Dimension reduction, embeddings

Sándor Kisfaludi-Bak

Computaional Geometry
Summer semester 2020

max planck institut
informatik

Overview

- Embeddings, distortion, Johnson-Lindenstrauss

Overview

- Embeddings, distortion, Johnson-Lindenstrauss
- Random partitions

Overview

- Embeddings, distortion, Johnson-Lindenstrauss
- Random partitions
- Embedding into HSTs

Overview

- Embeddings, distortion, Johnson-Lindenstrauss
- Random partitions
- Embedding into HSTs
- Further embeddings into Euclidean space

Embeddings, distortion

Definition. An embedding f from the metric space (X, dist_X) to (Y, dist_Y) is a K-bi-Lipschitz if there exists a $c > 0$ such that for all $x, x' \in X$ we have

$$c\mathrm{dist}_X(x, x') \leq \mathrm{dist}_Y(f(x), f(x')) \leq cK\mathrm{dist}_X(x, x').$$

Embeddings, distortion

Definition. An embedding f from the metric space (X, dist_X) to (Y, dist_Y) is a K-bi-Lipschitz if there exists a $c > 0$ such that for all $x, x' \in X$ we have

$$c\mathrm{dist}_X(x, x') \leq \mathrm{dist}_Y(f(x), f(x')) \leq cK\mathrm{dist}_X(x, x').$$

Definition. The distortion of an embedding $f : X \to Y$ is the smallest Δ s.t. f is Δ-bi-Lipscchitz.

Embeddings, distortion

Definition. An embedding f from the metric space (X, dist_X) to (Y, dist_Y) is a K-bi-Lipschitz if there exists a $c > 0$ such that for all $x, x' \in X$ we have

$$c\mathrm{dist}_X(x, x') \leq \mathrm{dist}_Y(f(x), f(x')) \leq cK\mathrm{dist}_X(x, x').$$

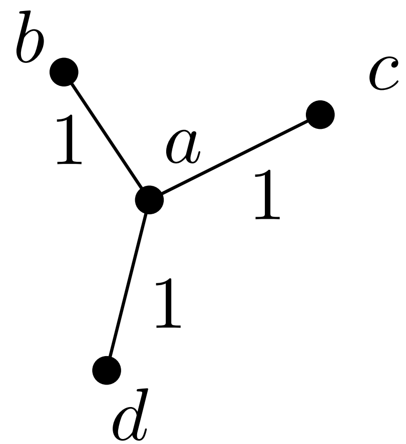
Definition. The distortion of an embedding $f : X \to Y$ is the smallest Δ s.t. f is Δ-bi-Lipscchitz.

If $Y = \mathbb{R}^d$, then we want

$$\mathrm{dist}(x, x') \leq \|f(x) - f(x')\|_2 \leq \Delta\mathrm{dist}(x, x')$$

Why distortion is necessary

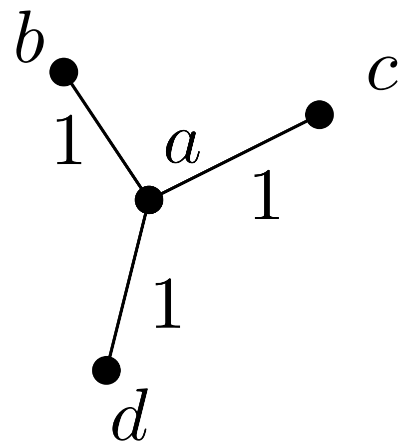
Take $Y = \mathbb{R}^d$, and

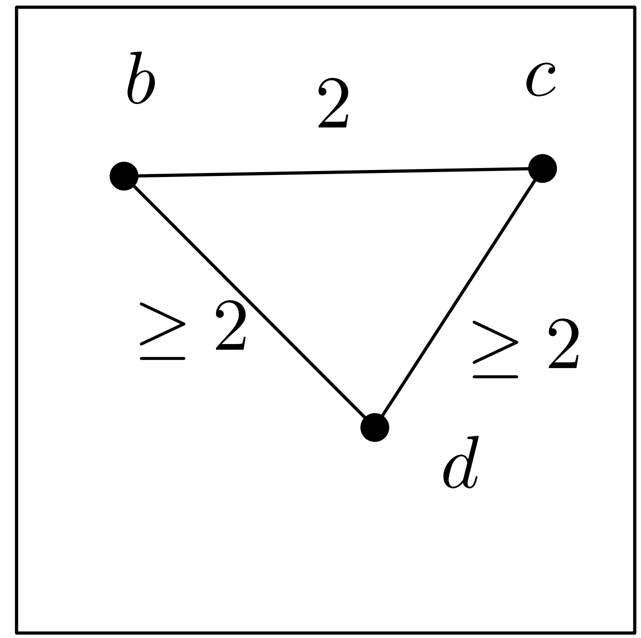
$X =$ 

Why distortion is necessary

Take $Y = \mathbb{R}^d$, and

$X =$





Where to put a?

