Lecture 25

Last time

- Isoperimetric number $\theta(S) = \dfrac{|\partial S|}{|S|} = R_L(1_S)$, where $|\partial S| = e(S, S^c)$

- Cheeger constant $\min_{1 \le |S| \le n/2} \theta(S) = h(G)$

- Spectral gap $\Rightarrow$ expansion

$$h(G) \geq \lambda_2/2$$

- Lem $\forall M$ positive semidefinite, $\forall x \perp y$

$$R_M(x+y) \le 2 \cdot \max\{R_M(x), R_M(y)\}$$

When dealing w/ sets of vxs $\rightsquigarrow$ Cheeger const.
edge expansion $\rightsquigarrow$ conductance

Def: For a d-reg graph $G = (V, E)$, the conductance (or edge expansion) of a set $S \subseteq V$ is

$$\phi(S) = \frac{|\partial S|}{d|S|} = R_N(x) = \frac{\sum_{uv \in E} (x_u - x_v)^2}{d \sum_{v \in V} x_v^2}$$

$|\partial S|$ ← amount going across the cut (S, S^c)

total # edges incident to S → $d|S|$

$x = 1_S$

i.e. $\phi(S)$ is the average fraction of neighbours (of vxs of S) lying outside S.

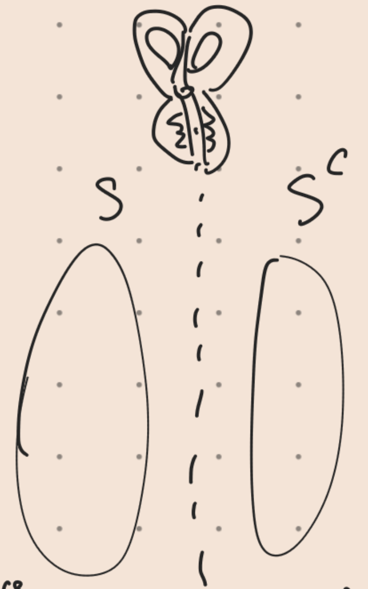
• The conductance (edge expansion) of G

is $\phi(G) = \min_{1 \le |S| \le n/2} \phi(S) = \min_{S} \phi(S, S^c)$

$$\phi(S, S^c) = \max\{\phi(S), \phi(S^c)\}$$

d-reg • $h(G) \geq \lambda_2/2$

$N = I - \frac{1}{d}A = \frac{1}{d}L \Rightarrow$ • $\phi(G) \geq \nu_2/2$



conductance of G $\phi(G)$ $\longleftrightarrow$ sparsest cut

OPEN: if there exists a poly-time approx. w/ constant-factor approx. ratio.

• Find sparse cut

application : • clustering problem.

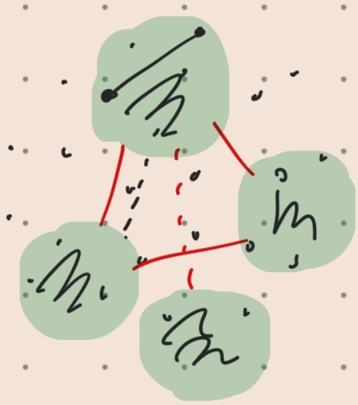


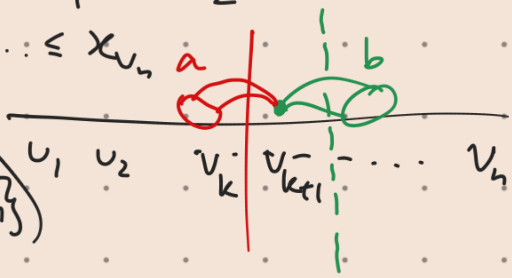
image segmentation

• divide & conquer alg.

