Lecture 16

Friday, May 12, 2017 9:42 AM

The Sparsest Cut Problem

Input: • Undirected $G = (V, E)$
• Costs c_e on edges

Output: $S \subseteq V$

Objective: Minimize $\dfrac{c(\partial S)}{|S| \cdot |\bar{S}|}$ $\boxed{\text{RATIO}}$

LP-Relaxation

As last time, we have "distance" variables where in the "integral" soln, $d(u,v) = 1$ if u, v are in sep. sides of the cut & 0 o/w

$$\therefore \text{OPT} \geq \min \frac{\sum_{e \in E} c(e)\, d(e)}{\sum_{u,v \in V \times V} d(u,v)}$$

note this is precisely $|S| \cdot |\bar{S}|$ if $d(u,v) = 1$ $\forall$ pairs $|\{u,v\} \cap S| = 1$

"d satisfies Δ-ineq"

Ratio $\rightsquigarrow$ Constraint

$$\text{OPT} \geq \min \sum_{e \in E} c(e)\, d(e)$$

- $\sum_{u,v \in V \times V} d_{uv} = 1$
- d satisfies metric constraints.

ROUNDING – Try 1

- Solve LP
- Pick a node s (arbitrarily for now; we will choose appropriately later)
- Let $R := \max_{v} d(s,v)$
- Sample $\rho \in (0,R]$ u.a.r.
- $S := \{u \mid d(s,u) \leq \rho\}$

As last time,

$$\mathbb{E}[c(\partial S)] = \frac{1}{\rho} \cdot LP$$

But this time we "also have a denominator".

Ideally, we would've liked to analyze

$$\mathbb{E}\left[\frac{c(\partial S)}{|S| \cdot |\bar{S}|}\right] \quad \ldots\ldots (*)$$

This is a rather unwieldy object. Instead let's look at

$$\frac{\mathbb{E}[c(\partial S)]}{\mathbb{E}[|S| \cdot |\bar{S}|]} \quad \ldots (**)$$

Note these two expressions are NOT the same!

Why is $(**)$ interesting?

$\because$ if we show $\dfrac{\mathbb{E}[c(\partial S)]}{\mathbb{E}[|S| \cdot |\bar{S}|]} \leq \alpha \cdot LP$

$$\mathbb{E}\left[|S| \cdot |\bar{S}|\right]$$

then, by moving terms & LINEARITY OF EXPECTATION,

$$\mathbb{E}\left[\, c(\partial S) - \alpha \cdot LP \cdot |S| \cdot |\bar{S}| \,\right] \leq 0$$

$\Rightarrow$ One of the possible cuts that our algorithm returns satisfies

$$c(\partial S^*) \leq \alpha \cdot LP \cdot |S^*| \cdot |\bar{S}^*|$$

$$\Rightarrow \text{sparsity}(S^*) \leq \alpha \cdot LP$$

Why would we be able to get our hands on S^*?
$\because$ Only a $O(m)$ different ρ's are "interesting" & we can deterministically enumerate.

$\therefore$ It suffices to upper bound $\dfrac{\mathbb{E}[c(\partial S)]}{\mathbb{E}[|S| \cdot |\bar{S}|]}$

- The numerator we already known to be LP/R
- The denominator is again a bit problematic since $|S|$ & $|\bar{S}|$ are not independent (in fact they sum to n)

$\therefore$ For now we use a trivial LOWER BOUND on $|S|$, i.e, $|S| \geq 1$

$$\therefore \mathbb{E}\left[|S| \cdot |\bar{S}|\right] \geq \mathbb{E}\left[|\bar{S}|\right] \cdots\cdots \#$$

The $\mathbb{E}[|\bar{S}|]$ we can calculate.

$$\mathbb{E}[|\bar{S}|] = \frac{1}{R} \sum_{\rho} (\text{\# of vertices } u \text{ s.t. } d(s,u) > \rho)$$

this really should be an integration.

$$= \frac{1}{R} \sum_{\rho} \sum_{v : d(s,v) > \rho} 1 = \frac{1}{R} \sum_{v \in V} \sum_{\rho \le d(s,v)} 1$$

$$= \frac{1}{R} \cdot \sum_{v \in V} d(s,v)$$

- What do we know about d's?

$$\sum_{u,v} d(u,v) = 1$$

$$\therefore \exists s \text{ s.t. } \sum_{v} d(s,v) \geq \frac{1}{n}$$

$\therefore$ If our algorithm starts from this s ...

$$\mathbb{E}[|\bar{S}|] \geq \frac{1}{nR}$$

$$\therefore \frac{\mathbb{E}[c(\partial S)]}{\mathbb{E}[\quad]} \leq n \cdot LP$$

$$\frac{L}{\mathbb{E}[|S| \cdot |\bar{S}|]} \leq n \cdot LP$$

$\Rightarrow$ An n-factor algorithm ..... ☹

Try 2

- The algorithm started cutting around s, a singleton, & $\therefore$ $|S| \geq 1$ was the only thing we could use.