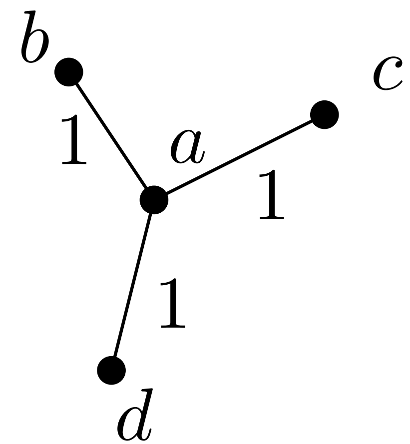
$$\min(\max\{\|a-b\|, \|a-c\|, \|a-d\|\})$$

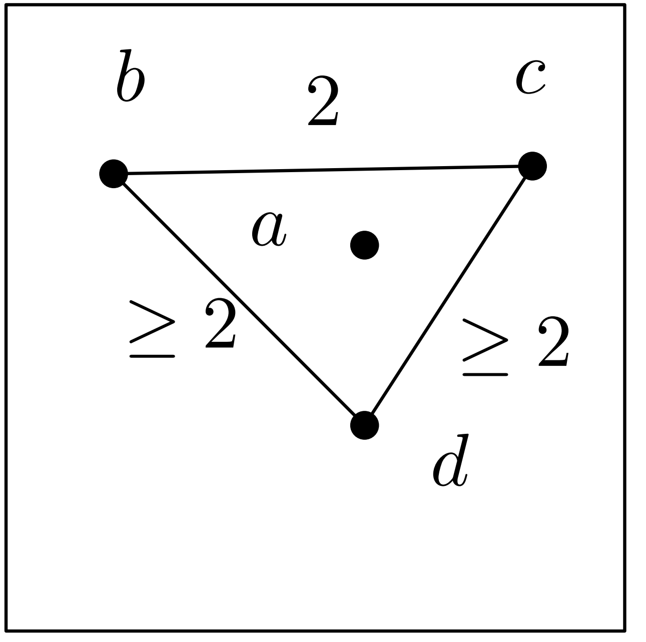
attained when a is circumcenter

Why distortion is necessary

Take $Y = \mathbb{R}^d$, and

$X =$





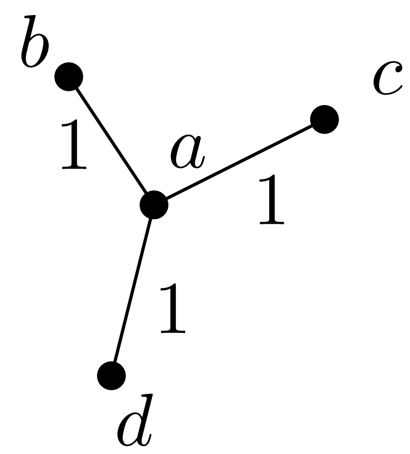
Where to put a?

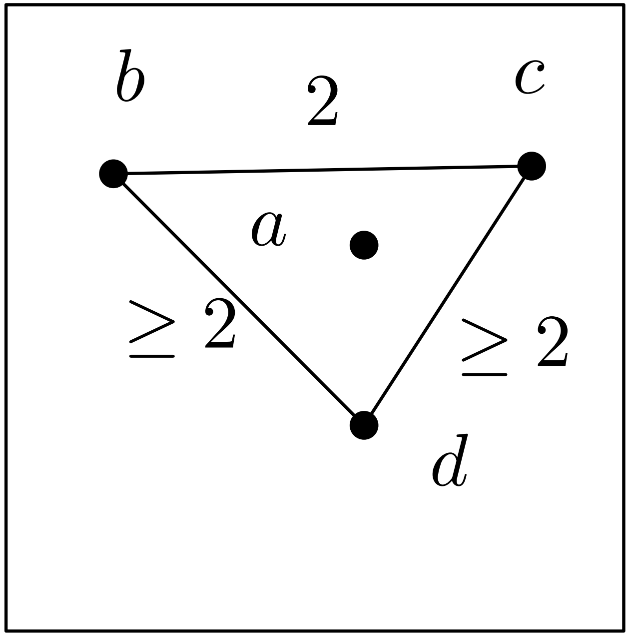
$$\min(\max\{\|a-b\|, \|a-c\|, \|a-d\|\})$$

attained when a is circumcenter

Why distortion is necessary

Take $Y = \mathbb{R}^d$, and

$X =$ 



Where to put a?

$$\min(\max\{\|a-b\|, \|a-c\|, \|a-d\|\})$$

attained when a is circumcenter

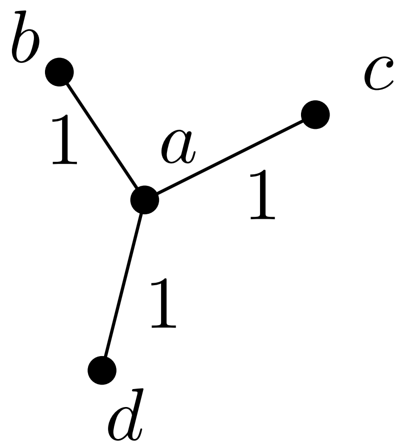
... and when bcd is equilateral of sidelength 2.

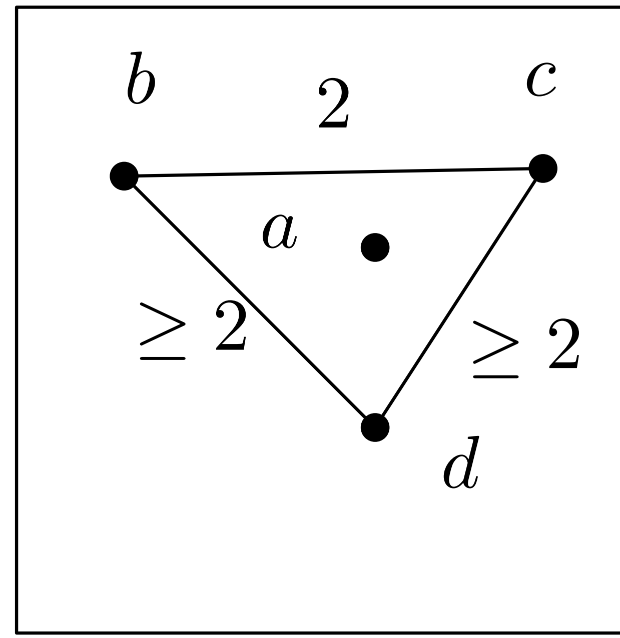
Distortion is $\|b-a\|/\mathrm{dist}_X(a,b) = 2/\sqrt{3}$

Why distortion is necessary

Take $Y = \mathbb{R}^d$, and

$X =$





Where to put a?

