[CS3958: Lecture 14] Graph Expansion(Cont'), Cheeger's Inequality

Instructor: Chihao Zhang, Scribed by Yulin Wang

January 1, 2023

1 Graph Expansion (Cont')

Review: Different views of analyzing mixing time/rate of convergence of Markov Chains.

- Probabilistic view ~ Coupling;
- Algebraic view ~ Spectrum; (Algebraic Graph Theory)
- Geometric view ~ Expansion. (KLS Conjecture)

Expansion can be defined on any weighted undirected graph, not only for Markov chains: Let $G = (V, E)$ be a weighted undirected graph with a weight function $w(i, j) > 0$ for each edge $\{i, j\} \in E$. Then we define the expansion as

$$\Phi(S, \bar{S}) = \frac{\sum_{i\in S, j\in \bar{S}} w(i, j)}{\sum_{i,j\in S} w(i, j)}.$$

Note this definition is consistent with that of Markov chains. If we let $P(i, j) = \frac{w(i,j)}{\sum_j w(i,j)}$ i.e. P is a natural random walk on G, then we have $\pi(i) \sim \sum_j w(i, j)$ and $\pi(i)P(i, j) = \pi(j)P(j, i)$. Therefore, the Markov chain P can imply some results on the expansion of G.

1.1 Applications for Sampling Colorings

Assume we want to sample from all proper $[q]$-colorings on $G = ([n], E)$ with maximum degree Δ. The Markov chain is

- Pick $v \in [n]$ and $c \in [q]$ uniformly at random.
- Recolor v with c if possible.

Recall that we proved $\tau_{mix}(\varepsilon) \le qn \log \frac{n}{\varepsilon}$ when $q > 4\Delta$. Now we want to argue that when q is rather small, the expansion is large for some special graph. Consider the case when G is a star and 1 is the vertex at the center. Let Z be the number of all proper colorings on G, and S be the set of proper colorings that the color of 1 is 1. Then we have

If we want to upperbound $\Phi(P)$, we need to argue that for any S such that $\pi(S) \le \frac{1}{2}$, $\Phi(S, \bar{S})$ has an upper bound.

The problem to find a cut $S, \bar{S}$ with maximum expansion is dual with multi-commodity flow problem. If you are interested in this topic, search for "canonical paths" or "multi-commodity flow".

$$\begin{aligned} Q(S, \bar{S}) &= \sum_{i\in S, j\in \bar{S}} \pi(i)P(i, j) \\ &= (q-1)(q-2)^{n-1}\frac{1}{Z\cdot nq}. \end{aligned}$$

Since $|S| = (q-1)^{n-1}$, we have

$$\Phi(S) = \frac{Q(S, \bar{S})}{\pi(S)} = \frac{(q-2)^{n-1}}{(q-1)^{n-2}}\frac{1}{nq} = \frac{q-1}{nq}\left(1 - \frac{1}{q-1}\right)^{n-2} \le \frac{1}{n}\exp\left(-\frac{n-2}{q-1}\right).$$

Therefore, $\tau_{mix} = \Omega\left(n \cdot \exp\left(\frac{q-1}{n-2}\right)\right)$, which means that when $q = o(\frac{n}{\log n})$, τ_{mix} is $n^{\omega(1)}$.

2 Cheeger's Inequality

Sometimes it is more convenient to work with $L = I - P$, the *Laplacian* of P. Then

$$L = \sum_{i=1}^{n} (1 - \lambda_i) \mathbf{v}_i \mathbf{v}_i^{\mathsf{T}} \Pi.$$

For every $i = 1, 2, \ldots, n$, we use γ_i denote $1 - \lambda_i$. Then $0 = \gamma_1 \le \gamma_2 \le \cdots \le \gamma_n \le 2$ are the eigenvalues of L.

The Cheeger's inequality is

$$\frac{\gamma_2}{2} \le \Phi(P) \le \sqrt{2\gamma_2}.$$

There are high-order Cheeger's inequalities indicating the relation between graph expansion and $\lambda_3, \lambda_4, \ldots$.

