Sparsest - CUT (G is a d-reg graph)

$$h(S) = \frac{E(S, V\setminus S)}{d \min(|S|, |V\setminus S|)} \qquad \phi(S) = \frac{\frac{1}{d} E(S, V\setminus S)}{\frac{1}{n} |S| \cdot |V\setminus S|}$$

"how many edges leave S"

$$h \le \phi \le 2h$$

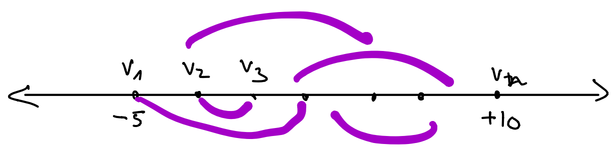
$$1 - \lambda_2 \le \phi \le \sqrt{2(1-\lambda_2)} \qquad \text{Cheeger's inequality}$$

relaxation

$$1-\lambda_2 = \min_{x \perp \mathbb{1}} \frac{x^t L x}{\|x\|^2} = \min_{x \perp \mathbb{1}} \frac{\frac{1}{d} \sum_{i \sim j} (x_i - x_j)^2}{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2} = \min_{\substack{x \ne 0 \\ \ne \mathbb{1} \\ x \in \mathbb{R}^n}} \text{(same)}$$

$$\le \min_{\substack{x \in \{0,1\}^n \\ x \ne \bar{0}, \bar{1}}} \text{(same)} = \phi = \min_{\substack{S \ne 0 \\ S \ne V}} \frac{\frac{1}{d} E(S, V\setminus S)}{\frac{1}{n} |S| \cdot |V\setminus S|}$$

$x \in \mathbb{R}^n$



Spectral Partitioning Alg:

- Given graph G, vector $x \in \mathbb{R}^n$ $(|V| = n)$
  - sort entries of x $\alpha_1 \le \alpha_2 \le \dots \le \alpha_n$
  - look at $(n-1)$ cuts $S_\alpha = \{ i \mid x_i \le \alpha \}$

  & output the sparsest one.

<u>Thm:</u> The cut output by the alg, $(S, \bar{S})$ satisfies

$$\phi(S) \le \sqrt{2\delta} \quad \text{where} \quad \delta = \delta(x) = \frac{\sum_{i,j} M_{ij} (x_i - x_j)^2}{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2}$$

norm. adj. matrix

$$= \frac{\frac{1}{d} \sum_{i \sim j} (x_i - x_j)^2}{\sum x_i^2}$$

Remark: if X is an evec. of λ_2 then alg will output a cut $(S,\bar{S})$ s.t.

$$\underline{h(S)} = \frac{E(S,\bar{S})}{\min(|S|,|\bar{S}|)} \leq \sqrt{2(1-\lambda_2)}.$$

Lemma 1: if $x \perp 1$ there is $y \geq 0$ s.t. $\delta(y) \leq \delta(x)$

& cuts generated by ALG(y) are a subset of cuts generated by ALG(x).

Suppose $y \geq 0$

Lemma 2: $\exists t$ s.t. $\{i \mid y_i \geq t\}$ has $\phi(S) \leq \sqrt{2\delta(y)} \leq \sqrt{2\delta(x)}$

Proof 1: Fix $x \perp 1$. $\delta(x + c\cdot \bar{1}) \leq \delta(x)$

$$\delta(x) = \frac{\frac{1}{d}\sum_{i\sim j}(x_i - x_j)^2}{\sum x_i^2}$$

true $\forall c$, take $-c = \text{median}\{x_1, x_2, \dots, x_n\}$. $x' = x + c\cdot \bar{1}$

$$x_i^+ = \begin{cases} x_i' & x_i' > 0 \\ 0 & x_i \leq 0 \end{cases} \qquad x_i^- = \begin{cases} -x_i' & x_i' < 0 \\ 0 & x_i' \geq 0 \end{cases}$$

$$x' = x^+ - x^-.$$

observe that $\underbrace{\{i \mid x_i^+ \geq t\}}$ is one of the sets the alg generates for X.