$\min(\max\{\|a-b\|, \|a-c\|, \|a-d\|\})$

attained when a is circumcenter

... and when bcd is equilateral of sidelength 2.

Distortion is $\|b-a\|/\mathrm{dist}_X(a,b) = 2/\sqrt{3}$

In general, n-star needs distortion $\Omega(n^{1/d})$ when $Y = \mathbb{R}^d$

The Johnson-Lindenstrauss Lemma

Theorem (Johnson, Lindenstrauss 1984) Given n points $P \subseteq \mathbb{R}^{n-1}$ and $\varepsilon \in (0,1]$, there is an embedding $f : P \to \mathbb{R}^d$ with distortion $1+\varepsilon$ where $d = O(\frac{\log n}{\varepsilon^2})$.

a.k.a. "dimension reduction", "JL lemma"

The Johnson-Lindenstrauss Lemma

Theorem (Johnson, Lindenstrauss 1984) Given n points $P \subseteq \mathbb{R}^{n-1}$ and $\varepsilon \in (0, 1]$, there is an embedding $f : P \to \mathbb{R}^d$ with distortion $1 + \varepsilon$ where $d = O(\frac{\log n}{\varepsilon^2})$.

a.k.a. "dimension reduction", "JL lemma"

- works for $\mathbb{R}^{any}$
- f can be: orthogonal projection to random d-subspace
- can be derandomized (Engebretsen et al. 2002)

Almost equidistant set in $\mathbb{R}^{O(\log n)}$

Let $e_i = (0, \ldots, 0, \underset{i}{1}, 0 \ldots, 0)$.

The set $e_1, \ldots, e_n$ is *equidistant*. (unit simplex).

Can't be embedded isometrically into $\mathbb{R}^d$ if $d < n - 1$. But!

Almost equidistant set in $\mathbb{R}^{O(\log n)}$

Let $e_i = (0, \ldots, 0, \underset{i}{1}, 0 \ldots, 0)$.

The set $e_1, \ldots, e_n$ is *equidistant*. (unit simplex).

Can't be embedded isometrically into $\mathbb{R}^d$ if $d < n - 1$. But!

Folklore. For any fixed $\varepsilon > 0$, there is a set P of n points in $\mathbb{R}^{O(\log n)}$ s.t. $\|p - p'\|_2 \in [1, 1 + \varepsilon]$ for all $p, p' \in P$.

Almost equidistant set in $\mathbb{R}^{O(\log n)}$

Let $e_i = (0, \ldots, 0, \underset{i}{1}, 0 \ldots, 0)$.

The set $e_1, \ldots, e_n$ is *equidistant*. (unit simplex).

Can't be embedded isometrically into $\mathbb{R}^d$ if $d < n - 1$. But!

Folklore. For any fixed $\varepsilon > 0$, there is a set P of n points in $\mathbb{R}^{O(\log n)}$ s.t. $\|p - p'\|_2 \in [1, 1 + \varepsilon]$ for all $p, p' \in P$.

Proof. Use JL lemma on simplex above. □

Random partitions

Partitions, probabilistic partitions

Goal: partition (X, dist) into clusters of diameter at most Δ, s.t. $x, y \in X$ are in the same cluster iff $\mathrm{dist}(x, y) \leq \Delta$.

Partitions, probabilistic partitions

Goal: partition (X, dist) into clusters of diameter at most Δ, s.t. $x, y \in X$ are in the same cluster iff $\mathrm{dist}(x, y) \leq \Delta$.

Clearly unattainable!

Partitions, probabilistic partitions

Goal: partition (X, dist) into clusters of diameter at most Δ, s.t. $x, y \in X$ are in the same cluster iff $\mathrm{dist}(x, y) \leq \Delta$.

Clearly unattainable!

$\mathcal{P}_X$: set of all partitions of X. Pick a random partition $\Pi \in \mathcal{P}_X$ from some distribution $\mathcal{D}$ over $\mathcal{P}_X$.

Revised goal: $\Pr(x, x'$ are separated in $\Pi)$ is small if $\mathrm{dist}(x, x')$ is small.

Partitions, probabilistic partitions

Goal: partition (X, dist) into clusters of diameter at most Δ, s.t. $x, y \in X$ are in the same cluster iff $\mathrm{dist}(x, y) \leq \Delta$.

Clearly unattainable!

$\mathcal{P}_X$: set of all partitions of X. Pick a random partition $\Pi \in \mathcal{P}_X$ from some distribution $\mathcal{D}$ over $\mathcal{P}_X$.

Revised goal: $\Pr(x, x'$ are separated in $\Pi)$ is small if $\mathrm{dist}(x, x')$ is small.

Example: $X = \mathbb{R}$.
Partition: $[x_0 + i\Delta, x_0 + (i+1)\Delta]$, where x_0 is random shift.

$$\Pr(x, y \text{ are separated}) \leq \frac{|x - y|}{\Delta}$$

Random partition for any metric space

Set $\Delta = 2^u$.

Let σ be uniform random permutation of X,

$\alpha \in [1/4, 1/2]$ uniform random.

