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

Sándor Kisfaludi-Bak

Computaional Geometry
Summer semester 2020

max planck institut
informatik

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

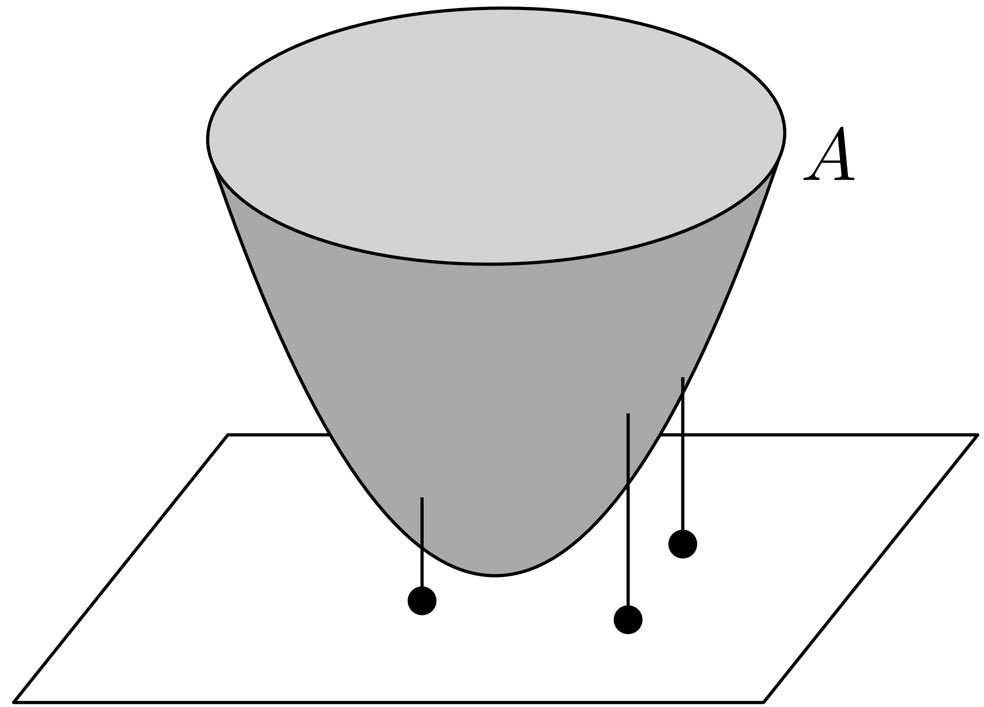
Overview

- 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 $A : z = x^2 + y^2$

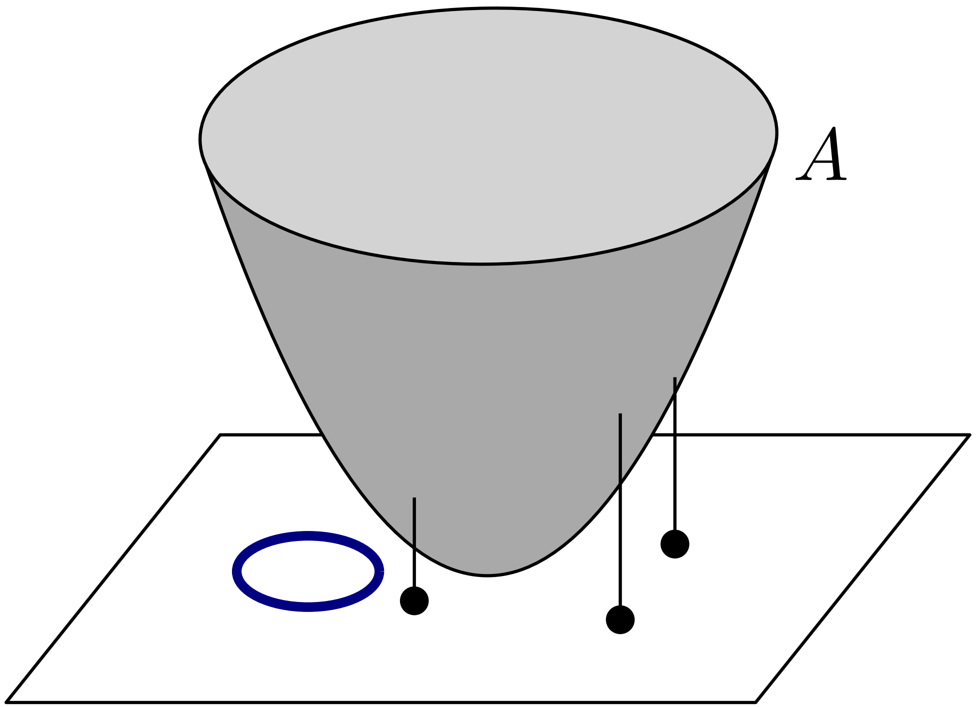


Lifting to a paraboloid

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

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

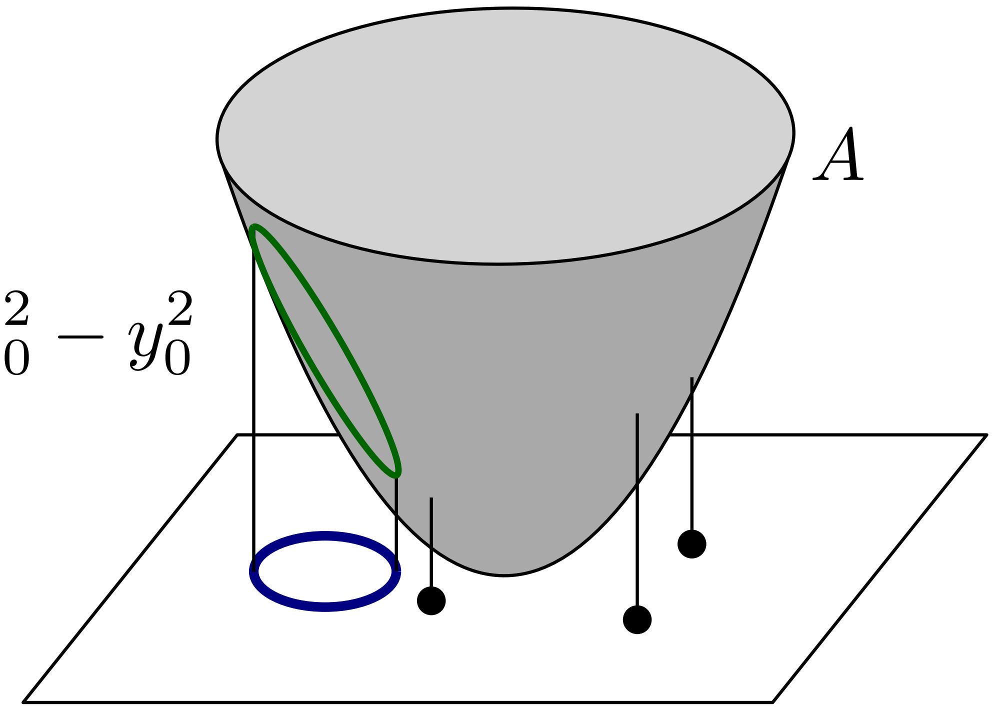
$$L(x, y) = (x,\ y,\ x^2 + y^2)$$

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

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

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

$$\begin{aligned} x^2 + y^2 &= r^2 + 2xx_0 + 2yy_0 - x_0^2 - y_0^2 \\ &= \alpha_1 x + \alpha_2 y + c \end{aligned}$$



