Lifting to paraboloids
Clustering — k-center, k-median

Sándor Kisfaludi-Bak

Computaional Geometry
Summer semester 2020
Overview

• Lifting to paraboloids: Delaunay, Voronoi
  Edelsbrunner–Seidel (1986)
Overview

- Lifting to paraboloids: Delaunay, Voronoi
  Edelsbrunner–Seidel (1986)

- Metric space, clustering
Overview

- Lifting to paraboloids: Delaunay, Voronoi
  Edelsbrunner–Seidel (1986)

- Metric space, clustering

- k-center, greedy clustering
Overview

• Lifting to paraboloids: Delaunay, Voronoi
  Edelsbrunner–Seidel (1986)

• Metric space, clustering

• k-center, greedy clustering

• k-median, local search 1
Lifting to a paraboloid

\[ L(x, y) = (x, y, x^2 + y^2) \]

\( L \) projects \((x, y)\) vertically up to the paraboloid \( A : z = x^2 + y^2 \)
Lifting to a paraboloid

\[ L(x, y) = (x, y, x^2 + y^2) \]

\( L \) projects \((x, y)\) vertically up to the paraboloid \( A : z = x^2 + y^2 \)

\[ \gamma : (x - x_0)^2 + (y - y_0)^2 = r^2 \]
Lifting to a paraboloid

\[ L(x, y) = (x, y, x^2 + y^2) \]

\( L \) projects \((x, y)\) vertically up to the paraboloid \( A: z = x^2 + y^2 \)

\[ \gamma: (x - x_0)^2 + (y - y_0)^2 = r^2 \]

\((x, y) \in \gamma \Rightarrow \)

\[ x^2 + y^2 = r^2 + 2xx_0 + 2yy_0 - x_0^2 - y_0^2 \]

\[ = \alpha_1 x + \alpha_2 y + c \]
Lifting to a paraboloid

\[ L(x, y) = (x, y, x^2 + y^2) \]

\( L \) projects \((x, y)\) vertically up to the paraboloid \( A : z = x^2 + y^2 \)

\[ \gamma : (x - x_0)^2 + (y - y_0)^2 = r^2 \]

\((x, y) \in \gamma \Rightarrow \)

\[ x^2 + y^2 = r^2 + 2xx_0 + 2yy_0 - x_0^2 - y_0^2 = \alpha_1 x + \alpha_2 y + c \]

\[ L(x, y) = (x, y, \alpha_1 x + \alpha_2 y + c) \]

\[ L(\gamma) \subset H_\gamma := \{(x, y, z) \mid -\alpha_1 x - \alpha_2 y + z = c\} \]
Lifting an empty circumcircle

\( pp'p'' \) is a Delaunay-triangle of \( P \)

\[ \Leftrightarrow \]

\( \gamma = \text{circumcircle of } pp'p'' \) is empty

\[ \Leftrightarrow \]

\( A \cap H_\gamma \) is empty

\[ \Leftrightarrow \]

\( H_\gamma \) is a face of \( \text{conv}^\perp(L(P)) \)
Lifting an empty circumcircle

$pp'p''$ is a Delaunay-triangle of $P$

$\iff$

$\gamma = \text{circumcircle of } pp'p''$ is empty

$\iff$

$A \cap H_\gamma \downarrow$ is empty

$\iff$

$H_\gamma$ is a face of $\text{conv} \downarrow (L(P))$

$DT(P) = \text{proj}_{z=0}(\text{conv} \downarrow (L(P)))$
Lifting all of $\mathbb{R}^3$:

$L(x, y, z) = (x, y, z + x^2 + y^2)$

$B_{x', y'} = \{ (x, y, z) \mid z = -(x - x')^2 - (y - y')^2 \}$

Lifting a paraboloid
Lifting all of $\mathbb{R}^3$:

$L(x, y, z) = (x, y, z + x^2 + y^2)$

$B_{x', y'} = \{(x, y, z) \mid z = -(x - x')^2 - (y - y')^2\}$

$L(x, y, -(x' - x)^2 - (y' - y)^2) = (x, y, x'^2 + y'^2 + 2x'x + 2y'y)$
Lifting a paraboloid

Lifting all of $\mathbb{R}^3$:

$L(x, y, z) = (x, y, z + x^2 + y^2)$

$B_{x', y'} = \{(x, y, z) \mid z = -(x - x')^2 - (y - y')^2\}$

$L(x, y, -(x' - x)^2 - (y' - y)^2) = (x, y, x'^2 + y'^2 + 2x'x + 2y'y)$

A plane! touches $A$ at $L(x', y')$
Lifting many paraboloids: Voronoi

Opaque hanging paraboloid $B_p$ for each $p \in P$.

$$\text{dist}(q, p') = \text{dist}(q, p) \iff q^* \in B_p \cap B_q$$
Lifting many paraboloids: Voronoi

Opaque hanging paraboloid $B_p$ for each $p \in P$.

\[ \text{dist}(q, p') = \text{dist}(q, p) \iff q^* \in B_p \cap B_q \]

upper envelope of $\bigcup_{p \in P} B_p$ looks like Vor($P$) from $(0, 0, \infty)$

Apply $L(.)$: polyhedron $\hat{B}$ with face $L(B_p)$ touching $A$ at $L(p)$. $L$ does not change view from $(0, 0, \infty)$
Lifting many paraboloids: Voronoi

Opaque hanging paraboloid $B_p$ for each $p \in P$.

$$\text{dist}(q, p') = \text{dist}(q, p) \iff q^* \in B_p \cap B_q$$

upper envelope of $\bigcup_{p \in P} B_p$ looks like $\text{Vor}(P)$ from $(0, 0, \infty)$

Apply $L(.):$ polyhedron $\hat{B}$ with face $L(B_p)$ touching $A$ at $L(p)$. $L$ does not change view from $(0, 0, \infty)$

$$\text{Vor}(P) = \text{proj}_{z=0}(\hat{B}) = \text{proj}_{z=0} \left( \bigcap_{p \in P} \text{touchplane}_{A(L(p))}^\uparrow \right)$$
Voronoi and Delaunay in higher dimensions?

Paraboloid lifting works in $\mathbb{R}^d$.
Vor($P$) and $DT(P)$ are projections of convex hulls in $\mathbb{R}^{d+1}$.
Voronoï and Delaunay in higher dimensions?

Paraboloid lifting works in $\mathbb{R}^d$. Vor($P$) and $DT(P)$ are projections of convex hulls in $\mathbb{R}^{d+1}$.

- Vor($P$) and $DT(P)$ in $\mathbb{R}^d$ have complexity $O(n^{\lceil d/2 \rceil})$.
Voronoi and Delaunay in higher dimensions?

Paraboloid lifting works in $\mathbb{R}^d$. Vor($P$) and DT($P$) are projections of convex hulls in $\mathbb{R}^{d+1}$.

- Vor($P$) and DT($P$) in $\mathbb{R}^d$ have complexity $O(n^{\lceil d/2 \rceil})$
- Vor($P$) and DT($P$) in $\mathbb{R}^d$ can be computed by convex hull algorithm in $\mathbb{R}^{d+1}$
Voronoi and Delaunay in higher dimensions?

Paraboloid lifting works in $\mathbb{R}^d$. Vor($P$) and $DT(P)$ are projections of convex hulls in $\mathbb{R}^{d+1}$.

- Vor($P$) and $DT(P)$ in $\mathbb{R}^d$ have complexity $O(n^{\lceil d/2 \rceil})$

- Vor($P$) and $DT(P)$ in $\mathbb{R}^d$ can be computed by convex hull algorithm in $\mathbb{R}^{d+1}$

$\mathbb{R}^3$: e.g. skew lines have Vor($P$) complexity $\Theta(n^2)$
Clustering variants in metric spaces
Definition. \((X, \text{dist})\) metric space with distance \(\text{dist} : X \times X \rightarrow \mathbb{R}_{\geq 0}\) iff \(\forall a, b, c \in X:\)

- \(\text{dist}(a, b) = \text{dist}(b, a)\) (symmetric)
- \(\text{dist}(a, b) = 0 \iff a = b\)
- \(\text{dist}(a, b) + \text{dist}(b, c) \geq \text{dist}(a, c)\) (triangle ineq.)
Metric spaces and clustering

**Definition.** \((X, \text{dist})\) metric space with distance \(\text{dist} : X \times X \rightarrow \mathbb{R}_{\geq 0}\) iff \(\forall a, b, c \in X:\)
- \(\text{dist}(a, b) = \text{dist}(b, a)\) (symmetric)
- \(\text{dist}(a, b) = 0 \iff a = b\)
- \(\text{dist}(a, b) + \text{dist}(b, c) \geq \text{dist}(a, c)\) (triangle ineq.)