Fiedler's algorithm (70s) (applied to x = eigenvector of ν_2)

- Sort vxs so that $x_{v_1} \le x_{v_2} \le \dots \le x_{v_n}$

- Find k min. $\phi(\{v_1, \dots, v_k\}, \{v_{k+1}, \dots, v_n\})$



Time : $O(\underbrace{|E|} + \underbrace{|V|\log|V|})$ ↳ sorting

$e_k = e(\{v_1, \dots, v_k\}, \{v_{k+1}, \dots, v_n\})$

$e_{k+1} = e(\{v_1, \dots, v_{k+1}\}, \{v_{k+2}, \dots, v_n\}) = e_k - a + b$

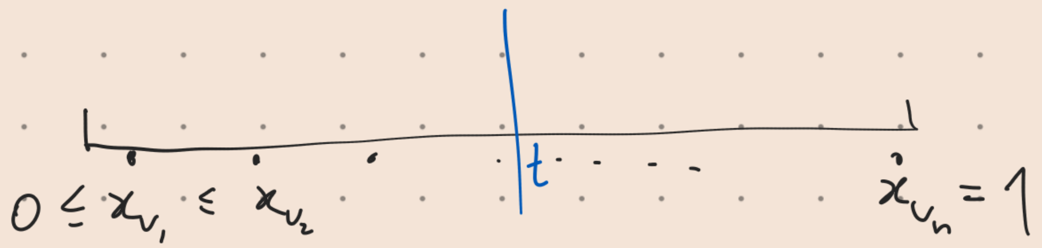
k^{th} step requires $\rightsquigarrow d(v_{k+1})$, $\sum \deg = O(|E|)$.

<u>Thm</u> (Cheeger's ineq) Let $G = (V,E)$ be a regular graph and ν_i be eigenvalues of its normalised Laplacian N.

Then $$\frac{\nu_2}{2} \le \phi(G) \le \sqrt{2\nu_2} = \sqrt{2R_N(x)}$$ ↖ eigenvect. of ν_2

We shall see an analyse of Fiedler's alg. by Trevisan and show that F's alg $\Rightarrow$ cut $\phi(S, S^c) \le \sqrt{2\nu_2}$.

Trevisan's <u>Idea</u>: define a clever distribution to cut randomly and show in expectation (w.r.t. this dist.) the cut is sparse (in term of ν_2) ← x eigenvector of ν_2



random $t \in [0,1]$: t^2 is uniformly distributed in $[0,1]$

i.e. the probability density funct. $f(x) = 2x$.

Thm (robust version) Let $x \perp \mathbb{1}$ and let the cut (S, S^c) be the minimiser in Fiedler's alg.

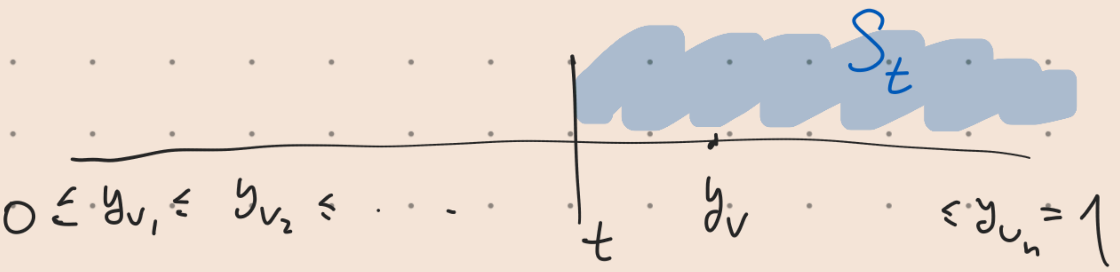
Then $\Rightarrow$ $$\phi(S, S^c) \le \sqrt{2 R_N(x)}$$

Rmk: This robust version can be used on approximate eigenvectors (i.e. vectors w/ small Rayleigh quotient) and produce a sparse cut.

Consider first non-negative vectors

Lem: Let $y \in \mathbb{R}^V_{\ge 0}$. Then $\exists\ t > 0$ s.t.

$$\phi(S_t) := \phi(\{v : y_v \ge t\}) \le \sqrt{2 R_N(y)}.$$



Pf:
- As multiplying by a scalar does not change Rayleigh quotient, i.e. $R_N(cy) = R_N(y)$, we may assume $y_{v_n} = 1$
- Let $t \in [0,1]$ be a random var. w/ prob. density funct $f(x) = 2x$, so that t^2 is unif dist. in $[0,1]$
- It suffices to prove

$$(\heartsuit) \cdots \frac{\mathbb{E}_t |\partial S_t|}{\underbrace{\mathbb{E}(d|S_t|)}_{>0}} \le \sqrt{2 R_N(y)}$$

Indeed, $(\heartsuit) \iff \mathbb{E}\,|\partial S_t| \leq \sqrt{2R_N(y)} \cdot \mathbb{E}(d|S_t|)$

$$\overset{\text{linearity}}{\iff} \quad \mathbb{E}\Big(|\partial S_t| - \sqrt{2R_N(y)} \cdot d|S_t| \Big) \leq 0$$

$\Rightarrow \exists$ a choice of $t \in (0,1]$ s.t.