Lifting to a paraboloid

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

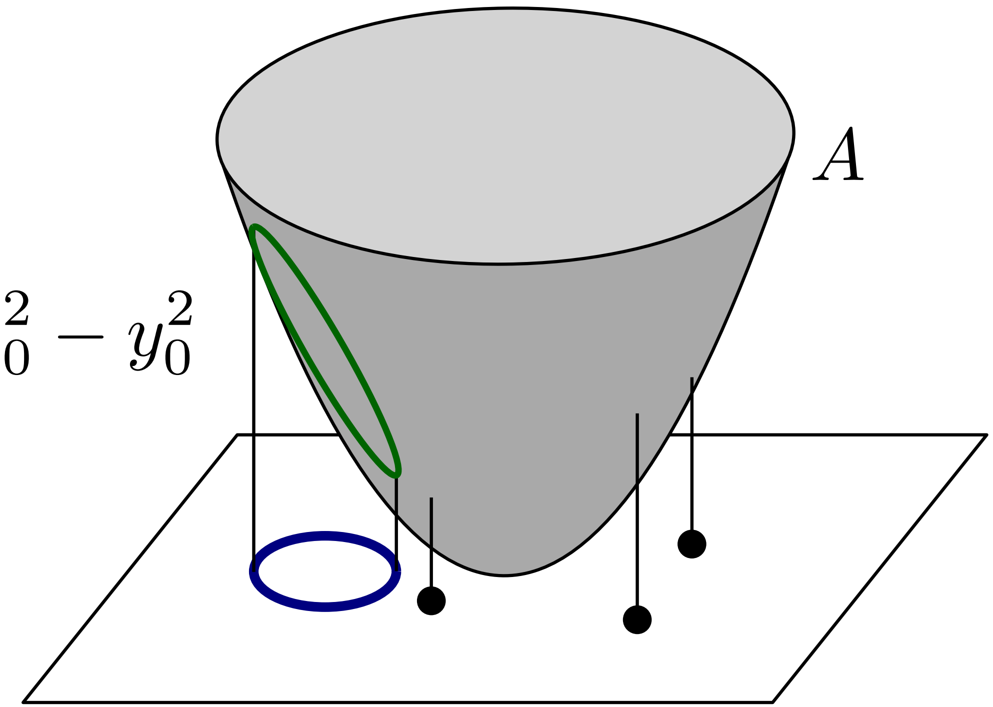
L projects (x, y) vertically up to the paraboloid $A : z = x^2 + y^2$

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

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

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

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



$$L(\gamma) \subset H_\gamma := \{(x, y, z) \mid -\alpha_1 x - \alpha_2 y + z = c\}$$

Lifting an empty circumcircle

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

$$\Updownarrow$$

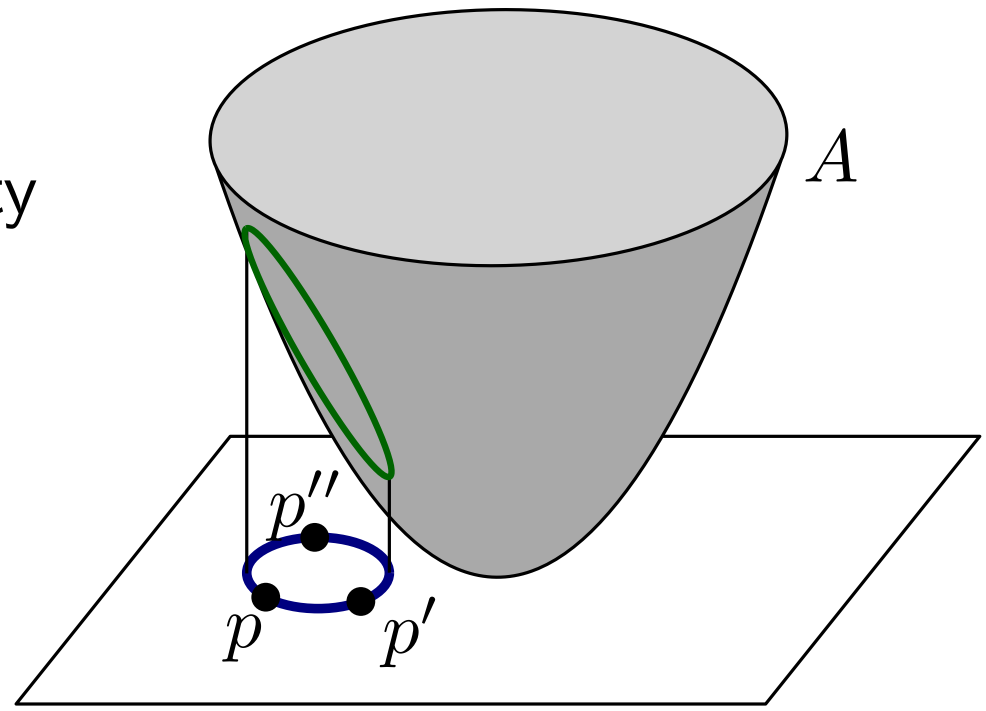
$\gamma =$ circumcircle of $pp'p''$ is empty

$$\Updownarrow$$

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

$$\Updownarrow$$

H_γ is a face of $\mathrm{conv}^\downarrow(L(P))$



Lifting an empty circumcircle

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

$$\Updownarrow$$

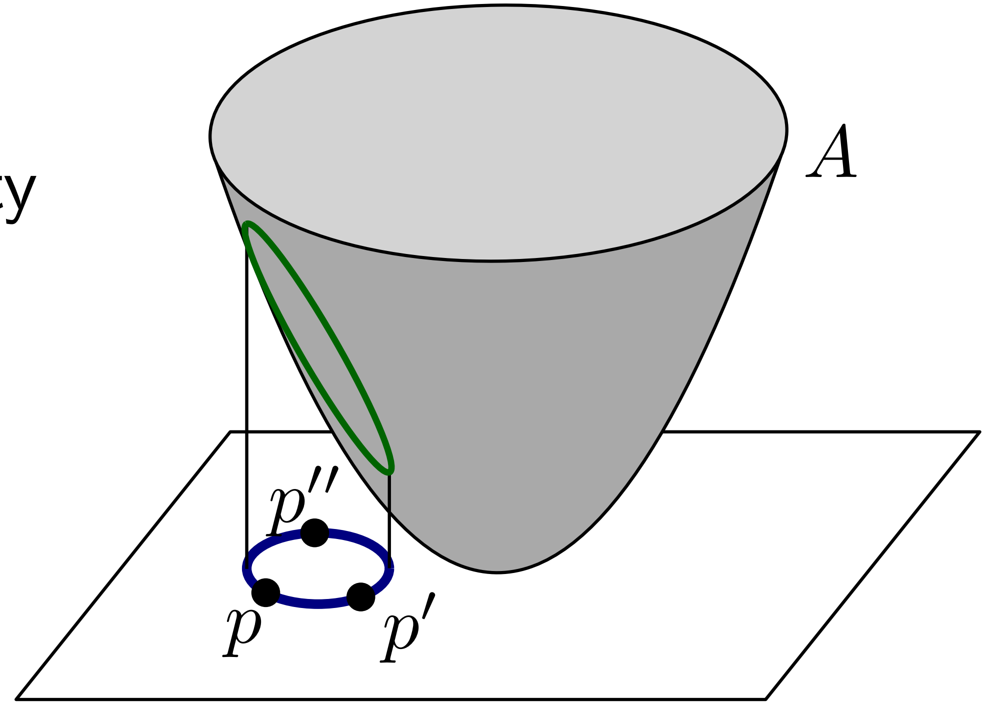
$\gamma =$ circumcircle of $pp'p''$ is empty