Random partition for any metric space

Set $\Delta = 2^u$.
Let σ be uniform random permutation of X,
$\alpha \in [1/4, 1/2]$ uniform random.

Greedy partiton:
Put all points within distance $R := \alpha\Delta$ of σ_1 into first cluster.
Remove the cluster from σ, repeat.

Random partition for any metric space

Set $\Delta = 2^u$.
Let σ be uniform random permutation of X,
$\alpha \in [1/4, 1/2]$ uniform random.

Greedy partiton:
Put all points within distance $R := \alpha\Delta$ of σ_1 into first cluster.
Remove the cluster from σ, repeat.

Cluster dimater is $2R = 2\alpha\Delta \leq \Delta$ ✓

Clustering quality

Lemma. For any $x \in X$ and $t \leq \Delta/8$,

$$\Pr\left(B(x,t) \not\subseteq \Pi(x)\right) \leq \frac{8t}{\Delta} \ln \frac{M}{m}$$

where $m = \#$ of pts at distance $\leq \Delta/8$
and $M = \#$ of pts at distance $\leq \Delta$

Clustering quality

Lemma. For any $x \in X$ and $t \leq \Delta/8$,

$$\Pr\Big(B(x,t) \not\subseteq \Pi(x)\Big) \leq \frac{8t}{\Delta} \ln \frac{M}{m}$$

where $m = \#$ of pts at distance $\leq \Delta/8$
and $M = \#$ of pts at distance $\leq \Delta$

Proof. Let $U =$ pts w where $B(w, \Delta/2) \cap B(x,t) \neq \emptyset$
$U = (w_1, \ldots, w_{|U|}) :=$ sorted by increasing distance from x.

$\mathcal{E}_k :=$ event that w_k is first in σ s.t. $\Pi(w_k) \cap B(x,t) \neq \emptyset$,
BUT $B(x,t) \not\subseteq \Pi(w_k)$

If $B(x,t) \not\subseteq \Pi(x)$ then some $\mathcal{E}_k$ must occur.

$\mathcal{E}_k$ only if R in some range

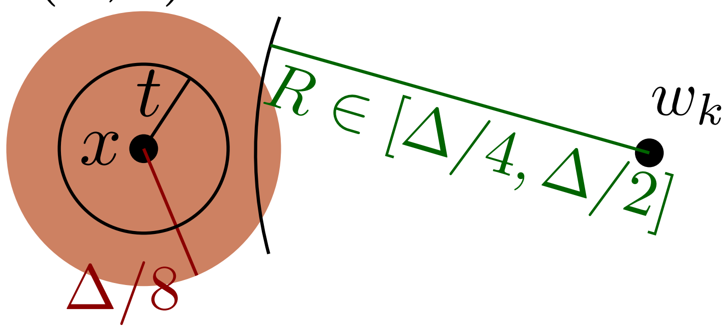
Let $I_k = [\mathrm{dist}(x, w_k) - t, \mathrm{dist}(x, w_k) + t]$.
Claim: $R \notin I_k \Rightarrow \Pr(\mathcal{E}_k) = 0$

$\mathcal{E}_k$ only if R in some range

Let $I_k = [\text{dist}(x, w_k) - t, \text{dist}(x, w_k) + t]$.
Claim: $R \notin I_k \Rightarrow \Pr(\mathcal{E}_k) = 0$

If $d(x, w_k) < R - t$, then $B(w_k, R) \supseteq B(x, t)$, so $\Pr(\mathcal{E}_k) = 0$.
If $d(x, w_k) > R + t$, then $B(w_k, R) \cap B(x, t) = \emptyset$, so $\mathcal{E}_k$ is impossible.

$\Rightarrow \Pr(w_i) = 0$ if $i \leq m$ or $i > M$

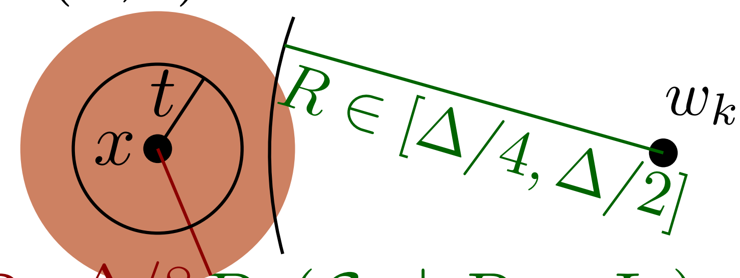


$\mathcal{E}_k$ only if R in some range

Let $I_k = [\mathrm{dist}(x, w_k) - t, \mathrm{dist}(x, w_k) + t]$.
Claim: $R \notin I_k \Rightarrow \Pr(\mathcal{E}_k) = 0$

If $d(x, w_k) < R - t$, then $B(w_k, R) \supseteq B(x, t)$, so $\Pr(\mathcal{E}_k) = 0$.
If $d(x, w_k) > R + t$, then $B(w_k, R) \cap B(x, t) = \emptyset$, so $\mathcal{E}_k$ is impossible.