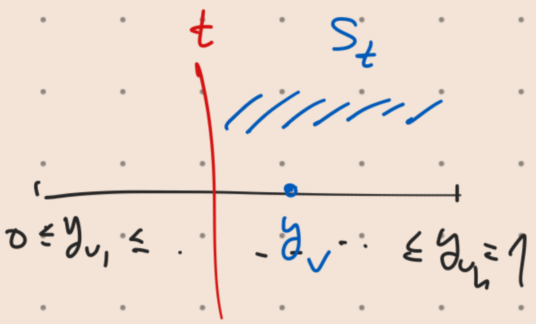
$$|\partial S_t| - \sqrt{2R_N(y)} \cdot d|S_t| \leq 0$$

$$\iff \quad \frac{|\partial S_t|}{d|S_t|} = \phi(S_t) \leq \sqrt{2R_N(y)} \quad \text{as desired.}$$

$$\boxed{(\heartsuit) \quad \cdots \quad \frac{\mathbb{E}_t|\partial S_t|}{\underbrace{\mathbb{E}(d|S_t|)}_{>0}} \leq \sqrt{2R_N(y)}}$$

$S_t = \{v : y_v \geq t\}$

- denominator: $\mathbb{E}(d|S_t|)$



$$= d \sum_{v \in V} \mathbb{P}(v \in S_t)$$

$$= d \sum_{v \in V} \mathbb{P}(t \leq y_v)$$

$$= d \sum_{v \in V} \mathbb{P}(t^2 \leq y_v^2)$$

$$= d \sum_{v \in V} y_v^2$$

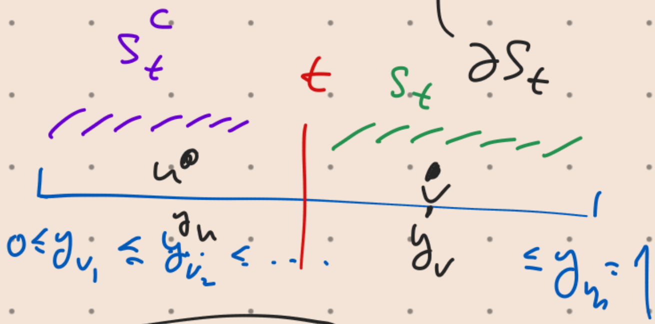
$$R_N(y) = \frac{\sum_{uv \in E} (y_u - y_v)^2}{d \sum y_v^2}$$

• Numerator:

$$\mathbb{E}|\partial S_t| = \sum_{uv\in E} \mathbb{P}(uv \in \partial S_t)$$

say $y_u < y_v$

$$= \sum_{uv\in E} \mathbb{P}(y_u \le t \le y_v)$$



$$= \sum \mathbb{P}(y_u^2 \le t^2 \le y_v^2)$$

$$= \sum_{uv\in E} |y_u^2 - y_v^2|$$

$$= \sum |y_u - y_v|(y_u + y_v)$$

numerator of $R_G(y)$: $\sum_{uv\in E} (y_u - y_v)^2$

C-S

$$\le \sqrt{\sum_{uv\in E} (y_u - y_v)^2} \cdot \sqrt{\sum_E (y_u + y_v)^2}$$

$(y_u + y_v)^2 \le 2(y_u^2 + y_v^2)$

$$\le \sqrt{\sum_E (y_u - y_v)^2} \cdot \sqrt{2d \sum_{v\in V} y_v^2}$$

$$\Rightarrow \frac{\mathbb{E}|\partial S_t|}{\mathbb{E}(d|S_t|)} \le \frac{\sqrt{\sum (y_u - y_v)^2} \cdot \sqrt{2d \sum_v y_v^2}}{d \sum_v y_v^2} = \sqrt{\frac{2 \sum_E (y_u - y_v)^2}{\sum_v y_v^2}}$$

$$= \sqrt{2 R_G(y)}.$$

∎