Small Frequencies:

Q: Why does vertex cover have a 2-approx but Set Cover doesn't?

Def: The frequency of $e \in U$ is

$$f_e = |\{ i \in [m] : e \in S_i \}|$$

In vertex cover:

every element (edge) has frequency 2!

Thm: Let $f = \max_{e \in U} f_e$. There is an f-approximation for Set Cover.

Pf:

Alg:

while not all elements covered, choose uncovered element, add all sets containing it.

Analysis: like VC!

Sps alg runs for k iterations

$\Rightarrow ALG \le kf$

No two elements chosen by algorithm are in same set

$\Rightarrow OPT \ge k$

Max Coverage (Max k-Cover):

Input: - Universe $\mathcal{U}$, $|\mathcal{U}| = n$
- Family of sets $S_1, S_2, \ldots, S_m$, each $S_i \subseteq \mathcal{U}$
- Integer $k \leq n$

Feasible Solutions: $I \subseteq [m]$, $|I| \leq k$

Objective: Maximize $\left| \bigcup_{i \in I} S_i \right|$

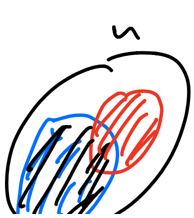
Greedy Algorithm:

Same as for Set Cover, but stop after k iterations!

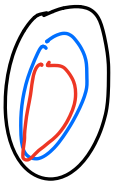
Thm: Greedy is a $(1 - \frac{1}{e})$-approximation

Pf: Like for Set Cover:

- g_t index of set picked by greedy in iteration t
- $I_t = \{g_1, g_2, \ldots, g_t\}$ the sets picked by greedy in first t iterations
- $J_t = \mathcal{U} \setminus \left(\bigcup_{i \in I_t} S_i \right)$ the uncovered elements after t iterations
- $x_t = |S_t \cap J_{t-1}|$ # elements covered at iteration t
- OPT = # elements covered in optimal solution (and solution itself, abusing notation)



$-z_t = OPT - \sum_{j \le t} x_j = OPT - \left|\bigcup_{j \le t} S_j\right|$ not $\left|\bigcup_{i \in OPT} S_i \setminus \bigcup_{j \le t} S_j\right|$



Note: # elements covered by OPT, not covered by $\bigcup_{j \le t} S_j$, is at least z_t.

Claim: $x_{i+1} \ge \frac{z_i}{k}$

Pf: In iteration $i+1$, the k sets in OPT cover at least z_i uncovered elements

$\Rightarrow$ at least one of them covers $\ge \frac{z_i}{k}$ uncovered elements

$\Rightarrow$ greedy covers $\ge \frac{z_i}{k}$ uncovered elements

Claim: $z_i \le (1-\frac{1}{k})^i \cdot OPT$

Pf: induction on iteration i

$i=0$: $z_0 = OPT - 0 = OPT$ ✓

inductive step:

$$OPT - \sum_{j \le i} x_j = OPT - \sum_{j \le i-1} x_j - x_i$$

$$z_i = z_{i-1} - x_i \quad \text{(def of } z_i\text{)}$$

$$\le z_{i-1} - \frac{z_{i-1}}{k} \quad \text{(previous claim)}$$

$$= (1-\tfrac{1}{k})\, z_{i-1}$$

$$\le (1-\tfrac{1}{k})(1-\tfrac{1}{k})^{i-1} \cdot OPT \quad \text{(IH)}$$

$$= (1-\tfrac{1}{k})^i \cdot OPT$$

$$\text{Greedy} = \sum_{i=1}^{k} x_i = OPT - z_k$$

$$z_k = OPT - \sum_{i=1}^{k} x_i$$

$$\geq OPT - \left(1-\frac{1}{k}\right)^k \cdot OPT \qquad \text{(previous claim)}$$

$$\geq OPT - \frac{1}{e} \cdot OPT \qquad \left(\left(1-\frac{1}{k}\right)^k \leq \frac{1}{e}\right)$$

$$= \left(1-\frac{1}{e}\right) \cdot OPT$$

<u>Extensions</u>:

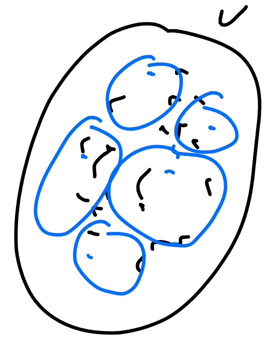
- Submodular optimization
- Minimum k-Union

$O(m^{1/4+\varepsilon})$-approx [CDM '17]

k-center:

Input: - (finite) metric (V,d), $|V|=n$
- integer k with $1 \le k \le n$

Feasible Solution: $F \subseteq V$ with $|F| = k$



Objective: minimize $\max_{v \in V} d(v,F)$

$d(v,S) = \min_{u \in S} d(v,u)$

Greedy:

Init: $F = \{u\}$ for some arbitrary $u \in V$

while $(|F| < k)$ {

Let $u \in V$ be node maximizing $d(u,F)$

Add u to F

}

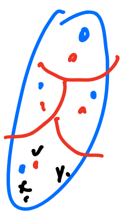
Thm: Greedy is a 2-approximation

Pf: Let F^* optimal solution, $OPT = \max_{u \in V} d(u, F^*)$

F solution returned by greedy

WTS: $\forall u \in V,\ d(u,F) \le 2 \cdot OPT = 2 \cdot \max_{v \in V} d(v, F^*)$

(not $d(u,F) \le 2\,d(u,F^*)$)

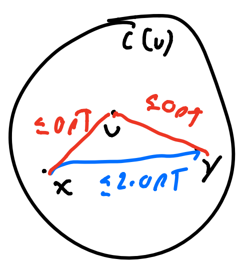


- For each $v \in F^*$, let cluster of v be

$$C(v) = \{u \in V : d(u,v) = d(u,F^*)\}$$

Lemma: Let $x, y \in C(v)$. Then $d(x,y) \le 2 \cdot OPT$

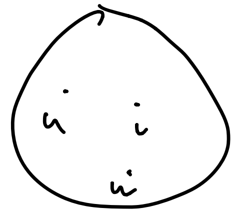
Pf:



$$\begin{aligned} d(x,y) &\le d(x,v) + d(v,y) \qquad \Delta\text{-ineq} \\ &= d(x,F^*) + d(y,F^*) \\ &\le OPT + OPT = 2 \cdot OPT \end{aligned}$$

Consider arbitrary $u \in V$. WTS $d(u,F) \le 2 \cdot OPT$

Case 1: $\forall v \in F^*$, $C(v) \cap F \neq \emptyset$



Let $v \in F^*$ s.t. $u \in C(v)$.

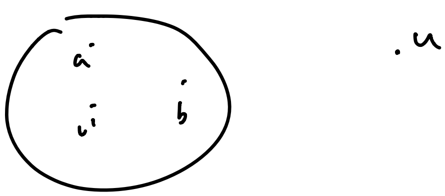
$\Rightarrow \exists\, w \in C(v) \cap F$

$$\Rightarrow d(u,F) \underset{\text{by def}}{\le} d(u,w) \underset{\text{by lemma}}{\le} 2\cdot OPT$$

Case 2: $\exists\, v \in F^*$ s.t. $C(v) \cap F = \emptyset$

$\Rightarrow$ by pigeonhole, $\exists\, v' \in F^*$ s.t. $|C(v') \cap F| \ge 2$

Let $a,b \in C(v') \cap F$, with a added by greedy before b



Let F' be nodes added by greedy until b added

$$\begin{aligned} d(u,F) &\le d(u,F') && (F' \subset F) \\ &\le d(b,F') && (\text{greedy alg}) \\ &\le d(b,a) && (a \in F') \\ &\le 2\cdot OPT && (\text{by lemma}) \end{aligned}$$

Q: Is analysis tight?

A: Yes!

Q: Is there a better algorithm?

Thm: Assuming $P \neq NP$, there is no c-approx for k-center for any $c < 2$.

Pf: A dominating set in $G = (V, E)$ is a set $S \subseteq V$ s.t. every $v \in V$ is either in S or is adjacent to node in S

Dominating set problem: Given G, k, YES if G has a DS of size $\leq k$, NO otherwise

- NP-complete

Reduction: given $G = (V, E)$, k, create metric space (V, d) where

$$d(u,v) = \begin{cases} 1 & \text{if } \{u,v\} \in E \\ 2 & \text{otherwise} \end{cases}$$

Lemma: If G has a dominating set of size $\leq k$, then (V,d) has a k-center solution of cost 1

Pf: Let $S \subseteq V$ be DS of G with $|S| \leq k$

$\Rightarrow \forall u \in V, \exists s(u) \in S$ s.t. $\{u, s(u)\} \in E$

$\Rightarrow \forall u \in V, d(u,S) \leq d(u, s(u)) = 1$

Lemma: If G does not have DS of size $\leq k$, then OPT of k-center on $(V,d), k$ is ≥ 2

Pf: Contrapositive.

Sps (V,d) has k-center solution S of cost < 2

$\Rightarrow \forall u, d(u,S) < 2 \Rightarrow d(u,S) = 1$ or 0

$\Rightarrow u \in S$ or adjacent to node in S

$\Rightarrow S$ a DS of size $\leq k$

So sps had polytime $c < 2$-approx for k-center.

Could solve Dominating Set:

- Given DS instance $(G=(V,E), k)$, create k-center instance
- Run c-approx alg, get back solution of cost α

If $\alpha < 2$, $OPT \leq 1 \Rightarrow$ YES of DS

If $\alpha \geq 2$, $OPT > 1 \Rightarrow OPT \geq 2 \Rightarrow$ NO of DS