ_1, we have $\max_{\mathbf{x}\neq\mathbf{0}} R_A(\mathbf{x}) = R_A(\mathbf{v}_1) = \lambda_1$. $\square$

With the same argument, we have $\lambda_2 = \max_{\mathbf{x}\neq\mathbf{0},\mathbf{x}\perp\mathbf{v}_1} R_A(\mathbf{x})$. This can be generalized to the k-th largest eigenvalue:

$$\lambda_k = \max_{\substack{\mathbf{x}\neq\mathbf{0} \\ \mathbf{x}\perp\mathrm{span}(\mathbf{v}_1,\dots,\mathbf{v}_{k-1})}} R_A(\mathbf{x}).$$

We also have

$$\lambda_k = \max_{\substack{V\subseteq\mathbb{R}^n \\ \dim(V)=k}} \min_{\mathbf{x}\in V\setminus\{\mathbf{0}\}} R_A(\mathbf{x}). \tag{1}$$

Equation (1) can be interpreted as the competition between the max player and min player. The best choice of the max player is to set $V = \mathrm{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ and the min player will choose $\mathbf{x} = \mathbf{v}_k$ to minimize $R_A(\mathbf{x})$.

Proposition 3 *For a simple* d*-regular graph* $G = (V, E)$*,* G *is connected iff* $\lambda_2 \neq d$.

Proof. Recall that $\mathbf{v}_1 = \frac{1}{\sqrt{n}} \cdot \mathbf{1}$ for d-regular graphs. Then by the Courant-Ficher theorem, $\lambda_2 = \max_{\mathbf{x}\neq\mathbf{0},\mathbf{x}\perp\mathbf{1}} R_{A_G}(\mathbf{x})$. Note that

$$R_{A_G}(\mathbf{x}) = \frac{\sum_{(i,j)\in E} 2x_i x_j}{\sum_{i=1}^n x_i^2} = d - \frac{d\sum_{i=1}^n x_i^2 - \sum_{(i,j)\in E} 2x_i x_j}{\sum_{i=1}^n x_i^2} = d - \frac{\sum_{(i,j)\in E}(x_i - x_j)^2}{\sum_{i=1}^n x_i^2}.$$

Therefore $\lambda_2 = d$ iff $(x_i - x_j)^2 = 0$ for all $(i, j) \in E$. Since $\mathbf{x} \perp \mathbf{1}$, $\sum_{(i,j)\in E}(x_i - x_j)^2 = 0$ indicates that G is not connected. □

Proposition 4 *Suppose* $G = (V, E)$ *is a simple* d*-regular graph which is connected. Then* G *is bipartite iff* $\lambda_n = -d$.

Proof. By the Courant-Ficher theorem,

$$\lambda_n = \min_{\mathbf{x}\neq\mathbf{0}} R_{A_G}(\mathbf{x}) = \min_{\mathbf{x}\neq\mathbf{0}} \frac{\sum_{i,j\in[n]} a_{ij}x_i x_j}{\sum_{i=1}^n x_i^2}.$$

Note that

$$\frac{\sum_{i,j\in[n]} a_{ij}x_i x_j}{\sum_{i=1}^n x_i^2} = \frac{\sum_{(i,j)\in E} 2x_i x_j}{\sum_{i=1}^n x_i^2} + d - d = \frac{\sum_{(i,j)\in E}(x_i + x_j)^2}{\sum_{i=1}^n x_i^2} - d.$$

Therefore $\lambda_n = -d$ iff $x_i = -x_j$ for all $(i, j) \in E$. This indicates that G is bipartite. □

3 Cauchy Interlacing Theorem

Theorem 5 (Cauchy Interlacing Theorem) *Let* $A \in \mathbb{R}^{n\times n}$ *be a symmetric matrix with eigenvalues* $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$*. Let* $B \in \mathbb{R}^{m\times m}$ *be a principal submatrix*[2] *of* A *with eigenvalues* $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_m$*. Then for all* $k \in m$*,* $\lambda_k \geqslant \mu_k \geqslant \lambda_{k+n-m}$.

[2] A principal submatrix is a square submatrix obtained by removing certain rows and columns. The indices of removed rows are the same with removed columns.

Proof. It is sufficient to prove the case when $m = n - 1$. W.l.o.g, assume B is generated by deleting the first row and first column in A. By the Courant-Ficher theorem, we have

$$\lambda_k = \max_{\substack{V\subseteq\mathbb{R}^n \\ \dim(V)=k}} \min_{\mathbf{x}\in V\setminus\{\mathbf{0}\}} R_A(\mathbf{x}) \quad \text{and} \quad \mu_k = \max_{\substack{U\subseteq\mathbb{R}^{n-1} \\ \dim(U)=k}} \min_{\mathbf{y}\in U\setminus\{\mathbf{0}\}} R_B(\mathbf{y}).$$

For any $\mathbf{y} \in \mathbb{R}^{n-1}$, let $\mathbf{y}' = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}$. Then $\mathbf{y}' \in \mathbb{R}^n$ and $\mathbf{y}'$ satisfies that $\langle \mathbf{y}, B\mathbf{y}\rangle = \langle \mathbf{y}', A\mathbf{y}'\rangle$ and $\langle \mathbf{y}, \mathbf{y}\rangle = \langle \mathbf{y}', \mathbf{y}'\rangle$. Therefore,

$$\mu_k = \max_{\substack{U\subseteq\mathbb{R}^{n-1} \\ \dim(U)=k}} \min_{\mathbf{y}\in U\setminus\{\mathbf{0}\}} R_A(\mathbf{y}').$$

This indicates that $\mu_k \leqslant \lambda_k$.

For the proof of $\mu_k \geqslant \lambda_{k+n-m}$, consider $-A$ and $-B$. Then the spectrum of $-A$ is $-\lambda_n \geqslant -\lambda_{n-1} \geqslant \cdots \geqslant -\lambda_1$ and the spectrum of $-B$ is $-\mu_{n-1} \geqslant -\mu_{n-2} \geqslant \cdots \geqslant -\mu_1$. With the same argument, we can verify that $-\mu_k \leqslant -\lambda_{k+1}$.

□

Let $n^+(A)$ be the number of positive eigenvalues of A and $n^-(A)$ be the number of negative eigenvalues. Let $\alpha(G)$ be the independent number of G. We can derive an upper bound of $\alpha(G)$ using the Cauchy interlacing theorem.

Theorem 6 (Cvetkovic Theorem) $\alpha(G) \leqslant \min\{n - n^+(A), n - n^-(A)\}$.

Proof. Let $S \subseteq V$ be an independent set with size $\alpha(G)$. Let B be the principal submatrix of A indexed by S. Then B must be a zero matrix with each eigenvalue $\mu_k = 0$ for $k \in [\alpha(G)]$.

For $k \in [\alpha(G)]$, by the Cauchy interlacing theorem, $\lambda_k \geqslant \mu_k \geqslant \lambda_{k+n-\alpha(G)}$. Therefore we have

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{\alpha(G)} \geqslant 0.$$

This indicates $n^-(A) \leqslant n - \alpha(G)$. Apply the same argument on $-A$ and $-B$, we can similarly yield that $n^+(A) \leqslant n - \alpha(G)$.

□