K-median : Local Search

K-median

$n=|P|,\ m=|Q|.$

Input: $P, Q \subseteq X$ in metric space (X, d). $k \in \mathbb{N}$.

Goal: Output $\underset{|Q'|=k}{Q' \subseteq Q}$ to minimize $\underbrace{\sum_{i \in P} \min_{j \in Q'} d_{ij}}_{cost(Q')}$.

Local Search with width w.

symmetric difference: $(Q' \setminus Q'') \cup (Q'' \setminus Q')$

$Q' \leftarrow$ arbitrary subset of Q with k centers.

While $\exists Q''$ s.t. $|Q' \Delta Q''| \le w$, $|Q''| = k$, $cost(Q'') < cost(Q')$,

or $cost(Q'') < (1-\varepsilon)\, cost(Q')$

$\quad Q' \leftarrow Q''$.

Output Q'.

Running time: Each while iteration takes $(n+m)^{O(w)}$ time.

$\#$ while iterations = can be a lot, but by doing this,

can be bounded by $\log_{\frac{1}{1-\varepsilon}} \left(\frac{cost(\text{first } Q')}{OPT} \right)$.

Assume: Final Q' has no Q'' s.t. $|Q' \Delta Q''| \le w$, $|Q''| = k$, $cost(Q'') < cost(Q')$!

Q' is a "local optimum".

Analysis

(by duplicating facilities)

Let $Q^* \subseteq Q$ be the optimal solution. WLOG, assume $Q' \cap Q^* = \emptyset$.

Strategy: Let $S \subseteq Q' \cup Q^*$ be a "swap" if $|S \cap Q'| = |S \cap Q^*| \leq w/2$.

Then by local optimality of Q', for any swap S, we know $\text{cost}(Q') \leq \text{cost}(S \Delta Q')$.

We will construct swaps $S_1, \ldots S_t$ with their "weights" $p_1, \ldots, p_t \in \mathbb{R}^+$. Then we have

$$\text{cost}(Q') \leq \text{cost}(S_1 \Delta Q')$$

$$\vdots$$

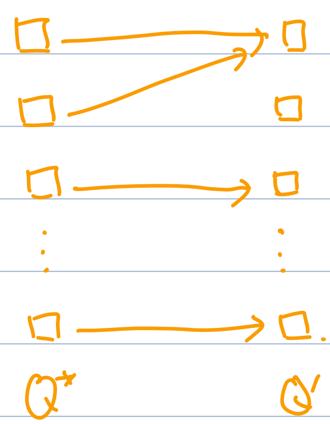
$$\text{cost}(Q') \leq \text{cost}(S_t \Delta Q')$$

$$\Rightarrow \sum_{i=1}^{t} p_i\left(\text{cost}(Q') - \text{cost}(S_i \Delta Q')\right) \leq 0. \qquad (*)$$

$(*)$ will imply $\text{cost}(Q') < \alpha \cdot \text{cost}(Q^*)$ for some $\alpha \in \mathbb{R}^+$.

Swaps: $\forall j \in Q^*$, let $\pi(j) \in Q'$ be closest to j.

$S_1 \ldots S_t$ will have following properties.



- $\forall j \in Q^*$, $\sum_{S_i \ni j} p_i = 1$.

 (let $p' = \max_{j \in Q'} \sum_{S_i \ni j} p_i$.)

- If S_i contains $j \in Q'$

 $\Rightarrow \pi^{-1}(j) \subseteq S_i$

(in particular, if $|\pi^{-1}(j)| > w/2$, then j is not contained in any swap)

Lemma: $\exists$ swaps $S_1 \ldots S_t$ with $p_1 \ldots p_t$ s.t. above are true with $p' = 1 + 2/w$.

$\min_{j\in Q'} d(i,j)$.

Fix $i\in P$, $\ell\in[t]$. Contribution of i,ℓ to (*) is $\delta_{i,\ell} := d(i,Q') - d(i,Q'\triangle S_\ell)$.

Let $j'\in Q'$, $j^*\in Q^*$ be the closest facilities to i. (let $d_i' = d_{ij'}$, $d_i^* = d_{ij^*}$).

