# Sparsest Cut Generalizations

15-750
4/11/16
4/13/16

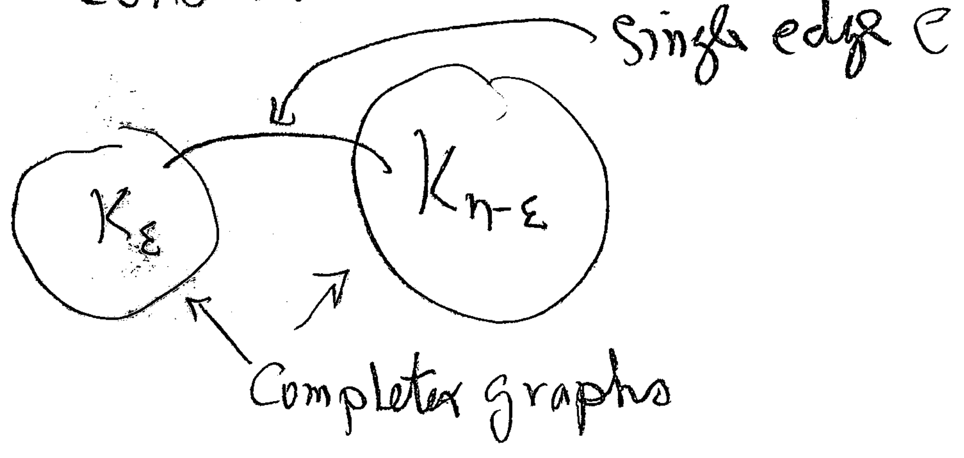
So far we have considered 2 cut probs.

1) S-t-Min Cut
2) Separators (Edge version)

Both wanted small # edges give a vertex partition A, B.

1) required $s \in A$ & $t \in B$
2) $|A|$ & $|B|$ not too big. (Strongly Balanced)

Goal: Consider relaxations of partition condition!



η $G=(V,E)$ network

3) All nodes communicate to all other nodes.

Find network bottle neck!

i.e. $$\min_{S \subseteq V} \frac{|\mathrm{cut}(S,\bar{S})|}{|S|\cdot|\bar{S}|}$$ (Sparsest Cut) (Uniform)

---

Here the communication pattern $K_n$ (the complete graph)

$$\sigma_G(S) = \frac{|\mathrm{Cut}_G(S,\bar{S})|}{|\mathrm{Cut}_H(S,\bar{S})|} \qquad H = K_n$$

$\min\limits_S \sigma_G(S)$ more generally $\sigma(G,H) = \min\limits_S \sigma_{G,H}(S)$

<u>note</u> $G$ flow graph & $H \equiv \underset{s}{\circ}\!-\!\underset{t}{\circ}$ then

$\sigma(G,H) \equiv$ min $s$-$t$ cut.

As in Min s-t Cut lets consider assigning lengths to edges of G, say $d_{uv}$.

Given d over edges to all pairs by shortest path.

Recall: Metric $d: V \times V \to R$

1) $d(x,y) \geq 0$
2) $d(x,y) = d(y,x)$
3) $d(x,z) \leq d(x,y) + d(y,z)$ $\Delta$ ineq
4) $d(x,y) = 0$ iff $x = y$

Def d is semi-metric if 1), 2), 3)

---

Leighton-Rao Relaxation

$$LR(G,H) = \min_{d \text{ semi-metric}} \frac{\sum_{uv \in G} d_{uv}}{\sum_{uv \in H} d_{uv}}$$

---

Simplest semi-metric on V

$f: V \to \{0,1\}$

$d_{uv} = |f(u) - f(v)|$

$$\underline{\text{Sparsest Cut}} = \min_{f: V \to \{0,1\}} \frac{\sum_{uv \in G} |f(u) - f(v)|}{\sum_{uv \in H} |f(u) - f(v)|} = \nabla(G, H)$$

---

$\underline{\text{Thus}}$ $LR(G,H) \leq \nabla(G,H)$

$\underline{\text{Thm}}$ $\nabla(G,H) \leq O(\log n) \cdot LR(G,H)$

What LP for Leighton-Rao?

$\underline{\text{Claim}}$ $\min \sum_{(uv) \in G} d_{uv} = W$

Subject: $\sum_{uv \in H} d_{uv} \geq 1$ $(=1)$ *

$d_{uw} \leq d_{uw} + d_{wv}$ $\forall u, v, w \in V$

$d_{vu} = d_{uv} \geq 0$ $\forall u, v \in V$

---

$\underline{\text{Claim}}$ Opt W to LP $= LR(G,H)$

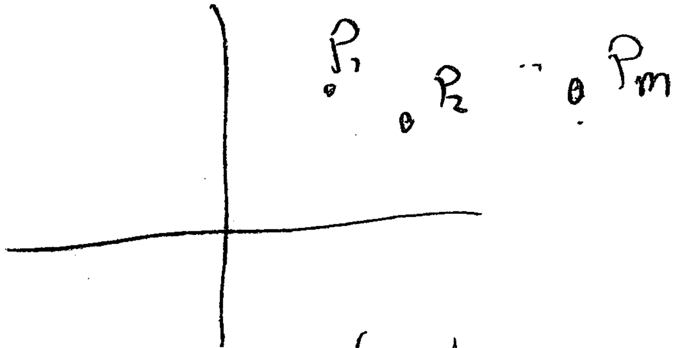
Trevisan: Picks $\frac{|E_H|}{|E_G|}$ as normalization factor!