$\Rightarrow \Pr(w_i) = 0$ if $i \leq m$ or $i > M$



$$\Pr(\mathcal{E}_k) = \Pr\left(\mathcal{E}_k \cap (R \in I_k)\right) = \Pr(R \in I_k) \Pr(\mathcal{E}_k \mid R \in I_k)$$

$$\leq \frac{length(I_k)}{\Delta/2 - \Delta/4} = \frac{2t}{\Delta/4} = \frac{8t}{\Delta}$$

If $w_1, \ldots, w_{k-1}$ are closer to x than w_k, so if one of them (w_i) occurs before w_k in σ, then w_k is not first to scoop from $B(x, t)$ as $\mathrm{dist}(x, w_i) \leq d(x, w_t) \leq R + t$

$\Rightarrow \Pr(\mathcal{E}_k \mid R \in I_k) \leq \frac{1}{k}$

Random partition quality estimate

$$\Pr(B(x,t) \not\subseteq \Pi(x)) = \sum_{k=1}^{|U|} \Pr(\mathcal{E}_k) = \sum_{k=m+1}^{M} \Pr(\mathcal{E}_k)$$

$$= \sum_{k=m+1}^{M} \Pr(R \in I_k) \Pr(\mathcal{E}_k \mid R \in I_k)$$

$$\leq \sum_{k=m+1}^{M} \frac{8t}{\Delta} \frac{1}{k}$$

$$< \frac{8t}{\Delta} \ln \frac{M}{m}$$

Embedding into HSTs

HSTs and quadtrees

Definition. A hierarchically well-separated tree (HST) is a metric space on the leaves of a rooted tree T where each vertex has a label $\Delta \geq 0$ s.t.

- leaves have label $\Delta_v = 0$
- each internal vertex v has $\Delta_v > 0$, and for any child u: $\Delta_u \leq \Delta_v$.
- if x, x' leaves, then $\mathrm{dist}_T(x, x') = \Delta_{lca(x,x')}$

HSTs and quadtrees

Definition. A hierarchically well-separated tree (HST) is a metric space on the leaves of a rooted tree T where each vertex has a label $\Delta \geq 0$ s.t.

- leaves have label $\Delta_v = 0$
- each internal vertex v has $\Delta_v > 0$, and for any child u: $\Delta_u \leq \Delta_v$.
- if x, x' leaves, then $\mathrm{dist}_T(x, x') = \Delta_{lca(x,x')}$

Example: quadtree.
$T =$ quadtree, $\Delta_v =$ diameter of cell v.

$$\|x - x'\|_2 \leq \Delta_{lca(x,x')} = \mathrm{dist}_T(x, x')$$

a bad embedding of $P \subset \mathbb{R}^d$ into a tree metric

HSTs and quadtrees

Definition. A hierarchically well-separated tree (HST) is a metric space on the leaves of a rooted tree T where each vertex has a label $\Delta \geq 0$ s.t.

- leaves have label $\Delta_v = 0$
- each internal vertex v has $\Delta_v > 0$, and for any child u: $\Delta_u \leq \Delta_v$.
- if x, x' leaves, then $\mathrm{dist}_T(x, x') = \Delta_{lca(x,x')}$

Example: quadtree.
$T =$ quadtree, $\Delta_v =$ diameter of cell v.

$$\|x - x'\|_2 \leq \Delta_{lca(x,x')} = \mathrm{dist}_T(x, x')$$

a bad embedding of $P \subset \mathbb{R}^d$ into a tree metric

k-HST: a HST where $\Delta_u \leq \Delta_v / k$

Probabilistic embedding into a 2-HST

Randomized alg. for non-contracting embedding from X into a HST T has probabilistic distortion:

$$\max_{x,y \in X} \frac{\mathbf{E}(\mathrm{dist}_T(x,y))}{\mathrm{dist}_X(x,y)}$$

Theorem. Given (X, dist), there is a randomized embedding into a 2-HST with prob. distortion $\leq 24 \ln n$

Probabilistic embedding into a 2-HST

Randomized alg. for non-contracting embedding from X into a HST T has probabilistic distortion:

$$\max_{x,y \in X} \frac{\mathbf{E}(\mathrm{dist}_T(x,y))}{\mathrm{dist}_X(x,y)}$$

Theorem. Given (X, dist), there is a randomized embedding into a 2-HST with prob. distortion $\leq 24 \ln n$

Proof. Wlog. scale X so $\mathrm{diam}(X) = 1$.
Start with $P = X$, set T's root label to 1.
Compute random partition with $\Delta = \mathrm{diam}(P)/2$, set the diam of partition classes as child labels. Recurse on each child.

Probabilistic embedding into a 2-HST

Randomized alg. for non-contracting embedding from X into a HST T has probabilistic distortion:

$$\max_{x,y \in X} \frac{\mathbf{E}(\mathrm{dist}_T(x,y))}{\mathrm{dist}_X(x,y)}$$

Theorem. Given (X, dist), there is a randomized embedding into a 2-HST with prob. distortion $\leq 24 \ln n$

Proof. Wlog. scale X so $\mathrm{diam}(X) = 1$.
Start with $P = X$, set T's root label to 1.
Compute random partition with $\Delta = \mathrm{diam}(P)/2$, set the diam of partition classes as child labels. Recurse on each child.

level of node v in T: $\lceil \log(\Delta_v) \rceil \leq 0$

Bounding distortion of rand. HST embedding

$x, y \in X$ have lca u in T.

$$\mathrm{dist}_T(x, y) = \Delta_u \leq 2^{level(u)}$$

σ: path from root of T to leaf x.
σ_i: level i node in σ (if exists)
$\mathcal{E}_i$: event that $B_X\big(x, \mathrm{dist}_X(x, y)\big) \not\subseteq \Pi(\sigma_i)$.
Y_i: indicator that $\mathcal{E}_i$ occurs but for all $j > i$ event $\mathcal{E}_j$ does not

Bounding distortion of rand. HST embedding

$x, y \in X$ have lca u in T.

$$\mathrm{dist}_T(x, y) = \Delta_u \leq 2^{level(u)}$$

σ: path from root of T to leaf x.
σ_i: level i node in σ (if exists)
$\mathcal{E}_i$: event that $B_X\big(x, \mathrm{dist}_X(x, y)\big) \not\subseteq \Pi(\sigma_i)$.
Y_i: indicator that $\mathcal{E}_i$ occurs but for all $j > i$ event $\mathcal{E}_j$ does not

We have $d_T(x, y) \leq \sum_i 2^i Y_i$.
Set $j := \lfloor \log \mathrm{dist}_X(x, y) \rfloor$.
If $i < j$, then $\Pr(\mathcal{E}_i) = 0 \Rightarrow \mathbb{E}(Y_i) = 0$.