**Clustering:**
given data, find similar entries and put them together
Metric spaces and clustering

**Definition.** \((X, \text{dist})\) metric space with distance \(\text{dist} : X \times X \rightarrow \mathbb{R}_{\geq 0}\) iff \(\forall a, b, c \in X:\)

- \(\text{dist}(a, b) = \text{dist}(b, a)\) (symmetric)
- \(\text{dist}(a, b) = 0 \iff a = b\)
- \(\text{dist}(a, b) + \text{dist}(b, c) \geq \text{dist}(a, c)\) (triangle ineq.)

**Clustering:**
given data, find similar entries and put them together

Given \(P \subseteq X\), find a set of \(k\) centers \(C \subseteq X\) s.t.

\[
vec_C := \left( \text{dist}(p_1, C), \text{dist}(p_2, C), \ldots, \text{dist}(p_n, C) \right)
\]
is

”small”
Clustering variants

- $k$-center:

$$\min_{C \subset X, |C|=k} \|vec_C\|_\infty = \min_{C \subset X, |C|=k} \max_{p \in P} \text{dist}(p, C)$$

“minimize the max distance to nearest center”
Clustering variants

• $k$-center:

$$\min_{C \subset X, |C| = k} \| vec_C \|_{\infty} = \min_{C \subset X, |C| = k} \max_{p \in P} \text{dist}(p, C)$$

“minimize the max distance to nearest center”

a.k.a. cover $X$ with $k$ disks of radius $r$, minimizing $r$
Clustering variants

- **k-center:**
  \[ \min_{C \subset X, |C| = k} \| vec_C \|_\infty = \min_{C \subset X, |C| = k} \max_{p \in P} \text{dist}(p, C) \]

  "minimize the max distance to nearest center"
  a.k.a. cover \( X \) with \( k \) disks of radius \( r \), minimizing \( r \)

- **k-median:**
  \[ \min_{C \subset X, |C| = k} \| vec_C \|_1 = \min_{C \subset X, |C| = k} \sum_{p \in P} \text{dist}(p, C) \]

  "minimize sum of distances to nearest center"
Clustering variants

- **k-center:**

  \[
  \min_{C \subset X, |C| = k} \|vec_C\|_\infty = \min_{C \subset X, |C| = k} \max_{p \in P} \text{dist}(p, C)
  \]

  "minimize the max distance to nearest center"

  a.k.a. cover \(X\) with \(k\) disks of radius \(r\), minimizing \(r\)

- **k-median:**

  \[
  \min_{C \subset X, |C| = k} \|vec_C\|_1 = \min_{C \subset X, |C| = k} \sum_{p \in P} \text{dist}(p, C)
  \]

  "minimize sum of distances to nearest center"

- **k-means:**

  \[
  \min_{C \subset X, |C| = k} \|vec_C\|_2 = \min_{C \subset X, |C| = k} \sqrt{\sum_{p \in P} \left(\text{dist}(p, C)\right)^2}
  \]

  "minimize sum of squared distances to nearest center"
Clustering variants

- **k-center:**
  \[
  \min_{C \subseteq X, |C| = k} \| \text{vec}C \|_\infty = \min_{C \subseteq X, |C| = k} \max_{p \in P} \text{dist}(p, C)
  \]
  “minimize the max distance to nearest center”
  a.k.a. cover \(X\) with \(k\) disks of radius \(r\), minimizing \(r\)

- **k-median:**
  \[
  \min_{C \subseteq X, |C| = k} \| \text{vec}C \|_1 = \min_{C \subseteq X, |C| = k} \sum_{p \in P} \text{dist}(p, C)
  \]
  “minimize sum of distances to nearest center”

- **k-means:**
  \[
  \min_{C \subseteq X, |C| = k} \| \text{vec}C \|_2 = \min_{C \subseteq X, |C| = k} \sqrt{\sum_{p \in P} (\text{dist}(p, C))^2}
  \]
  “minimize sum of squared distances to nearest center”
Facility location

Opening a center at $x \in X$ has cost $\gamma(x)$. Total cost is

$$\sum_{x \in C} \gamma(x) + \|\text{vec}C\|_1$$

“Hip” topic.
$k$-center via greedy
Hardness of $k$-center

Theorem (Feder–Greene 1988). There is no polynomial time 1.8-approximation for $k$-center in $\mathbb{R}^2$, unless $P = NP$. 
Theorem (Feder–Greene 1988). There is no polynomial time 1.8-approximation for \( k \)-center in \( \mathbb{R}^2 \), unless \( P = NP \).