$$\Updownarrow$$

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

$$\Updownarrow$$

H_γ is a face of $\mathrm{conv}^\downarrow(L(P))$

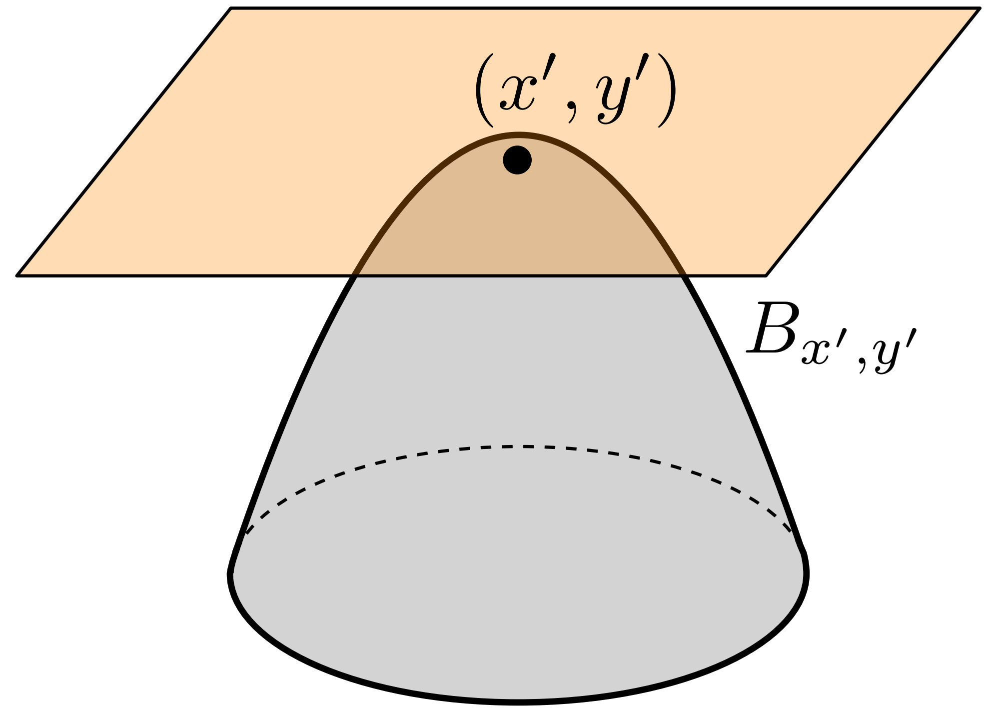


$$DT(P) = \mathrm{proj}_{z=0}(\mathrm{conv}^\downarrow(L(P)))$$

Lifting a paraboloid

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

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

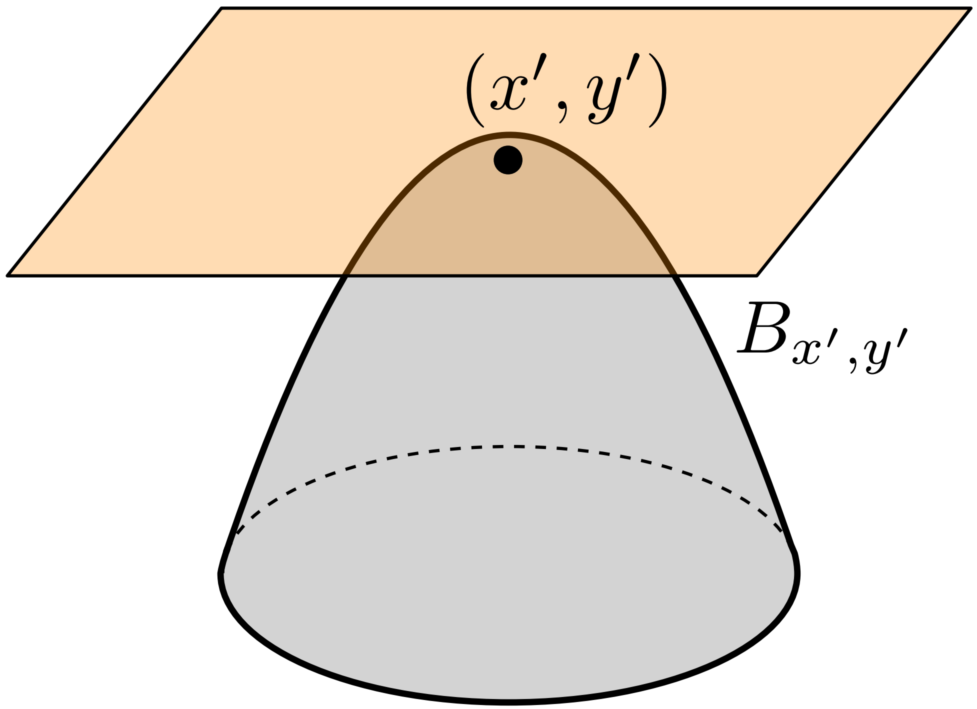


$$B_{x',y'} = \{(x, y, z) \mid z = -(x - x')^2 - (y - y')^2\}$$

Lifting a paraboloid

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

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



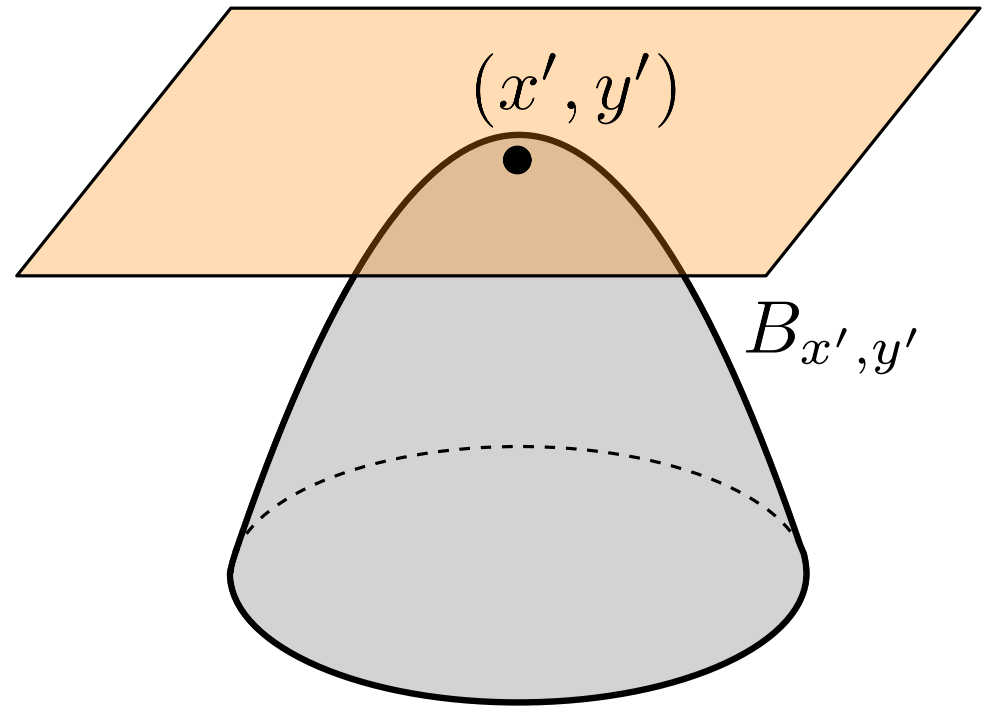
$$B_{x',y'} = \{(x, y, z) \mid z = -(x - x')^2 - (y - y')^2\}$$

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

Lifting a paraboloid

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

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



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

$$L(x, y, -(x' - x)^2 - (y' - y)^2) = \underbrace{(x, y, x'^2 + y'^2 + 2x'x + 2y'y)}_{\text{a plane!}}$$

