Cheeger Inequalities:

Recall:

Conductance $\varphi(S) := \dfrac{|\partial S|}{\min\{d(S), d(V\setminus S)\}}$; $S \subseteq V$

$$\varphi(G) := \min_{\emptyset \subsetneq S \subsetneq V} \varphi(S)$$

Normalized Laplacian: $N_G = D^{-1/2} L_G D^{-1/2}$

$d(S) \le d(V)/2$

$$\frac{x^T N x}{x^T x} \xleftrightarrow[x = \sqrt{D}y]{} \frac{y^T L y}{y^T D y} \quad \left[\frac{\mathbb{1}_S^T L \mathbb{1}_S}{\mathbb{1}_S^T D \mathbb{1}_S} = \frac{|\partial S|}{d(S)} \right]$$

$\underline{N_G}$

Eigenvalues: $0 = \nu_1 \le \nu_2 \quad . \quad . \quad . \quad \le \nu_n \le 2$

Eigen vectors $\sqrt{D}\mathbb{1} = u_1 \quad u_2 \quad u_n$.

$$\nu_2 = \min_{x \perp \sqrt{D}\mathbb{1}} \frac{x^T N_G x}{x^T x} = \min_{y \perp D\mathbb{1}} \frac{y^T L y}{y^T D y}$$

— Random Walk matrix : $W_G = D_G^{-1} A_G$

$$p^T \longmapsto p^T W_G.$$

$$W_G \sim I - N_G.$$

Cheeger Inequalities:

$$\underbrace{\nu_2/2 \le \varphi}_{} \le \sqrt{2\nu_2}$$

Easy direction (last time)

Today: $\varphi \le \sqrt{2\nu_2}$ [Trevisan's Proof]

Algorithmic:

Input: $y \perp D\mathbb{1}$ s.t $\dfrac{y^T L y}{y^T D y} = \rho$ (or equivalently

$z \perp \sqrt{D}\mathbb{1}$ s.t $z^T N_G z / z^T z = \rho$

Output: $S \subseteq V$ s.t. $\varphi(S) \le \sqrt{2\rho}$.

Spectral Partitioning (Fiedler's Algorithm).

1. Sort vertices in non-decreasing order of values in y

$$y_1 \le y_2 \le \quad . \quad . \quad \le y_n$$

2. Compute $\varphi(\{1, 2, \ldots, i\})$ for all $i \in [n]$ & output the minimum.

Remark: Can be implemented in time $O(|V|\log|V| + |E|)$ (given y)

Cuts considered are of form

$$S_t := \{i \mid y_i \le t\} \quad - \text{Threshold Cuts.}$$

Trevisan's Idea: Design distn on t s.t

$$\frac{\mathbb{E}[|\partial S_t|]}{\mathbb{E}[\min\{d(S_t), d(V \setminus S_t)\}]} \le \sqrt{2\rho} \quad \ldots \quad (*)$$

$(*)$ Implies there exist t s.t $\varphi(S_t) \le \sqrt{2\rho}$.

Distribution on t:

Preprocessing

Step 1: $j_* = \min\{j \in [n] \mid \sum_{i=1}^{j} d(i) \ge d(V)/2\}$

Translate $y \longmapsto z := y - y(j_*)\mathbb{1}$.

Observe: (i) $\text{median}(z) \approx 0$.

(ii) $z^T L z = y^T L y$.

(iii) What about $z^T D z$ vs $y^T D y$

Will show $z^T D z \geq y^T D y$

Hence $\dfrac{z^T L z}{z^T D z} \leq \dfrac{y^T L y}{y^T D y} = \rho$.

(Proof: $z = y + \alpha \mathbb{1}$, $f(\alpha) = z^T D z$

$f'(\alpha) = 2 \sum d(i) z(i)$

Hence, among all z on the line $\{y + \alpha \mathbb{1} \mid \alpha \in \mathbb{R}\}$, the least value of $z^T D z$ is attained when $z \perp d\mathbb{1}$ (ie, $z = y$).

Step 2: Renormalize z s.t

$$z_1^2 + z_n^2 = 1.$$

Distribution t (and S_t)

1. Pick $t \in [z_1, z_n]$ w/ prob density $f(t) = 2|t|$
2. Set $S_t \leftarrow \{i \mid z_i \leq t\}$.

Why this dist?