Reduction from planar vertex cover of max degree 3
Theorem (Feder–Greene 1988). There is no polynomial time 1.8-approximation for \( k \)-center in \( \mathbb{R}^2 \), unless \( P = NP \).

Reduction from planar vertex cover of max degree 3

Double subdivision:

Makes equivalent instance of VC with \( k \to k + 1 \).
Hardness of $k$-center

**Theorem (Feder–Greene 1988).** There is no polynomial time 1.8-approximation for $k$-center in $\mathbb{R}^2$, unless $P = NP$.

Reduction from planar vertex cover of max degree 3

Double subdivision:

Makes equivalent instance of $VC$ with $k \to k + 1$.

Subdivide, get length 2 edges and "smooth" turns only:

Theorem (Feder–Greene 1988). There is no polynomial time 1.8-approximation for $k$-center in $\mathbb{R}^2$, unless $P = NP$. 
Hardness of $k$-center: disk radii

\[ P := \text{edge midpoints of smooth drawing of } G' \]
Hardness of $k$-center: disk radii

$P := \text{edge midpoints of smooth drawing of } G'$

$\exists \text{ VC of size } k \text{ in } G' \iff \exists k\text{-center with radius } 1$
Hardness of $k$-center: disk radii

$P := \text{edge midpoints of smooth drawing of } G'$

$\exists \text{ VC of size } k \text{ in } G' \iff \exists k\text{-center with radius } 1$

Otherwise needs $\geq 1$ disk covering 2 non-neighbors $u, v$

$\text{dist}(u, v) \geq 2 \cdot 1.8 \Rightarrow r \geq 1.8$
Greedy centers

Given $C \subseteq P$, the greedy next center is $q \in P$ where $\text{dist}(q, C)$ is maximized.

Greedy clustering:
start with arbitrary $c_1 \in P$.
For $i = 2, \ldots, k$:
    Let $c_i = \text{GreedyNext}(c_1, \ldots, c_{i-1})$.
Return $\{c_1, \ldots, c_k\}$
### Greedy centers

Given $C \subseteq P$, the greedy next center is $q \in P$ where $\text{dist}(q, C)$ is maximized.

**Greedy clustering:**

start with arbitrary $c_1 \in P$.

For $i = 2, \ldots, k$:

Let $c_i = \text{GreedyNext}(c_1, \ldots, c_{i-1})$.

Return $\{c_1, \ldots, c_k\}$

Let $r_i = \max_{p \in P} \text{dist}(p, \{c_1, \ldots, c_i\})$.

Balls of radius $r_i$ with centers $\{c_1, \ldots, c_i\}$ cover $P$ for any $i$.

$\Rightarrow r_k, \{c_1, \ldots, c_k\}$ is valid $k$-center
Greedy centers

Given $C \subseteq P$, the greedy next center is $q \in P$ where $\text{dist}(q, C)$ is maximized.

**Greedy clustering:**
- start with arbitrary $c_1 \in P$.
- For $i = 2, \ldots, k$:
  - Let $c_i = \text{GreedyNext}(c_1, \ldots, c_{i-1})$.
  - Return $\{c_1, \ldots, c_k\}$

Let $r_i = \max_{p \in P} \text{dist}(p, \{c_1, \ldots, c_i\})$.
- Balls of radius $r_i$ with centers $\{c_1, \ldots, c_i\}$ cover $P$ for any $i$.
  - $r_k, \{c_1, \ldots, c_k\}$ is valid $k$-center

Store most distant center and update in each step
  - $\Rightarrow O(nk)$ time
Greedy $k$-center approximation quality

**Theorem.** Greedy $k$-center gives a 2-approximation.

*Proof*

\[ r_1 \geq r_2 \geq \cdots \geq r_k \]
Greedy $k$-center approximation quality

**Theorem.** Greedy $k$-center gives a 2-approximation.

**Proof**

$r_1 \geq r_2 \geq \cdots \geq r_k$

c_{k+1} := \text{point realizing } r_k$

If $i < j \leq k + 1$, then

$$\text{dist}(c_i, c_j) \geq \text{dist}(c_j, \{c_1, \ldots, c_{j-1}\}) = r_{j-1} \geq r_k$$
Greedy $k$-center approximation quality