Bounding distortion of rand. HST embedding

$x, y \in X$ have lca u in T.

$$\mathrm{dist}_T(x,y) = \Delta_u \leq 2^{level(u)}$$

σ: path from root of T to leaf x.
σ_i: level i node in σ (if exists)
$\mathcal{E}_i$: event that $B_X\big(x, \mathrm{dist}_X(x,y)\big) \not\subseteq \Pi(\sigma_i)$.
Y_i: indicator that $\mathcal{E}_i$ occurs but for all $j > i$ event $\mathcal{E}_j$ does not

We have $d_T(x,y) \leq \sum_i 2^i Y_i$.
Set $j := \lfloor \log \mathrm{dist}_X(x,y) \rfloor$.
If $i < j$, then $\Pr(\mathcal{E}_i) = 0 \Rightarrow \mathbb{E}(Y_i) = 0$.

If $i \geq j$, then

$$\mathbb{E}(Y_i) = \Pr(\mathcal{E}_i \cap \overline{\mathcal{E}_{i+1}} \cap \cdots \cap \overline{\mathcal{E}_0}) \leq \frac{8\mathrm{dist}_X(x,y)}{2^i} \ln \frac{|B_X(x, 2^i)|}{|B_X(x, 2^i/8)|}$$

Distortion bound wrap-up

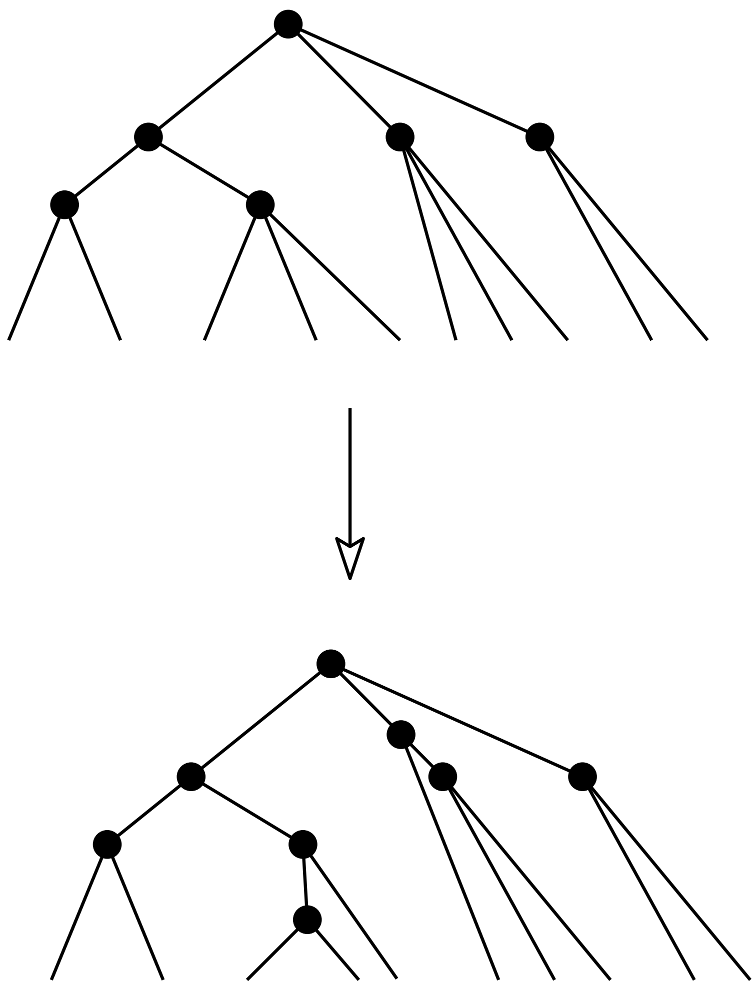
Set $n_i = B_X(x, 2^i)$, and $t := \mathrm{dist}_X(x, y)$.

$$\begin{aligned}
\mathbb{E}(d_T(x,y)) &\leq \mathbb{E}\Big(\sum_i 2^i Y_i\Big) = \sum_i 2^i \mathbb{E}(Y_i) \\
&\leq \sum_{i=j}^{0} 2^i \frac{8t}{2^i} \ln \frac{n_i}{n_{i-3}} = 8t \ln \left(\prod_{i=j}^{0} \frac{n_i}{n_{i-3}} \right) \\
&\leq 8t \ln(n_0 n_1 n_2) \leq 24t \ln n.
\end{aligned}$$