2.1 Proof of $\gamma_2 \le 2\Phi(P)$

Recall that

$$\gamma_2 = \min_{2-\dim V \subseteq \mathbb{R}^n} \max_{x \in V \setminus \{\mathbf{0}\}} R_L(\mathbf{x}).$$

Therefore, in order to prove an upper bound for γ_2, it suffices to construct some 2-dimensional space V such that any nonzero $\mathbf{x} \in V$ has small $R_L(\mathbf{x})$.

Suppose $\Phi(P) = \Phi(S)$ for some $S \subseteq V$. Let $\mathbf{1}_S$ and $\mathbf{1}_{\overline{S}}$ be the indicator vector of S and its complement $\overline{S}$ respectively. Consider the space $V = \mathrm{span}(\mathbf{1}_S, \mathbf{1}_{\overline{S}})$. Then every $\mathbf{x} \in V$ can be written as $\mathbf{x} = a\mathbf{1}_S + b\mathbf{1}_{\overline{S}}$ for some $a, b \in \mathbb{R}$. We have

$$\begin{aligned}
R_L(a\mathbf{1}_S) &= \frac{\sum_{i \in S, j \in \overline{S}} \pi(i) P(i,j)}{\pi(S)} \\
&= \frac{\sum_{i \in S, j \in \overline{S}} \mathbf{Pr}\left[X_t = i \wedge X_{t+1} = j\right]}{\pi(S)} \\
&= \frac{\mathbf{Pr}\left[X_t \in S \wedge X_{t+1} \in \overline{S}\right]}{\pi(S)} \\
&= \Phi(S).
\end{aligned}$$

Similarly, we have $R_L(b\mathbf{1}_{\overline{S}}) = \Phi(\overline{S})$.

The inequality then follows from the following proposition:

Proposition 1 *If at least one of* $\mathbf{x}$ *and* $\mathbf{y}$ *is not zero, then* $R_L(\mathbf{x} + \mathbf{y}) \le 2 \max\{R_L(\mathbf{x}), R_L(\mathbf{y})\}$.

Assume $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{v}_i$ and $\mathbf{y} = \sum_{i=1}^n b_i \mathbf{v}_i$. Then

$$\begin{aligned}
R_L(\mathbf{x} + \mathbf{y}) &= \frac{\langle \mathbf{x} + \mathbf{y}, L(\mathbf{x} + \mathbf{y}) \rangle_\Pi}{\langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle_\Pi} \\
&= \frac{\sum_{i=1}^n (a_i + b_i)^2 \lambda_i}{\sum_{i=1}^n (a_i + b_i)^2} \\
&\le \frac{2\sum_{i=1}^n (a_i^2 + b_i^2) \lambda_i}{\sum_{i=1}^n a_i^2 + b_i^2} \\
&\le 2 \cdot \max\{R_L(\mathbf{x}), R_L(\mathbf{y})\}.
\end{aligned}$$

2.2 Proof of $\Phi(P) \le \sqrt{2\gamma_2}$

In order to prove an upper bound for $\Phi(P)$, we give an approximation algorithm to estimate $\Phi(P)$ and the upper bound is a consequence of the analysis of its performance. The algorithm is called *Fiedler's Algorithm*: Input Ω and $\mathbf{x} \in \mathbb{R}^{\Omega}$.

Our goal is to find $(S, \bar{S})$ such that $\Phi(S) \vee \Phi(\bar{S}) \le \sqrt{2\gamma_2}$.

- Sort $\Omega = \{v_1, \dots, v_n\}$ according to $\mathbf{x}$ (namely $x(v_1) \le x(v_2) \le \dots$);
- For every $i \in [n]$, let $S_i = \{v_1, v_2, \dots, v_i\}$;
- Return $\min_{i\in[n]} \Phi(S_i) \vee \Phi(\overline{S}_i)$.

We prove the following stronger theorem:

Theorem 2 *For all* $\mathbf{x} \perp \mathbf{1}$*, let* S *be the set returned by Fiedler's algorithm on the input* $\mathbf{x}$*. Then*

$$\Phi(S) \le \sqrt{2R_L(\mathbf{x})}.$$

The Cheeger's inequality then follows by taking $\mathbf{x} = \mathbf{v}_2$.