**Theorem.** Greedy $k$-center gives a $2$-approximation.

**Proof**

$r_1 \geq r_2 \geq \cdots \geq r_k$

$c_{k+1} := \text{point realizing } r_k$

If $i < j \leq k + 1$, then

$$\text{dist}(c_i, c_j) \geq \text{dist}(c_j, \{c_1, \ldots, c_{j-1}\}) = r_{j-1} \geq r_k$$

$r_{\text{opt}} := \text{is optimal } k\text{-cover radius, suppose } 2r_{\text{opt}} < r_k$
Greedy $k$-center approximation quality

**Theorem.** Greedy $k$-center gives a 2-approximation.

**Proof**

$r_1 \geq r_2 \geq \cdots \geq r_k$

$c_{k+1} := $ point realizing $r_k$

If $i < j \leq k + 1$, then

$$\text{dist}(c_i, c_j) \geq \text{dist}(c_j, \{c_1, \ldots, c_{j-1}\}) = r_{j-1} \geq r_k$$

$r_{opt} := $ is optimal $k$-cover radius, suppose $2r_{opt} < r_k$

$\Rightarrow$ each ball in opt has $\leq 1$ pt from $c_1, \ldots, c_{k+1}$
Definition. $S \subset X$ is an $r$-packing if

- $r$-balls cover $X$: $\text{dist}(x, S) \leq r$ for each $x \in X$
- $S$ is sparse: $\text{dist}(s, s') \geq r$ for each $s, s' \in S$
Definition. $S \subseteq X$ is an $r$-packing if

- $r$-balls cover $X$: $\text{dist}(x, S) \leq r$ for each $x \in X$
- $S$ is sparse: $\text{dist}(s, s') \geq r$ for each $s, s' \in S$

Theorem. For any $i$, $\{c_1, \ldots, c_i\}$ is an $r_i$-packing.
Exact $k$-center in $\mathbb{R}^d$, approximating $k$

Trivial: $O(n^{k+1})$

$\mathbb{R}^2$, $n^{O(\sqrt{k})}$, or $2^{O(\sqrt{n})}$
Exact $k$-center in $\mathbb{R}^d$, approximating $k$

**Trivial:** $O(n^{k+1})$

<table>
<thead>
<tr>
<th>$\mathbb{R}^2$</th>
<th>$n^{O(\sqrt{k})}$ or $2^{O(\sqrt{n})}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{R}^d$, $d =$ const.</td>
<td>no $n^{o(k)}$ known $2^{O(n^{1-1/d})}$</td>
</tr>
</tbody>
</table>
Exact $k$-center in $\mathbb{R}^d$, approximating $k$

Trivial: $O(n^{k+1})$

<table>
<thead>
<tr>
<th>$\mathbb{R}^2$</th>
<th>$n^{O(\sqrt{k})}$ or $2^{O(\sqrt{n})}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{R}^d, d = \text{const.}$</td>
<td>no $n^{o(k)}$ known $2^{O(n^{1-1/d})}$</td>
</tr>
</tbody>
</table>

"optimal"
Exact $k$-center in $\mathbb{R}^d$, approximating $k$

Trivial: $O(n^{k+1})$

<table>
<thead>
<tr>
<th>$\mathbb{R}^2$</th>
<th>$n^{O(\sqrt{k})}$ or $2^{O(\sqrt{n})}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{R}^d, d = \text{const.}$</td>
<td>no $n^{o(k)}$ known</td>
</tr>
</tbody>
</table>

Fix $r$, approximate $k$ instead:

poly $(1 + \varepsilon)$-approximation for any fixed $d, \varepsilon$ (PTAS)
Exact $k$-center in $\mathbb{R}^d$, approximating $k$

Trivial: $O(n^{k+1})$

$\mathbb{R}^2$

\[ n^{O(\sqrt{k})} \]

or $2^{O(\sqrt{n})}$

"optimal"

$\mathbb{R}^d, d = \text{const.}$

no $n^{o(k)}$ known

2$^{O(n^{1-1/d})}$

Fix $r$, approximate $k$ instead:

poly $(1 + \varepsilon)$-approximation for any fixed $d, \varepsilon$ (PTAS)

Later lectures!
\( k \)-median
\( k \)-median via local search
$k$-median via local search