□

k-median in HST

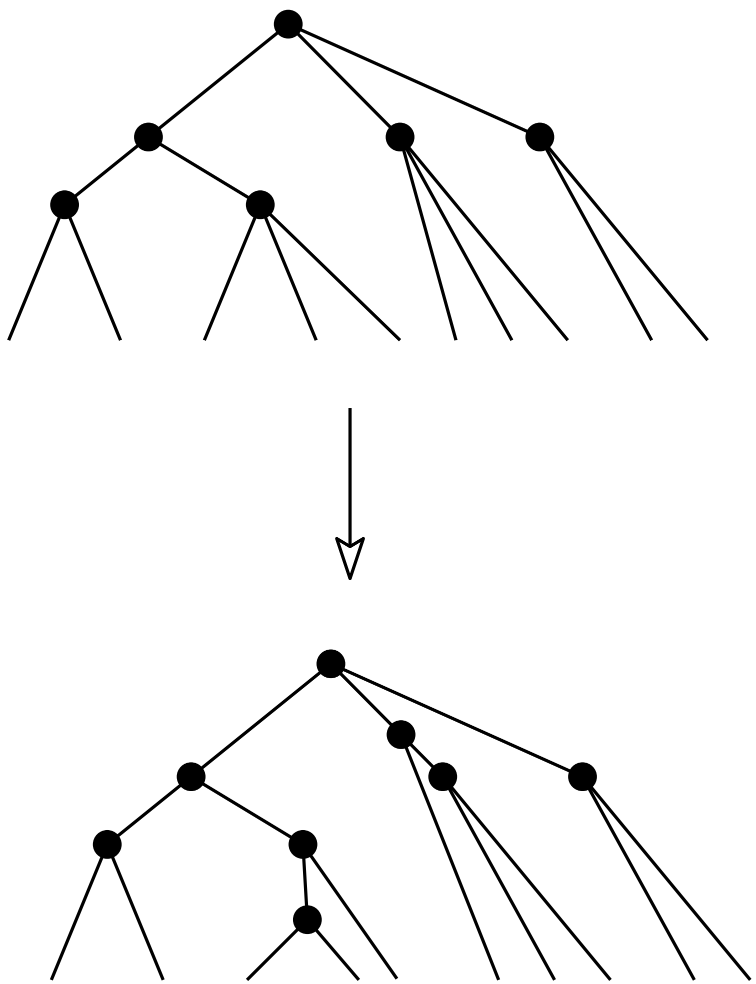
Computing k-median in HST is “easy”



- make it into binary HST (new nodes have same label)

k-median in HST

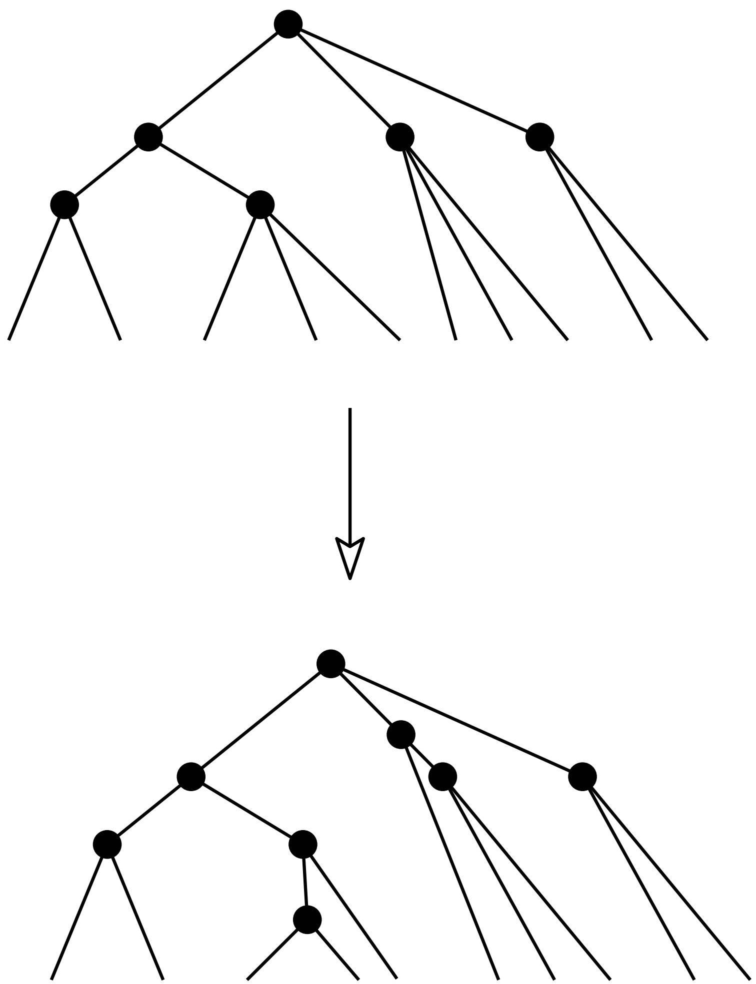
Computing k-median in HST is “easy”



- make it into binary HST (new nodes have same label)
- Dyanimc program. Subproblem at v, param $\ell \in [k]$: what is cheapest ℓ-median for descendants of v?