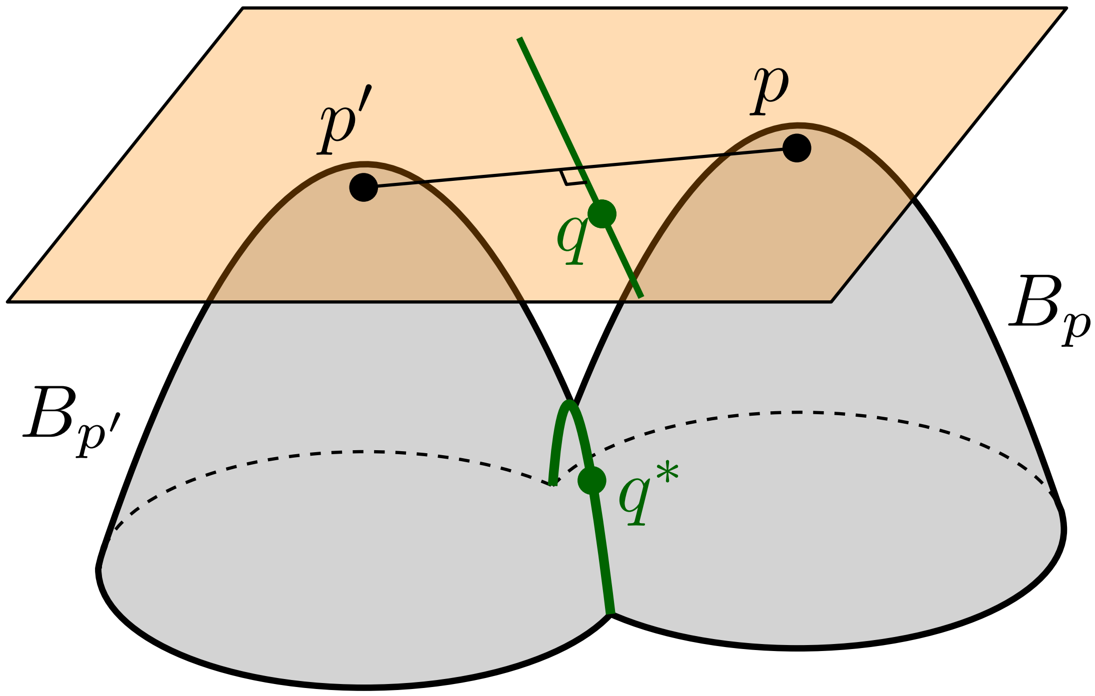
a plane!

touches A at $L(x', y')$

Lifting many paraboloids: Voronoi

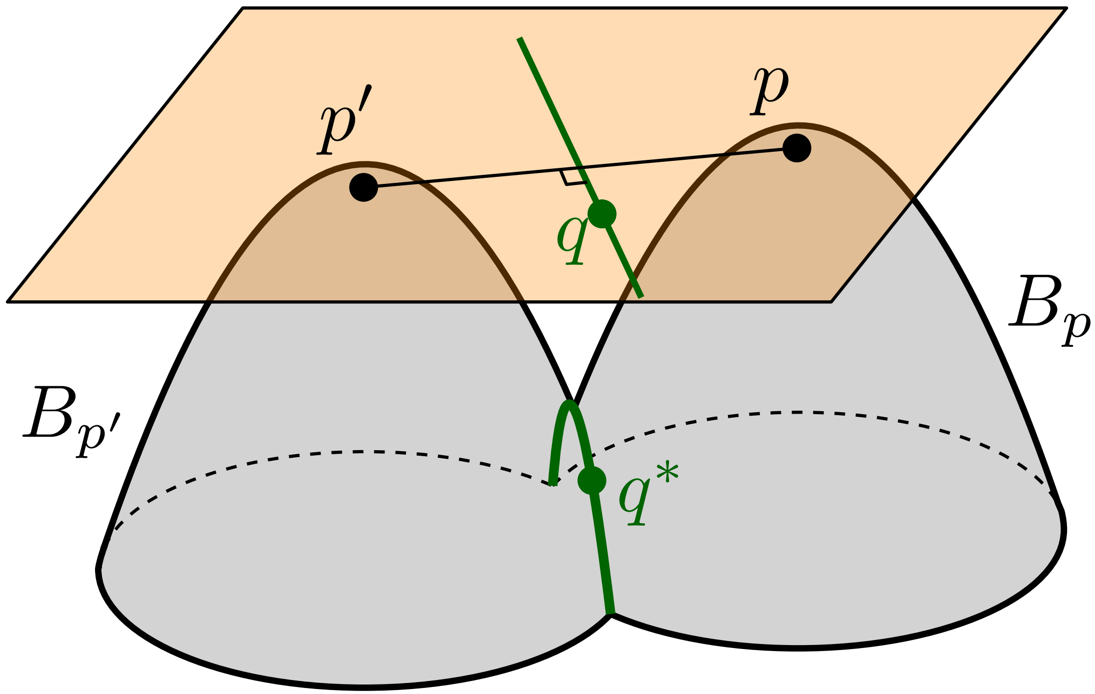
Opaque hanging paraboloid B_p for each $p \in P$.

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



Lifting many paraboloids: Voronoi

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



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

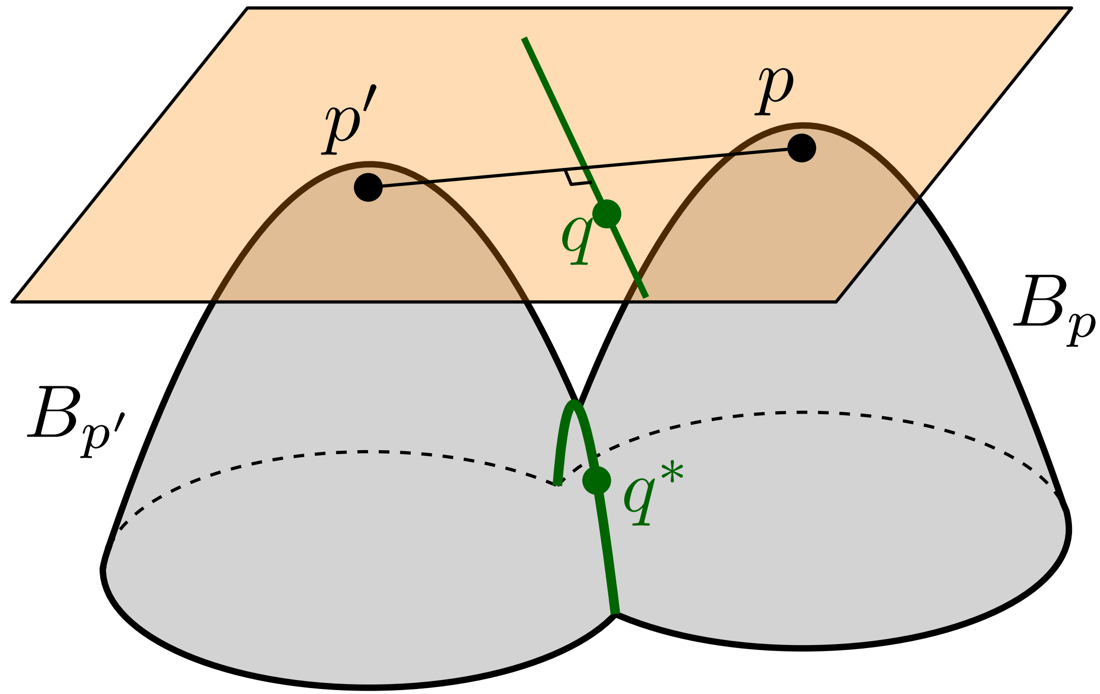
upper envelope of $\bigcup_{p \in P} B_p$ looks like Vor(P) from $(0, 0, \infty)$

Apply $L(.)$: polyhedron $\widehat{B}$ 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 $p \in P$.