- Compute $C = \{c_1, \ldots, c_k\}$ and $r_k : k$-center 2-approx..
  Gives $2n$-approx for $k$-median as
  \[ \|\text{vec}_C\|_1 \leq n\|\text{vec}_C\|_\infty \]
  so $\text{OPT}(k\text{-med}) \leq n\text{OPT}(k\text{-cent}) \leq 2nr_k$

- Iteratively replace $c \in C$ with $c'$ if it improves $\|\text{vec}_C\|_1$
  (by at least factor $1 - \tau$, $\tau = \frac{1}{10k}$)
  \[ \Rightarrow \text{Results in local opt center set } L \]
\textbf{$k$-median via local search}

- Compute $C = \{c_1, \ldots, c_k\}$ and $r_k : k$-center $2$-approx. Gives $2n$-approx for $k$-median as

$$\|\text{vec} C\|_1 \leq n \|\text{vec} C\|_\infty$$

so $\text{OPT}(k\text{-med}) \leq n \text{OPT}(k\text{-cent}) \leq 2nr_k$

- Iteratively replace $c \in C$ with $c'$ if it improves $\|\text{vec} C\|_1$ (by at least factor $1 - \tau$, $\tau = \frac{1}{10k}$)

$\Rightarrow$ Results in local opt center set $L$

\textbf{Running time:} $O(nk)$ possible swaps, $O(nk)$ to compute new distances. At most $\log \frac{1}{1 - \tau} 2n$ swaps.

$$O((nk)^2 \log \frac{1}{1 - \tau} 2n) = O((nk)^2 \log_{1 + \tau} n) = O((nk)^2 \cdot 10k \log n) = O(k^3 n^2 \log n)$$
Theorem. The local optimum $L$ gives a 5-approximation for $k$-median.

Challenge: $L$ and $OPT$ may be very different. 
Idea: use “intermediate” clustering $\Pi$ to relate them.
Theorem. The local optimum $L$ gives a 5-approximation for $k$-median.

Challenge: $L$ and $OPT$ may be very different. 
Idea: use "intermediate" clustering $\Pi$ to relate them

assign cluster of center $o \in OPT$ to $nn(o, L)$
Theorem. The local optimum $L$ gives a 5-approximation for $k$-median.

Challenge: $L$ and $OPT$ may be very different.
Idea: use "intermediate" clustering $\Pi$ to relate them like $L$, but respects clusters of $OPT$. 
Cost of moving from $L$ to $\Pi$

$\Pi(p), L(p), OPT(p)$ be the center (= nearest neighbor) of $p$ in each clustering.
Cost of moving from $L$ to $\Pi$

$\Pi(p), L(p), OPT(p)$ be the center (= nearest neighbor) of $p$ in each clustering.

Claim. $\|vec_\Pi\|_1 - \|vec_L\|_1 \leq 2\|vec_{OPT}\|_1$.

$$
\begin{align*}
\text{dist}(p, \Pi(p)) & \leq \text{dist}(p, OPT(p)) + \text{dist}(OPT(p), \Pi(p)) \\
& \leq \text{dist}(p, OPT(p)) + \text{dist}(OPT(p), L(p)) \\
& \leq \text{dist}(p, OPT(p)) + \text{dist}(OPT(p), p) \\
& \quad + \text{dist}(p, L(p)) \\
& = 2\text{dist}(p, OPT(p)) + \text{dist}(p, L(p))
\end{align*}
$$
Cost of moving from $L$ to $\Pi$

$\Pi(p), L(p), OPT(p)$ be the center (= nearest neighbor) of $p$ in each clustering.

Claim. $\|vec_{\Pi}\|_1 - \|vec_L\|_1 \leq 2\|vec_{OPT}\|_1.$

$$\begin{align*}
\text{dist}(p, \Pi(p)) &\leq \text{dist}(p, OPT(p)) + \text{dist}(OPT(p), \Pi(p)) \\
&\leq \text{dist}(p, OPT(p)) + \text{dist}(OPT(p), L(p)) \\
&\leq \text{dist}(p, OPT(p)) + \text{dist}(OPT(p), p) \\
&\quad + \text{dist}(p, L(p)) \\
&= 2\text{dist}(p, OPT(p)) + \text{dist}(p, L(p))
\end{align*}$$

For $c \in L$, the cost of reassigning its cluster to $\Pi$ is

$$\text{ran}(c) := \sum_{p \in \text{Cl}(L,c) \setminus \text{Cl}(\Pi,c)} \left(\text{dist}(p, \Pi(p)) - \text{dist}(p, L(p))\right)$$

Claim $\Rightarrow \sum_{c \in L} \text{ran}(c) \leq 2\|vec_{OPT}\|_1$
L_0, L_1, L_{\geq 2}, \text{OPT}_1, \text{OPT}_{\geq 2}

c \in L \text{ may be assigned to } 0, 1, \text{ or } \geq 2 \text{ centers of } \text{OPT}.