For simplicity, we assume $\Omega = [n]$ and $\mathbf{x}(1) \le \mathbf{x}(2) \le \dots$ here. To prove the theorem, we first normalize the vector $\mathbf{x}$. Let

$$\ell \triangleq \min_{i\in[n]} \sum_{i=1}^{\ell} \pi(i) \ge \sum_{i=\ell+1}^{n} \pi(i),$$

and for every $i \in [n]$, let $y_i = x_i - x_\ell$. By the definition, $y_\ell = 0$ and $y_i \le 0$ for all $i \le \ell$, $y_i \ge 0$ for all $i \ge \ell$.

We have the following proposition:

Proposition 3 $R_L(\mathbf{x}) \ge R_L(\mathbf{y})$.

To see why it holds, note that

$$R_L(\mathbf{x}) = \frac{\sum_{i,j} \pi(i)P(i,j)(x_i^2 - x_i x_j)}{\langle \mathbf{x}, \mathbf{x}\rangle_\Pi} = \frac{\sum_{\substack{i,j\in\Omega \\ i<j}} \pi(i)P(i,j)\left(x_i - x_j\right)^2}{\langle \mathbf{x}, \mathbf{x}\rangle_\Pi}.$$

Since $\mathbf{y} = \mathbf{x} - y_\ell \mathbf{1}$ is obtained from $\mathbf{x}$ by substracting a constant multiples of $\mathbf{1}$, this operation does not change the numerator and increase the denominator (because $\mathbf{x} \perp \mathbf{1}$). This can also be verified via direct calculation.

As a result, we only need to prove that $\Phi(S) \le \sqrt{2R_L(\mathbf{y})}$. We prove by the *probabilistic method*. That is, we randomly choose some $t \in [y_1, y_n]$ (following a certain tailored density) and consider the expected expansions of $\Phi(S_t)$ and $\Phi(\overline{S}_t)$ where $S_t \triangleq \{i \in [n] \mid y_i \le t\}$.

To this end, we can normalize $\mathbf{y}$ by dividing some constant and assume without loss of generality that $y_1^2 + y_n^2 = 1$. We sample t with density $p(t) = 2|t|$.

Note that for every $t \in [y(1), y(n)]$,

$$\max\left\{\Phi(S_t), \Phi(\overline{S}_t)\right\} = \frac{\sum_{i\in S_t, j\in \bar{S}_t} \pi(i)P(i,j)}{\min\left\{\pi(S_t), \pi(\overline{S}_t)\right\}} =: \frac{A}{B}.$$

We calculate the expectations of the numerator and denominator respectively.

$$
\begin{aligned}
\mathbf{E}\left[A\right] &= \sum_{\substack{i,j\in\Omega\\ i<j}} \pi(i)P(i,j)\mathbf{Pr}\left[i\in S_t, j\in \overline{S}_t\right]\\
&= \sum_{\substack{i,j\in\Omega\\ i<j}} \pi(i)P(i,j)\int_{y_j}^{y_i} 2|t|\mathrm{d}t\\
&= \sum_{\substack{i,j\in\Omega\\ i<j}} \pi(i)P(i,j)\left(\mathrm{sgn}(y_j)\cdot y_j^2-\mathrm{sgn}(y_i)\cdot y_i^2\right)\\
&\le \sum_{\substack{i,j\in\Omega\\ i<j}} \pi(i)P(i,j)\left(|y_i|+|y_j|\right)\left(y_j-y_i\right)\\
&= \sum_{\substack{i,j\in\Omega\\ i<j}} \left(\pi(i)P(i,j)\right)^{\frac{1}{2}}\left(|y_i|+|y_j|\right)\left(\pi(i)P(i,j)\right)^{\frac{1}{2}}\left(y_j-y_i\right)\\
&\overset{(\heartsuit)}{\le} \left(\sum_{\substack{i,j\in\Omega\\ i<j}} \pi(i)P(i,j)\left(|y_i|+|y_j|\right)^2\right)^{\frac{1}{2}}\cdot\left(\sum_{\substack{i,j\in\Omega\\ i<j}} \pi(i)P(i,j)\left(y_j-y_i\right)^2\right)^{\frac{1}{2}}\\
&\le \left(2\sum_{\substack{i,j\in\Omega\\ i<j}} \pi(i)P(i,j)\left(|y_i|^2+|y_j|^2\right)\right)^{\frac{1}{2}}\cdot\sqrt{\langle \mathbf{y}, L\mathbf{y}\rangle_\Pi}\\
&\le \sqrt{2\langle \mathbf{y},\mathbf{y}\rangle_\Pi}\cdot\sqrt{\langle \mathbf{y}, L\mathbf{y}\rangle_\Pi},
\end{aligned}
$$