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

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

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

$$\text{Vor}(P) = \mathrm{proj}_{z=0}(\widehat{B}) = \mathrm{proj}_{z=0}\left(\bigcap_{p \in P} \text{touchplane}_A(L(p))^{\uparrow}\right)$$

Voronoi and Delaunay in higher dimensions?

Paraboloid lifting works in $\mathbb{R}^d$.

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

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 $O(n^{\lceil d/2 \rceil})$

Voronoi and Delaunay in higher dimensions?

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

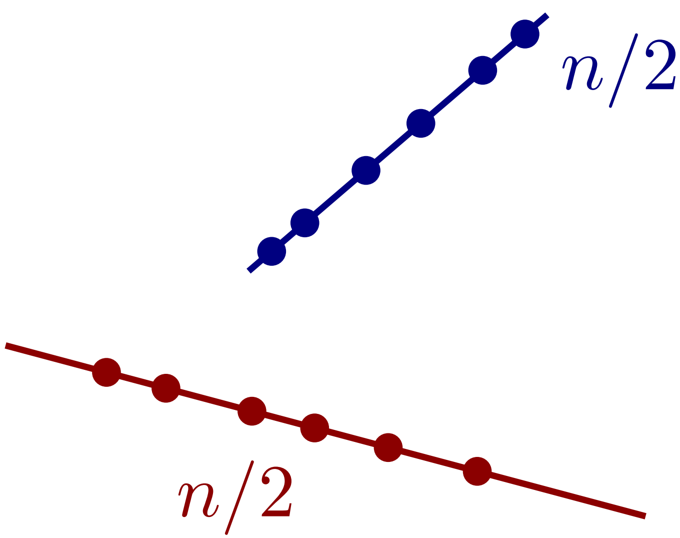
- Vor(P) and $DT(P)$ in $\mathbb{R}^d$ have complexity $O(n^{\lceil d/2 \rceil})$
- Vor(P) and $DT(P)$ in $\mathbb{R}^d$ can be computed by convex hull algorithm in $\mathbb{R}^{d+1}$

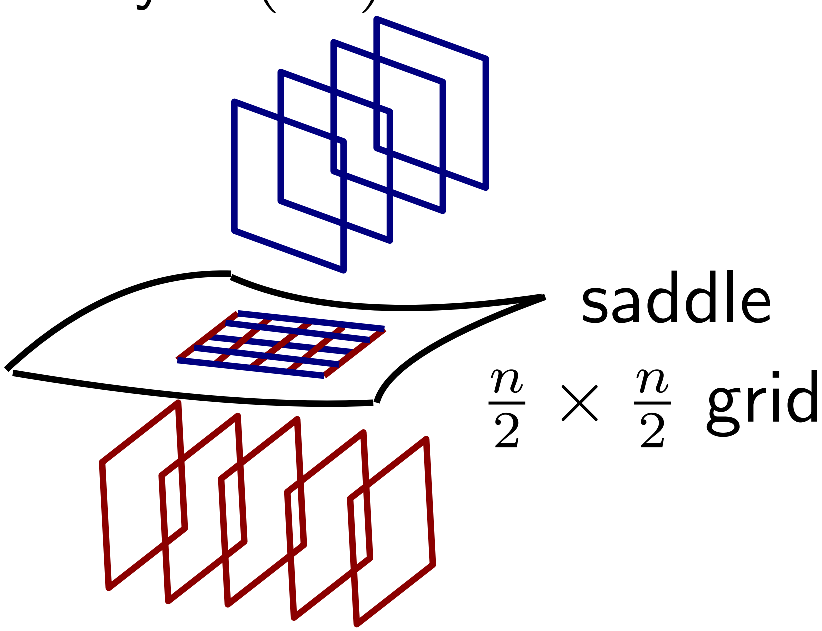
Voronoi and Delaunay in higher dimensions?

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

Metric spaces and clustering

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

- $\mathrm{dist}(a, b) = \mathrm{dist}(b, a)$ (symmetric)
- $\mathrm{dist}(a, b) = 0 \Leftrightarrow a = b$
- $\mathrm{dist}(a, b) + \mathrm{dist}(b, c) \geq \mathrm{dist}(a, c)$ (triangle ineq.)

Metric spaces and clustering

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

- $\mathrm{dist}(a, b) = \mathrm{dist}(b, a)$ (symmetric)
- $\mathrm{dist}(a, b) = 0 \Leftrightarrow a = b$
- $\mathrm{dist}(a, b) + \mathrm{dist}(b, c) \geq \mathrm{dist}(a, c)$ (triangle ineq.)

Clustering:

given data, find similar entries and put them together

Metric spaces and clustering

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

- $\text{dist}(a, b) = \text{dist}(b, a)$ (symmetric)
- $\text{dist}(a, b) = 0 \Leftrightarrow a = b$
- $\text{dist}(a, b) + \text{dist}(b, c) \geq \text{dist}(a, c)$ (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.

$$vec_C := \Big(\text{dist}(p_1, C), \text{dist}(p_2, C), \ldots, \text{dist}(p_n, C)\Big)$$

is

"small"

Clustering variants

- k-center:

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

"minimize the max distance to nearest center"

Clustering variants

- k-center:

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

"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:

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

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

- k-median:

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

"minimize sum of distances to nearest center"

- k-means:

$$\min_{C \subset X, |C|=k} \|vec_C\|_2 = \min_{C \subset X, |C|=k} \sqrt{\sum_{p \in P} \big(\mathrm{dist}(p, C)\big)^2}$$

"minimize sum of squared distances to nearest center"

Clustering variants

- k-center:

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

"minimize the max distance to nearest center"
a.k.a. cover X with ... imizing r

$C \subseteq P$: discrete clustering

$C \subseteq X$: continuous clustering

- k-median:

$$\min_{C \subset \cancel{X}\, P,\, |C|=k} \|ve \ldots \operatorname{dist}(p, C)$$

"minimize sum of distances to nearest center"

- k-means:

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

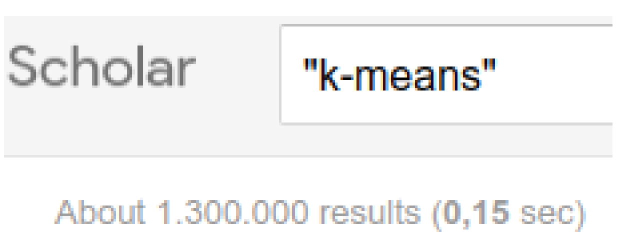
"minimize sum of squared distances to nearest center"

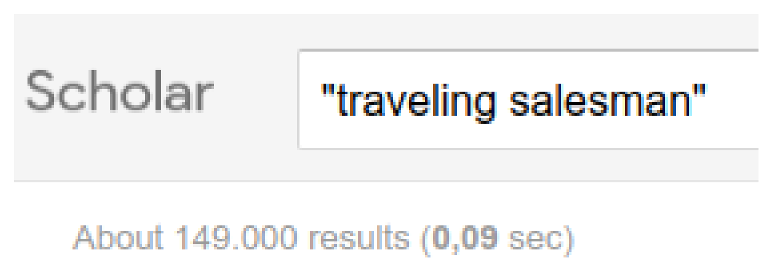
Facility location

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

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

"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 $P = NP$.

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

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 $k \to k+1$.