— $\Pr[t \in [a, b]] = \begin{cases} |b^2 - a^2| & \text{if } a, b \text{ same sign} \\ b^2 + a^2 & \text{if } a, b \text{ opposite signs} \end{cases}$

$\leq |b - a|(|b| + |a|)$ $\quad a < b$

— $d(S_t) = \sum_{u \in V} \mathbb{1}[u \in S_t]\, d(u) = \sum_{u \in V} \mathbb{1}[z_u \leq t]\, d(u)$

Similarly

$$d(V \setminus S_t) = \sum_{u \in V} \mathbb{1}[z_u > t]\, d(u)$$

Centering around 0 guarantees

$$\min\{d(S_t), d(V\setminus S_t)\}$$

$$= \sum_{u\in V} \mathbb{1}[z_u \le t]\cdot \mathbb{1}\left[\frac{t\le 0}{d(u)}\right] + \sum_{u\in V} \mathbb{1}[z_u > t]\cdot \mathbb{1}\left[\frac{t>0}{d(u)}\right]$$

Thus, $\mathbb{E}[\min\{d(S_t), d(V\setminus S_t)\}]$

$$= \sum_{u\in V} \Pr\left[\frac{z_u \le t \,\&\, t\le 0}{d(u)}\right] + \sum_{u\in V} \Pr\left[\frac{z_u > t \,\&\, t>0}{d(u)}\right]$$

$$= \sum_{i\le j} z_i^2 d(i) + \sum_{i\ge j} z_i^2 d(i)$$

$$= \sum_i z_i^2 d(i) = z^T D z.$$

Hence, $\mathbb{E}_t[\min\{d(S_t), d(V\setminus S_t)\}] = z^T D z.$

What about Numerator?

$$\mathbb{E}[|\partial S_t|] = \sum_{(u,v)\in E} \Pr[(u,v)\in \partial S_t] = \sum_{\substack{(u,v)\in E\\ z_u\le z_v}} \Pr[z(u)\le t\le z(v)]$$

$$\le \sum_{(u,v)\in E} |z_v - z_u|\,(|z_v| + |z_u|)$$

$$\le \sqrt{\sum_{(u,v)\in E}(z_v - z_u)^2}\cdot\sqrt{\sum_{(u,v)\in E}(|z_v|+|z_u|)^2}$$

$$= \sqrt{z^T L z}\,\sqrt{\sum_{(u,v)\in E}\left(2z_u^2 + 2z_v^2 - (z_v - z_u)^2\right)}$$

$$= \sqrt{z^T L z}\,\sqrt{2z^T D z - z^T L z}$$

$$\frac{\mathbb{E}[|\partial S_t|]}{\mathbb{E}[\min\{d(S_t), d(V\setminus S_t)\}]} \le \sqrt{\frac{z^T L z}{z^T D z}\left(2 - \frac{z^T L z}{z^T D z}\right)}$$

$$\le \sqrt{\rho(2-\rho)} \quad (\text{since } \rho \le 1).$$

Thus, proved

(In fact, slightly stronger

$\exists\, t, \quad \varphi(S_t) \leq \sqrt{\rho(2-\rho}$

Hence $\quad \varphi(G) \leq \sqrt{\nu_2(2-\nu_2)} \leq \sqrt{2\nu_2}$

Tightness of Cheeger Inequalities:

(a) $\varphi(G) \geq \nu_2/2$: Hypercube H_d

$\nu_2 = 2/d \;; \quad \varphi = \frac{1}{d}$ (for dimension cuts)

(b) $\varphi(G) \leq \sqrt{2\nu_2}$ $\quad \varphi(S) = \frac{2}{2\cdot|S|}$ (S-connected

$\nu_2 = 2\left(1-\cos\frac{2\pi}{n}\right)$

$= \Theta\left(\frac{1}{n^2}\right)$ $\quad \varphi = \frac{2}{n}$

(c) Fiedler's Algorithm: Hypercube H_d

Candidate 2^{nd} e.vector $\sum_{i=1}^{d} x_{e_i} : \{0,1\}^n \to \mathbb{R}$

w/ e.val $2/d$. $\quad (x_1 \ldots x_n) \mapsto \sum_{i=1}^{d} (-1)^{x_i}$

Best cut : majority cut

$\varphi(S) = \Omega\left(\frac{1}{\sqrt{d}}\right)$