where $(\heartsuit)$ is due to Cauchy-Schwarz.

$\langle \mathbf{y},\mathbf{y}\rangle_\Pi = \sum_{i,j}\pi(i)P(i,j)y_i^2 \ge \sum_{i<j}\pi(i)P(i,j)(y_i^2+y_j^2).$

On the other hand, we have

$$
\begin{aligned}
\mathbf{E}\left[B\right] &= \mathbf{E}\left[\min\left\{\pi(S_t),\pi(\overline{S}_t)\right\}\right]\\
&= \mathbf{Pr}\left[t<0\right]\mathbf{E}\left[\pi(S_t)\mid t<0\right]+\mathbf{Pr}\left[t\ge 0\right]\mathbf{E}\left[\pi(\overline{S}_t)\;\middle|\; t>0\right].
\end{aligned}
$$

Note that

$$\begin{aligned}
\mathbf{Pr}\left[t < 0\right] \mathbf{E}\left[\pi(S_t) \mid t < 0\right] &= \mathbf{Pr}\left[t < 0\right] \cdot \mathbf{E}\left[\sum_{i=1}^{n} \pi(i) \cdot \mathbf{1}[i \in S_t] \,\middle|\, t < 0\right] \\
&= \mathbf{Pr}\left[t < 0\right] \cdot \sum_{i=1}^{n} \pi(i) \cdot \mathbf{Pr}\left[i \in S_t \mid t < 0\right] \\
&= \sum_{i=1}^{n} \pi(i) \cdot \mathbf{Pr}\left[i \in S_t \wedge t < 0\right] \\
&= \sum_{i=1}^{n} \pi(i) \mathbf{Pr}\left[y_i \le t < 0\right] \\
&= \sum_{i \le \ell} \pi(i) \cdot \int_{y_i}^{0} 2|t| \mathrm{d}t \\
&= \sum_{i=1}^{\ell} \pi(i) y(i)^2.
\end{aligned}$$

Similarly

$$\mathbf{Pr}\left[t \ge 0\right] \mathbf{E}\left[\pi(\overline{S}_t) \,\middle|\, t > 0\right] = \sum_{i=\ell+1}^{n} \pi(i) y(i)^2.$$

Therefore,

$$\mathbf{E}\left[B\right] = \sum_{i=1}^{n} \pi(i) y(i)^2 = \langle \mathbf{y}, \mathbf{y} \rangle_{\Pi}.$$

Now we know that

$$\frac{\mathbf{E}\left[A\right]}{\mathbf{E}\left[B\right]} \le \frac{\sqrt{2\langle \mathbf{y}, \mathbf{y} \rangle_{\Pi}} \cdot \sqrt{\langle \mathbf{y}, L\mathbf{y} \rangle_{\Pi}}}{\langle \mathbf{y}, \mathbf{y} \rangle_{\Pi}} = \sqrt{2R_L(\mathbf{y})} \le \sqrt{2R_L(\mathbf{x})}.$$

Moreover, for any r, we have

$$\frac{\mathbf{E}\left[A\right]}{\mathbf{E}\left[B\right]} \le r \implies \mathbf{E}\left[A - rB\right] \le 0 \implies \mathbf{Pr}\left[\frac{A}{B} \le r\right] > 0.$$

The Cheeger's inequality is proved.