- Modifation:
  - Let $T \subseteq V$ s.t. $|T| \approx \Theta(n)$; say $|T| \geq n/3$
  - $R := \max_v \underbrace{d(T, v)}_{\min_{u \in T} d(u,v)}$; $\rho \in_R (0, R]$
  - $S := \{u \mid d(T, u) \leq \rho\}$

Once again :- $\mathbb{E}[\mathrm{cost}(\partial S)] \leq \frac{1}{R} \cdot LP$

- $|S| \geq |T| \geq n/3$ $\because T \subseteq S$

- What about $\mathbb{E}[|\bar{S}|]$?

As before,

$$\mathbb{E}[|\bar{S}|] = \frac{1}{R} \sum_{\rho} (\# \text{ of } u\text{'s st } d(T,u) \geq \rho)$$

$$= \frac{1}{R} \sum_{\rho} \sum_{u : d(T,u) \geq \rho} 1$$

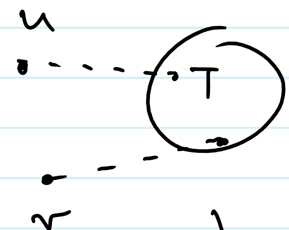
$$= \frac{1}{R} \sum_{\rho} \ \underset{u:d(T,u) \geq \rho}{1}$$

$$= \frac{1}{R} \cdot \sum_{u} \sum_{\rho : \rho \leq d(T,u)} 1$$

$$= \frac{1}{R} \underbrace{\sum_{u} d(T,u)}$$ . . . . . . (*)

When is this "large"? $\hookrightarrow \geq \Omega(1/n)$?

$$1 = \sum_{u,v} d(u,v)$$



$$\leq \sum_{(u,v)} \Big(d(T,u) + \mathrm{diam}(T) + d(T,v) \Big)$$

$$= \binom{n}{2} \cdot \mathrm{diam}(T) + 2n \cdot \sum_{u} d(T,u)$$

$\Rightarrow$ If $\mathrm{diam}(T) \leq \frac{1}{2n^2}$, $\sum_{u} d(T,u) \geq \frac{1}{4n}$

Indeed, if such a T exists, then from (***) we get $\mathbb{E}\big[|\bar{S}|\big] \geq \frac{1}{4nR}$

$\Rightarrow \mathbb{E}\big[|S| \cdot |\bar{S}|\big] \geq \Omega\Big(\frac{1}{\ldots}\Big)$

$$\Rightarrow \mathbb{E}\left[|S| \cdot |\bar{S}|\right] \geq \Omega\left(\frac{1}{R}\right)$$

$$\Rightarrow \frac{\mathbb{E}\left[\text{cost}(\gamma S)\right]}{\mathbb{E}\left[|S| \cdot |\bar{S}|\right]} = O(LP)$$

$\Rightarrow$ An $O(1)$-appx for sparsest cut.

To conclude, we would be happy if there was a $T \subseteq V$ s.t. (a) $|T| = \Omega(n)$ & (b) $\text{diam}(T)$ w.r.t. d is $\leq \frac{1}{2n^2}$

Low-diameter Decomposition Lemma (Leighton-Rao '1988)

Given any undirected $G = (V, E)$ with costs $c(e)$ on edges and distances $d : V \times V \to \mathbb{R}_{\geq 0}$ st

$$L = \sum_e c(e) d(e) \quad \& \quad \sum_{u,v} d_{uv} = 1,$$

and given any $R > 0$, there exists a partition $\{S_1, S_2, \ldots, S_k\} \subseteq V$ s.t.

1. $\text{diam}(S_i) \leq R$
2. Total cost of cross-edges, ie. edges with epts in diff. S_i's, is $\leq O\left(\frac{\log n}{R}\right) \cdot L$