Hardness of k-center

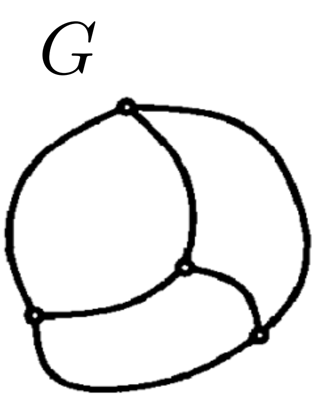
Theorem (Feder–Greene 1988). There is no polynomial time 1.8-approximation for k-center in $\mathbb{R}^2$, unless $P = NP$.

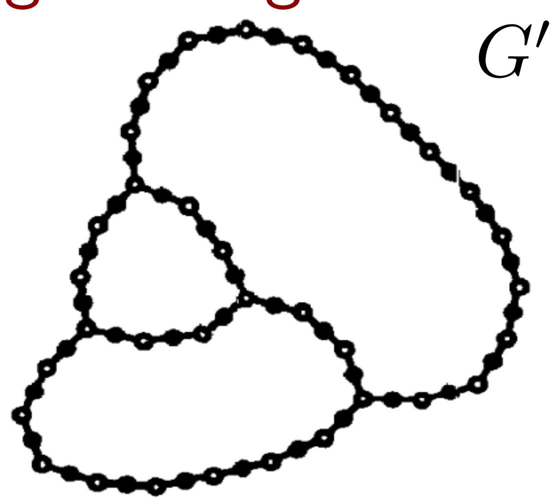
Reduction from planar vertex cover of max degree 3

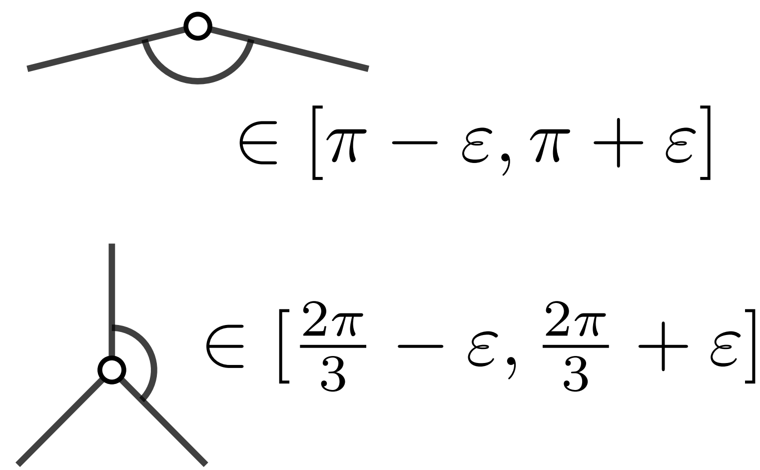
Double subdivision:

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

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

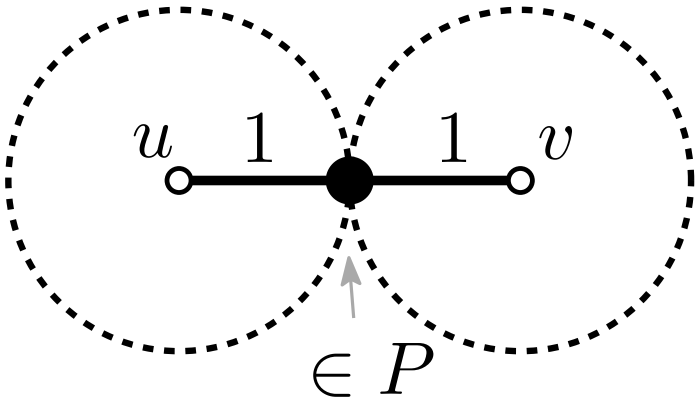






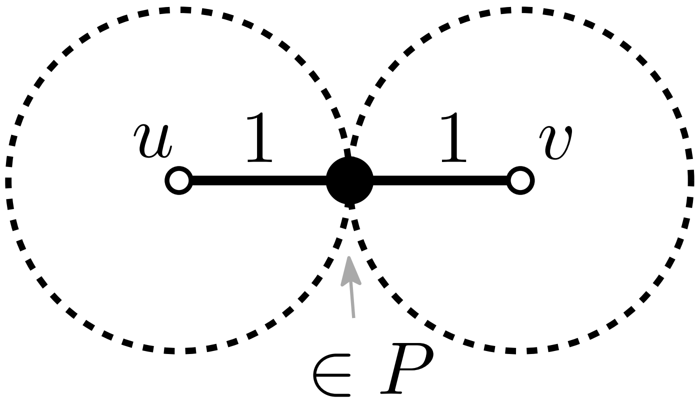
Hardness of k-center: disk radii

$P :=$ edge midpoints of smooth drawing of G'



Hardness of k-center: disk radii

$P :=$ edge midpoints of smooth drawing of G'



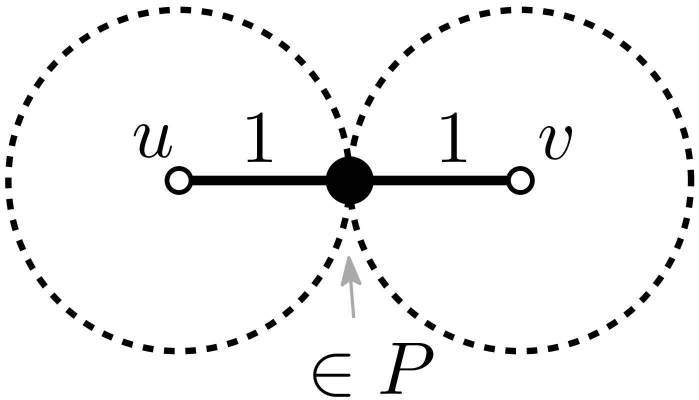
$\exists$ VC of size k in G'

$\Leftrightarrow$

$\exists$ k-center with radius 1

Hardness of k-center: disk radii

$P :=$ edge midpoints of smooth drawing of G'



$\exists$ VC of size k in G'

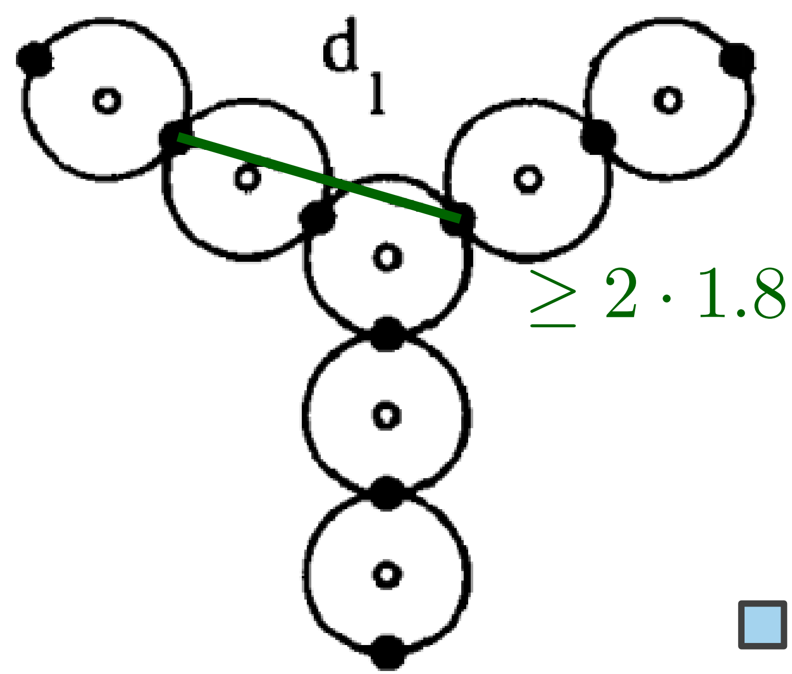
$\Leftrightarrow$

$\exists$ k-center with radius 1

Otherwise needs ≥ 1 disk covering 2 non-neighbors u, v

$\mathrm{dist}(u, v) \geq 2 \cdot 1.8$

$\Rightarrow r \geq 1.8$



Greedy centers

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

Greedy clustering:
start with arbitrary $c_1 \in P$.
For $i = 2, \ldots, k$:
  Let $c_i =$ 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 $\mathrm{dist}(q, C)$ is maximized.