L = L_0 \cup L_1 \cup L_{\geq 2}
$L_0, L_1, L_{\geq 2}, OPT_1, OPT_{\geq 2}$

c $\in L$ may be assigned to 0, 1, or $\geq 2$ centers of $OPT$.
$L = L_0 \cup L_1 \cup L_{\geq 2}$

$OPT_1$: subset of $OPT$ assigned to $L_1$

$OPT_{\geq 2}$: subset of $OPT$ assigned to $L_{\geq 2}$

$OPT = OPT_1 \cup OPT_{\geq 2}$
\(L_0, L_1, L_{\geq 2}, \text{OPT}_1, \text{OPT}_{\geq 2}\)

c \in L may be assigned to 0, 1, or \(\geq 2\) centers of OPT.

\(L = L_0 \cup L_1 \cup L_{\geq 2}\)

\(\text{OPT}_1\): subset of OPT assigned to \(L_1\)
\(\text{OPT}_{\geq 2}\): subset of OPT assigned to \(L_{\geq 2}\)

\(\text{OPT} = \text{OPT}_1 \cup \text{OPT}_{\geq 2}\)

For \(o \in \text{OPT}\), \(\text{cost}(o)\) and \(\text{localcost}(o)\) is the cost of \(\text{Cluster}(o, \text{OPT})\) in OPT and L
\[ L_0, L_1, L_{\geq 2}, OPT_1, OPT_{\geq 2} \]

\( c \in L \) may be assigned to 0, 1, or \( \geq 2 \) centers of \( OPT \).
\( L = L_0 \cup L_1 \cup L_{\geq 2} \)

\( OPT_1 \): subset of \( OPT \) assigned to \( L_1 \)
\( OPT_{\geq 2} \): subset of \( OPT \) assigned to \( L_{\geq 2} \)
\( OPT = OPT_1 \cup OPT_{\geq 2} \)

---

For \( o \in OPT \), \( cost(o) \) and \( localcost(o) \) is the cost of \( Cluster(o, OPT) \) in \( OPT \) and \( L \)

**Lemma.** For \( c \in L_0 \) and \( o \in OPT \) we have
\[
localcost(o) \leq ran(c) + cost(o).
\]

**Proof.** Removing \( c \) and adding \( o \) to \( L \) does not improve:
\[
0 \leq ran(c) - localcost(o) + cost(o).
\]
Bounding the contribution of $OPT_{\geq 2}$

Since $|L_1| = |OPT_1|$ (matching) and
$|L_0| + |L_1| + |L_{\geq 2}| = |OPT_1| + |OPT_{\geq 2}| = k$

$$|L_0| = |OPT_{\geq 2}| - |L_{\geq 2}| \geq |OPT_{\geq 2}|/2$$
Bounding the contribution of $OPT_{\geq 2}$

Since $|L_1| = |OPT_1|$ (matching) and
$|L_0| + |L_1| + |L_{\geq 2}| = |OPT_1| + |OPT_{\geq 2}| = k$

$|L_0| = |OPT_{\geq 2}| - |L_{\geq 2}| \geq |OPT_{\geq 2}|/2$

Lemma.

$$\sum_{o \in OPT_{\geq 2}} localcost(o) \leq 2 \sum_{c \in L_0} ran(c) + \sum_{o \in OPT_{\geq 2}} cost(o)$$
Bounding the contribution of $OPT_{\geq 2}$

Since $|L_1| = |OPT_1|$ (matching) and $|L_0| + |L_1| + |L_{\geq 2}| = |OPT_1| + |OPT_{\geq 2}| = k$

$$|L_0| = |OPT_{\geq 2}| - |L_{\geq 2}| \geq |OPT_{\geq 2}|/2$$

**Lemma.**

$$\sum_{o \in OPT_{\geq 2}} \text{localcost}(o) \leq 2 \sum_{c \in L_0} \text{ran}(c) + \sum_{o \in OPT_{\geq 2}} \text{cost}(o)$$