$\{i \mid x_i^- \geq t\}$ " "

moreover, these sets are $\frac{n}{2}$ or smaller.

we claim $\delta(x') \geq \min(\delta(x^+), \delta(x^-))$

$$\delta(x') = \frac{\frac{1}{d}\sum_{i\sim j}\left((x_i^+ - x_j^+) - (x_i^- - x_j^-)\right)^2}{\|x^+\|^2 + \|x^-\|^2}$$

$(x_i^+ + x_j^-)^2$ (if $i \geq 0$, $j \leq 0$)

$\mathbin{\rotatebox[origin=c]{90}{$\leq$}}$ $i \leq 0$, $j \geq 0$

case analysis

$x_j', x_i' \geq 0$

$x_j', x_i' \leq 0$

$$\geq \frac{\frac{1}{d}\sum_{i\sim j}(x_i^+ - x_j^+)^2 + \frac{1}{d}\sum_{i\sim j}(x_i^- - x_j^-)^2}{\|x^+\|^2 + \|x^-\|^2}$$

$(x_i^+)^2 + (x_j^-)^2$

$$= \frac{\delta(x^+)\cdot\|x^+\|^2 + \delta(x^-)\cdot\|x^-\|^2}{\|x^+\|^2 + \|x^-\|^2} \geq \min(\delta(x^-), \delta(x^+)).$$

Proof of Lemma 2:

Assume $y \geq 0$, s.t. $y_i = 0$ for $\geq \frac{n}{2}$ i's.

Define distrib over t: choose $\alpha \in [0,1]$ uniformly

set $t = \sqrt{\alpha}$ $\quad (t^2 = \alpha)$

let $S_t = \{ i \mid y_i > t \}$.

ⓐ $\mathbb{E}_t \left(|E(S_t, V \setminus S_t)| \right)$ vs. $\frac{1}{d} \sum_{i \sim j} (y_i - y_j)^2$

ⓑ $\mathbb{E}_t |S_t|$ vs. $\sum y_i^2$ ✓

ⓑ: Fix i

$$\mathbb{E}_t \mathbb{1}_{\{i \in S_t\}} = y_i^2$$

$$\uparrow$$

$$\operatorname*{Prob}_{\alpha \sim [0,1]} \big(\underset{\alpha}{\underset{\shortparallel}{t^2}} \leq y_i^2 \big)$$

$$\sum y_i^2 = \mathbb{E}_t \sum_i \mathbb{1}_{\{i \in S_t\}} = \mathbb{E}_t |S_t|$$

ⓐ: Fix an edge i, j $\qquad y_i^2 > y_j^2$

$$\mathbb{E}_t \mathbb{1}_{\{i,j \text{ cross } S_t \text{ to } V \setminus S_t\}} = |y_i^2 - y_j^2|$$

$$\mathbb{E}\, \tfrac{1}{d} E(S_t, \bar{S}_t) = \frac{1}{d} \mathbb{E}_t \sum_{i \sim j} \mathbb{1}_{\{ij \text{ crosses } S_t \sim V \setminus S_t\}} = \frac{1}{d} \sum_{i \sim j} |y_i^2 - y_j^2|$$

$$= \mathbb{E}\, \frac{1}{d} \sum_{i \sim j} |y_i - y_j| \cdot (y_i + y_j) \underset{CS}{\leq} \underbrace{\sqrt{\frac{1}{d} \sum_{i \sim j} (y_i - y_j)^2}} \cdot \underbrace{\sqrt{\frac{1}{d} \sum_{i \sim j} (y_i + y_j)^2}}_{2\|y\|^2}$$

$$\sqrt{\delta(y) \cdot \|y\|^2} \sqrt{2\|y\|^2} = \sqrt{2 \cdot \delta(y)}\, \|y\|^2$$

$$\therefore \quad \frac{1}{d}\mathop{\mathbb{E}}_{t}\, E(s_t, \bar{s}_t) \;-\; \sqrt{2\delta(y)} \cdot \mathop{\mathbb{E}}_{t} |s_t| \leq 0$$

$\rightarrow \exists t$ s.t. non positive

$$E(s_t, \bar{s}_t) \leq \sqrt{2\delta(y)} \cdot |s_t| .$$

$\square$