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

Let $r_i = \max_{p \in P} \mathrm{dist}(p, \{c_1, \ldots, c_i\})$.
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 $\mathrm{dist}(q, C)$ is maximized.

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

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

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} := point realizing r_k

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

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

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} := point realizing r_k

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

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

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

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

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

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

$\Downarrow$

each ball in opt has ≤ 1 pt from $c_1, \ldots, c_{k+1}$

$\lightning$ $\square$

r-packing from greedy

Definition. $S \subset X$ is an r-packing if

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

r-packing from greedy

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 $s, s' \in S$

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

$\mathbb{R}^2$	$n^{O(\sqrt{k})}$	or $2^{O(\sqrt{n})}$
$\mathbb{R}^d, d = \text{const.}$	no $n^{o(k)}$ known	$2^{O(n^{1-1/d})}$

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

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

$\mathbb{R}^2$	$n^{O(\sqrt{k})}$	or $2^{O(\sqrt{n})}$
$\mathbb{R}^d, d = \text{const.}$	no $n^{o(k)}$ known	$2^{O(n^{1-1/d})}$

"optimal"

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

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

$\mathbb{R}^2$	$n^{O(\sqrt{k})}$	or $2^{O(\sqrt{n})}$
$\mathbb{R}^d, d = \text{const.}$	no $n^{o(k)}$ known	$2^{O(n^{1-1/d})}$

"optimal"

Fix r, approximate k instead:

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

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

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

$\mathbb{R}^2$	$n^{O(\sqrt{k})}$	or $2^{O(\sqrt{n})}$
$\mathbb{R}^d, d = \text{const.}$	no $n^{o(k)}$ known	$2^{O(n^{1-1/d})}$

"optimal"

Fix r, approximate k instead:

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

$$\|vec_C\|_1 \leq n\|vec_C\|_\infty$$

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

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

  $\Rightarrow$ Results in local opt center set L

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

$$\|vec_C\|_1 \leq n\|vec_C\|_\infty$$

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

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

  $\Rightarrow$ Results in local opt center set L

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

$$O((nk)^2 \log_{\frac{1}{1-\tau}} 2n) = O((nk)^2 \log_{1+\tau} n)$$

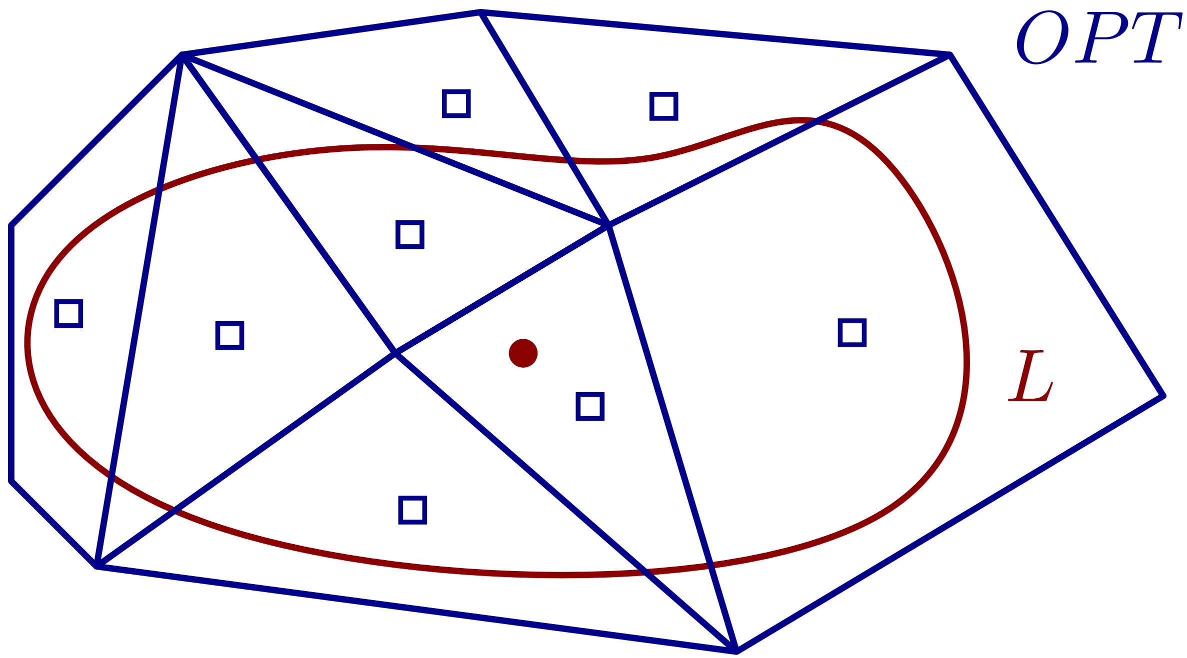
$$= O((nk)^2 \cdot 10k \log n) = O(k^3 n^2 \log n)$$

k-median: quality of approximation

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

Challange: L and OPT may be very different.

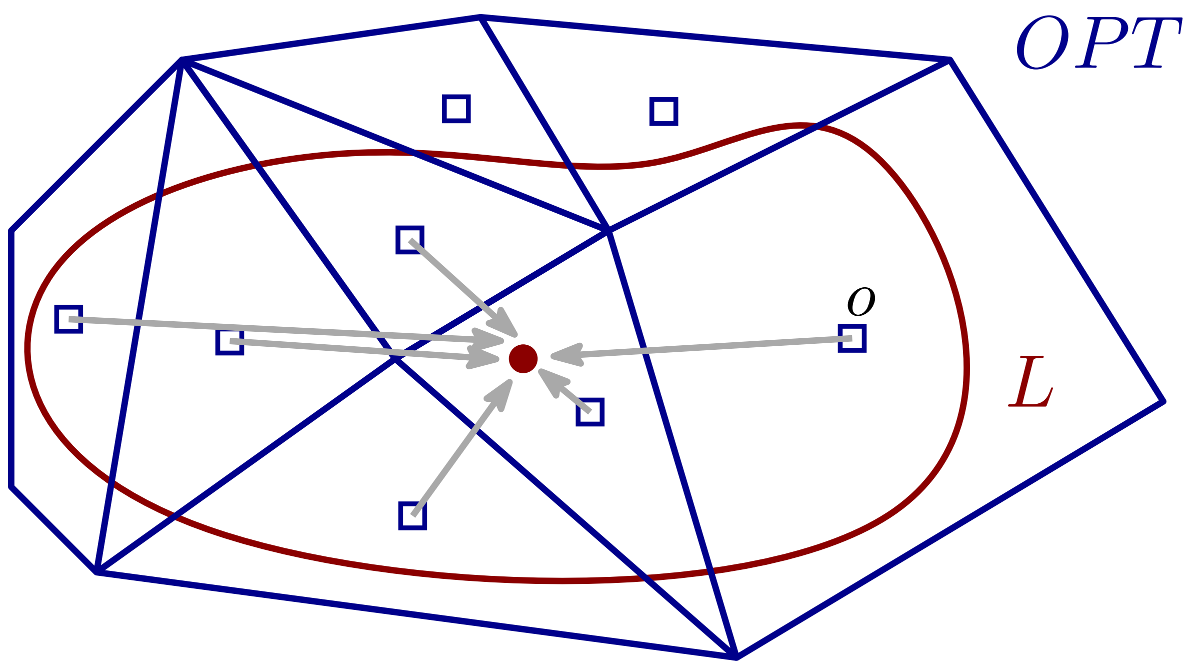
Idea: use "intermediate" clustering Π to relate them



k-median: quality of approximation

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

Challange: L and OPT may be very different.
Idea: use "intermediate" clustering Π to relate them