**Proof.** Let $c^* \in L_0$ minimize $\text{ran}(c)$. Earlier lemma:

$$\text{localcost}(o) \leq \text{ran}(c^*) + \text{cost}(o)$$

Summing over $o \in OPT_{\geq 2}$:

$$\sum_{o \in OPT_{\geq 2}} \text{localcost}(o) \leq |OPT_{\geq 2}| \text{ran}(c^*) + \sum_{o \in OPT_{\geq 2}} \text{cost}(o)$$
Bounding the contribution of $OPT_1$

**Lemma.**

\[
\sum_{o \in OPT_1} localcost(o) \leq \sum_{o \in OPT_1} \text{ran}(L(o)) + \sum_{o \in OPT_1} cost(o)
\]
Bounding the contribution of $OPT_1$

**Lemma.**

$$\sum_{o \in OPT_1} \text{localcost}(o) \leq \sum_{o \in OPT_1} \text{ran}(L(o)) + \sum_{o \in OPT_1} \text{cost}(o)$$

**Proof.** $o \in OPT_1$ is assigned to $L(o) = \Pi(o)$.

**Claim:** $\text{localcost}(o) \leq \text{ran}(L(o)) + \text{cost}(o)$.

Replacing $L(o)$ with $o$ in $L$ doesn’t improve.

Potential increased prices in $\text{Cl}(L, L(o)) \cup \text{Cl}(OPT, o)$.

Replace cost in $\left( \text{Cl}(L, L(o)) \setminus \text{Cl}(OPT, o) \right)$ is $\text{ran}(L(o))$.

Replace cost in $\text{Cl}(OPT, o)$ is $\leq -\text{localcost}(o) + \text{cost}(o)$.

$\Rightarrow 0 \leq \text{ran}(L(o)) - \text{localcost}(o) + \text{cost}(o)$. 


Theorem. The local optimum $L$ gives a 5-approximation for $k$-median.
Theorem. The local optimum $L$ gives a 5-approximation for $k$-median.

$$\|vec_L\|_1 = \sum_{o \in OPT_1} localcost(o) + \sum_{o \in OPT_{\geq 2}} localcost(o)$$

$$\leq \sum_{c \in L_0} ran(c) + \sum_{o \in OPT_{\geq 2}} cost(o)$$

$$+ \sum_{o \in OPT_1} ran(L(o)) + \sum_{o \in OPT_1} cost(o)$$

$$\leq 2 \sum_{c \in L} ran(c) + \sum_{o \in OPT} cost(o)$$

$$\leq 4 \|vec_{OPT}\|_1 + \|vec_{OPT}\|_1$$
Theorem. For any $\varepsilon > 0$ the local optimum $L$ wrp. $1 - \tau$-improvements ($\tau := \varepsilon / 10k$) gives a $5 + \varepsilon$-approximation for $k$-median in $O(n^2k^3 \log n/\varepsilon)$ time.
Theorem. For any $\varepsilon > 0$ the local optimum $L$ wrp. $1 - \tau$-improvements ($\tau := \varepsilon/10k$) gives a $5 + \varepsilon$-approximation for $k$-median in $O(n^2k^3 \log \frac{n}{\varepsilon})$ time.

→ Can get $3 + \frac{2}{p}$-approx with $p$-swaps (tight)
Theorem. For any $\varepsilon > 0$ the local optimum $L$ wrp. $1 - \tau$-improvements ($\tau := \varepsilon/10k$) gives a $5 + \varepsilon$-approximation for $k$-median in $O(n^2 k^3 \log n \over \varepsilon)$ time.

→ Can get $3 + 2/p$-approx with $p$-swaps (tight)

Theorem. For any $\varepsilon > 0$ local search gives a $25 + \varepsilon$-approximation for $k$-means in $O(n^2 k^3 \log n \over \varepsilon)$ time.

→ Can get $(3 + 2/p)^2$-approx with $p$-swaps (tight)
$k$-median, $k$-means in $\mathbb{R}^d$
\( k \)-median, \( k \)-means in \( \mathbb{R}^d \)

\( k \)-median is NP-hard if \( k, d \) both in input. (Guruswami–Indyk 2003), but if at least one is constant, there is a PTAS.

For \( k \)-means with constant \( d \), local search with \((1/\varepsilon)^{\Theta(1)}\)-swaps gives PTAS. (e.g. Cohen-Addad et al. 2019)
Next week:
SoCG 2020! Check it out.