## Simple fact

**Lemma** $a_1, \ldots, a_m \geq 0$ & $b_1, \ldots, b_m > 0$ then

$$\max_i \frac{a_i}{b_i} \geq \frac{a_1 + \cdots + a_m}{b_1 + \cdots + b_m} \geq \min_i \frac{a_i}{b_i}$$

**pf** consider $\left(\frac{a_i}{b_i}\right) \equiv$ point in $\mathbb{R}^2$ $\quad (b_i, a_i) = P_i$




$\text{mean}(P_1 \cdots P_m) = \left(\sum b_i, \sum a_i\right)/m$

$\text{slope}(\text{Line}(0, P_i)) = \left(\frac{a_i}{b_i}\right)$

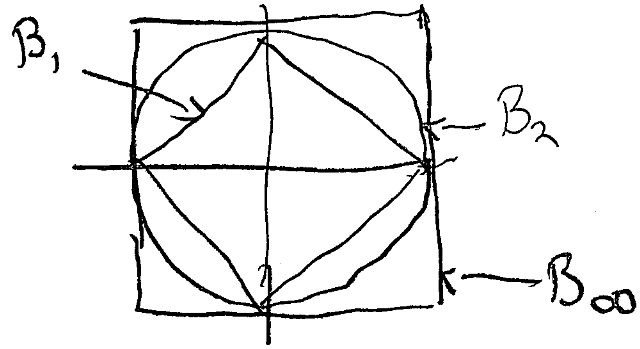
Slope of mean $\geq$ min slope

# $L_1$ Relaxation of Sparsest Cut

$x, y \in \mathbb{R}^n$ $\quad \|x\|_1 = \sum_i |x_i|$ $\quad$ metrix $\|x-y\|_1 = \sum_i |x_i - y_i|$

---

$L_1, L_2, L_\infty$ in $\mathbb{R}^2$ $\quad B_i = \{x \in \mathbb{R}^2 \mid \|x\|_{L_i} = 1\}$




---

$$\sigma_1(G,H) = \inf_{m, f: V \to \mathbb{R}^m} \frac{\sum_{uv \in G} \|f(u) - f(v)\|_1}{\sum_{uv \in H} \|f(u) - f(v)\|_1}$$

---

<u>Thm</u> $\forall G, H$ $\quad \sigma(G,H) = \sigma_1(G,H)$

<u>pf</u> Clearly $\sigma_1(G,H) \leq \sigma(G,H)$

to Show

$$\sigma(G,H) \leq \sigma_1(G,H) \qquad (*)$$

proof of $(*)$

$$f(v) = (f_1(v), \cdots, f_m(v))$$

$$a_i = \sum_{uv \in G} |f_i(u) - f_i(v)| \quad b_i = \sum_{uv \in H} |f_i(u) - f_i(v)|$$

$$\sigma_1(G,H) = \frac{\sum a_i}{\sum b_i} \geq \min_i \frac{a_i}{b_i}$$

Thus WLOG $m=1$ ie $f: V \to \mathbb{R}$

$\sigma$ is invariant under translation & scaling

$$\min_v g(v) = 0 \quad \& \quad \max_v g(v) = 1$$

We are back to a slight Min s-t cut generalization!

## Back to Min s-t Cut

Min-Cut LP (new one)

Variables $Y_{uv}$ $\forall u \neq v \in V$

Input: Flow graph $G=(V,E)$ cap $C_{uv}$, $s, t$

LP $\min \sum_{uv \in E} C_{uv} Y_{uv} \equiv W$

subject: $Y_{st} \geq 1$ $(Y_{st} = 1)$

$Y_{uv} \leq Y_{uw} + Y_{wv}$

$Y_{uv} = Y_{vu} \geq 0$

---

View $Y_{uv} \equiv$ length of $e_{uv}$

$d(u,v) \equiv$ dist in $(G, Y)$

Claim $d(u,v) \leq Y_{uv}$

Induct on # edges in first conterexample.

Old proof:

1) place vertices on $[0,1]$ by dist to s.

2) Some threshhold cut worked.

---

Another View:

Random Variable $C_x$ for $0 \le x \le 1$

$$C_x \equiv \sum \{ C_{uv} \mid Y_{su} \le x < Y_{sv} \}$$

<u>Claim</u> $\mathbb{E}(C_x) = W$

---

Back to Sparsest Cut

Random Variable

$C_x$ for $0 \le x \le 1$

$$\mathrm{Cut}(G,x) = \{ uv \in E_G \mid g(u) \leq x < g(v) \}$$

$$\mathrm{Cut}(H,x) = \{ \qquad \}$$

$$C_x = \frac{|\mathrm{Cut}(G,x)|}{|\mathrm{Cut}(H,x)|}$$

<u>Claim</u> $\mathbb{E}(C_x) \leq \sigma_1(G,H) \leq \sigma(G,H)$

thus $\sigma(G,H) = \mathbb{E}(C_x)$

Finishes Proof of Thm

---

LR embeds G & H into any arbitrary metric.

If they ebedded into $L_1$ we would get a sparsest cut.

<u>Note</u> a finite semi-metric is just dist in a Graph!

<u>Goal</u>: Embed G into $L_1$ with not much distortion

Thm (Bourgain) If $d$ is a semi-metric on $V$ then $\exists f: V \to \mathbb{R}^m$ st

$$\|f(u)-f(v)\|_1 \le d(u,v) \le c \log n \, \|f(u)-f(v)\|_1$$

---

Upto effect algorithm for Bourgain done.

$$LR = \frac{\sum_{uv \in G} d_{uv}}{\sum_{uv \in H} d_{uv}} \ge \frac{\sum_{uv \in G} \|f(u)-f(v)\|_1}{c \log n \sum_{uv \in H} \|f(u)-f(v)\|_1}$$

$$\ge \left(\frac{1}{c \log n}\right) \tau(G,H)$$