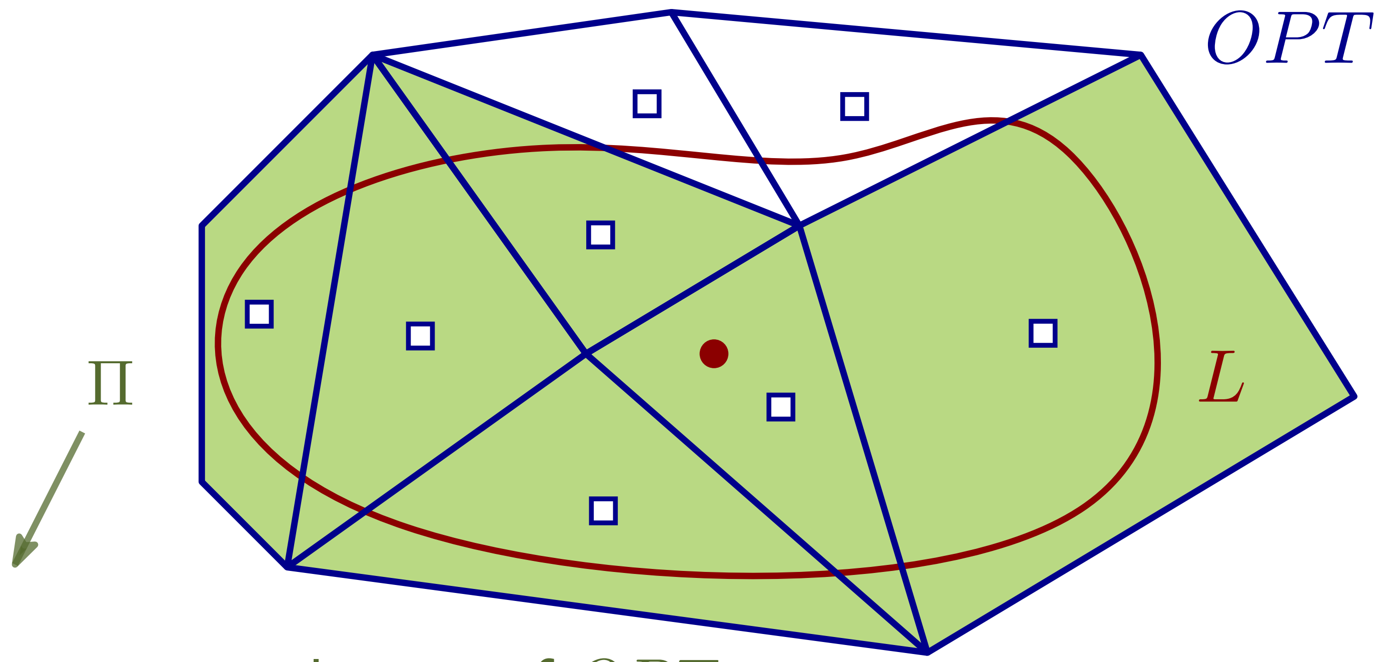
assign cluster of center $o \in OPT$ to $nn(o, L)$

k-median: quality of approximation

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

Challange: L and OPT may be very different.

Idea: use "intermediate" clustering Π to relate them



like L, but respects clusters of OPT

Cost of moving from L to Π

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

Cost of moving from L to Π

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

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

$$
\begin{aligned}
\mathrm{dist}(p, \Pi(p)) &\leq \mathrm{dist}(p, OPT(p)) + \mathrm{dist}(OPT(p), \Pi(p)) \\
&\leq \mathrm{dist}(p, OPT(p)) + \mathrm{dist}(OPT(p), L(p)) \\
&\leq \mathrm{dist}(p, OPT(p)) + \mathrm{dist}(OPT(p), p) \\
&\qquad + \mathrm{dist}(p, L(p)) \\
&= 2\mathrm{dist}(p, OPT(p)) + \mathrm{dist}(p, L(p)) \qquad \square
\end{aligned}
$$

Cost of moving from L to Π

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

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

$$\begin{aligned}
\mathrm{dist}(p, \Pi(p)) &\leq \mathrm{dist}(p, OPT(p)) + \mathrm{dist}(OPT(p), \Pi(p)) \\
&\leq \mathrm{dist}(p, OPT(p)) + \mathrm{dist}(OPT(p), L(p)) \\
&\leq \mathrm{dist}(p, OPT(p)) + \mathrm{dist}(OPT(p), p) \\
&\qquad + \mathrm{dist}(p, L(p)) \\
&= 2\mathrm{dist}(p, OPT(p)) + \mathrm{dist}(p, L(p)) \qquad \square
\end{aligned}$$

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

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

claim $\Rightarrow$ $\sum_{c \in L} ran(c) \leq 2\|vec_{OPT}\|_1$

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

$c \in L$ may be assigned to 0, 1 , or ≥ 2 centers of 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 ≥ 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 $L_{\geq 2}$
$OPT = OPT_1 \cup OPT_{\geq 2}$

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

$c \in L$ may be assigned to 0, 1 , or ≥ 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

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

$c \in L$ may be assigned to 0, 1 , or ≥ 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 $|L_1| = |OPT_1|$ (mathcing) 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 $|L_1| = |OPT_1|$ (mathcing) 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 $|L_1| = |OPT_1|$ (mathcing) 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.

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

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

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

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

$$\sum_{o \in OPT_{\geq 2}} localcost(o) \leq |OPT_{\geq 2}| ran(c^*) + \sum_{o \in OPT_{\geq 2}} cost(o)$$

□

Bounding the contribution of OPT_1

Lemma.

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

Bounding the contribution of OPT_1

Lemma.

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

Proof. $o \in OPT_1$ is assigned to $L(o) = \Pi(o)$.
Claim: $localcost(o) \leq ran(L(o)) + cost(o)$.
Replacing $L(o)$ with o in L doesn't improve.

Potential increased prices in $Cl(L, L(o)) \cup Cl(OPT, o)$.
Replace cost in $\Big(Cl(L, L(o)) \setminus Cl(OPT, o)\Big)$ is $ran(L(o))$.
Replace cost in $Cl(OPT, o)$ is $\leq -localcost(o) + cost(o)$.

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

□

k-median approximation quality wrap-up

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

k-median approximation quality wrap-up

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

$$\begin{aligned}
\|vec_L\|_1 &= \sum_{o \in OPT_1} localcost(o) + \sum_{o \in OPT_{\geq 2}} localcost(o) \\
&\leq \sum_{c \in L_0} ran(c) + \sum_{o \in OPT_{\geq 2}} cost(o) \\
&\quad + \sum_{o \in OPT_1} ran(L(o)) + \sum_{o \in OPT_1} cost(o) \\
&\leq 2 \sum_{c \in L} ran(c) + \sum_{o \in OPT} cost(o) \\
&\leq 4\|vec_{OPT}\|_1 + \|vec_{OPT}\|_1
\end{aligned}$$

□

k-median, k-means with local search

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 $O(n^2 k^3 \frac{\log n}{\varepsilon})$ time.

k-median, k-means with local search

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 $O(n^2 k^3 \frac{\log n}{\varepsilon})$ time.

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

k-median, k-means with local search

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 $O(n^2k^3\frac{\log n}{\varepsilon})$ 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 $O(n^2k^3\frac{\log n}{\varepsilon})$ 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 cosntant, 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.