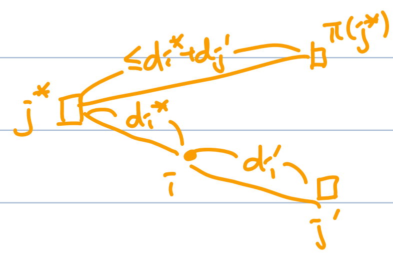
① $S_\ell \ni j^*$: $\delta_{i,\ell}$ is at least $d_i' - d_i^*$

(i can go to j^*).

② $S_\ell \ni j'$ and $S_\ell \not\ni j^*$: By property, either $\pi(j^*) \notin S_\ell$

If $\pi(j^*)\notin S_\ell$, $\delta_{i,\ell} \geq d_i' - (2d_i^* + d_i') = -2d_i^*$

(i can go to $\pi(j^*)$.)



③ Otherwise, $\delta_{i,\ell} \geq 0$.

(i can stay to j').

Then, ① has total weight 1.

② has total weight $\leq p'$.

Therefore, (*) implies

$$0 \geq \sum_{\ell\in[t]} p_\ell\left(\mathrm{cost}(Q') - \mathrm{cost}(Q'\triangle S_\ell)\right)$$

$$= \sum_{i\in P}\left[\sum_{\ell\in[t]} p_\ell\left(d(i,Q') - d(i,Q'\triangle S_\ell)\right)\right]$$

$$\geq \sum_{i\in P}\left[(d_i' - d_i^*)\cdot 1 - (2d_i^*)\cdot p'\right]$$

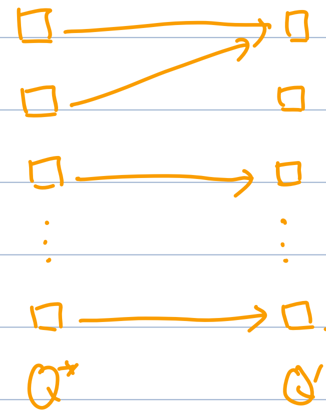
$\therefore$ $\mathrm{cost}(Q') < (1+2p')\,\mathrm{OPT}$.

Constructing Swaps.

- $\forall j \in Q^*$, $\sum_{S_i \ni j} p_i = 1$.

  (let $p' = \max_{j \in Q'} \sum_{S_i \ni j} p_i$.)
- If S_i contains $j \in Q'$

  $\Rightarrow \pi^{-1}(j) \subseteq S_i$



Lemma: $\exists$ swaps $S_1 \dots S_t$ with $p_1 \dots p_t$ s.t. above are true with $p' \le 1 + 2/w$.

Pf. $\forall j \in Q'$, call j big : $|\pi^{-1}(j)| > w/2$

small : $|\pi^{-1}(j)| \in [1, w/2]$

lonely : $|\pi^{-1}(j)| = 0$.

For each small or big j, create a "group" G_j that has

- $\pi^{-1}(j)$
- j $\quad$ ($|G_j| = 2|\pi^{-1}(j)|$)
- $|\pi^{-1}(j)| - 1$ lonely facilities. (Call them R_j).

Then, one can ensure each lonely facility belongs to exactly 1 group.

For small j, G_j becomes a swap with weight 1.

For big j, (let $w' = |\pi^{-1}(j)|$).

- For any $S \subseteq \pi^{-1}(j)$ and $T \subseteq R_j$ with $|S| = |T| = w/2$,

  $(S \cup T)$ becomes a swap with weight $1 / \binom{w'-1}{w/2-1} \cdot \binom{w'-1}{w/2}$.

For $j^* \in \pi^{-1}(j)$, (total weight of swaps containing j) $= 1$.

For $j' \in R_j$, (")

$$= \frac{1}{\binom{w'-1}{w/2-1}\binom{w'-1}{w/2}} \cdot \binom{w'}{w/2} \cdot \binom{w'-2}{w/2-1}$$

$$= \frac{w'}{(w/2)} \cdot \frac{(w/2)}{(w'-1)} = \frac{w'}{w'-1} \leq \frac{w/2+1}{w/2} = 1 + 2/w$$

$\square$