k-median in HST

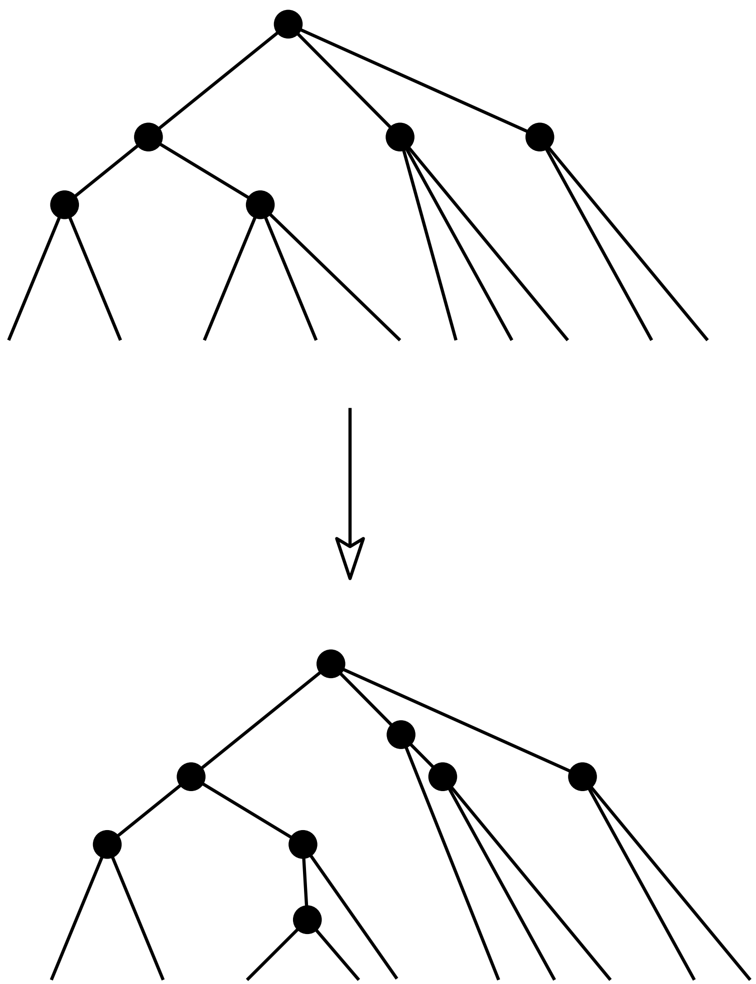
Computing k-median in HST is "easy"



- make it into binary HST (new nodes have same label)
- Dyanimc program. Subproblem at v, param $\ell \in [k]$: what is cheapest ℓ-median for descendants of v?
- Recursive step: for each a, b with $a + b = \ell$, compute a-median in left child subtree and b-median in right child subtree.

k-median in HST

Computing k-median in HST is "easy"



- make it into binary HST (new nodes have same label)
- Dyanimc program. Subproblem at v, param $\ell \in [k]$: what is cheapest ℓ-median for descendants of v?
- Recursive step: for each a, b with $a + b = \ell$, compute a-median in left child subtree and b-median in right child subtree.

$O(k^2 n)$

Application: k-median approximation in metric spaces

Theorem. There is an $O(\log n)$-approximation for k-median in any metric space (X, dist_X).

Application: k-median approximation in metric spaces

Theorem. There is an $O(\log n)$-approximation for k-median in any metric space (X, dist_X).

Proof. Emebed $P \subseteq X$ into a HST T.
Compute cluster centers C in T.
C induces clustering $\mathcal{X}$ in P (center of p is $nn_X(p, C)$).
Return $C, \mathcal{X}$. OPT: $(C_{opt}, \mathcal{X}_{opt})$

$$\gamma(C, \mathrm{dist}_X) \le \gamma(C, \mathrm{dist}_T) \le \gamma(C_{opt}, \mathrm{dist}_T)$$
$$= \sum_{p \in P} \mathrm{dist}_T(p, C_{opt}) \le \sum_{p \in P} \mathrm{dist}_T(p, nn_X(p, C_{opt}))$$

Application: k-median approximation in metric spaces

Theorem. There is an $O(\log n)$-approximation for k-median in any metric space (X, dist_X).

Proof. Emebed $P \subseteq X$ into a HST T.
Compute cluster centers C in T.
C induces clustering $\mathcal{X}$ in P (center of p is $nn_X(p, C)$).
Return $C, \mathcal{X}$. OPT: $(C_{opt}, \mathcal{X}_{opt})$

$$\gamma(C, \mathrm{dist}_X) \leq \gamma(C, \mathrm{dist}_T) \leq \gamma(C_{opt}, \mathrm{dist}_T)$$

$$= \sum_{p \in P} \mathrm{dist}_T(p, C_{opt}) \leq \sum_{p \in P} \mathrm{dist}_T(p, nn_X(p, C_{opt}))$$

$$\mathbb{E}(\gamma(C, \mathrm{dist}_X)) = \sum_{p \in P} \mathbb{E}\big(\mathrm{dist}_T(p, nn_X(p, C_{opt}))\big)$$

$$= \sum_{p \in P} O\big(\mathrm{dist}_X(p, nn_X(p, C_{opt}))\log n\big)$$

$$= O\big(\gamma(\mathcal{X}_{opt}, C_{opt}, \mathrm{dist}_X) \cdot \log n\big) \quad \square$$

Further embeddings into ℓ_2

Embedding into ℓ_2

Theorem (Bourgain 1985). Any n-pt metric space can be embedded into $\mathbb{R}^{O(\log n)}$ (with ℓ_2 metric) with distortion $O(\log n)$.

This is tight for coonstant-degree expanders.

Embedding into ℓ_2

Theorem (Bourgain 1985). Any n-pt metric space can be embedded into $\mathbb{R}^{O(\log n)}$ (with ℓ_2 metric) with distortion $O(\log n)$.

This is tight for coonstant-degree expanders.

Some proof ideas for weaker version:

- forget dimension (use JL in the end)
- for a given resolution r, use $O(\log n)$ random HST embedding of diameter r.
  Flip coin for each cluster; if heads, create an anchor set Y_j.
- embedding: j-th coord of x wrt. anchors Y_j is $\mathrm{dist}(x, Y)$.
  This is non-contracting.
  For each resolution we get $O(\log n)$ coords each

Proof ideas for weak Bourgain, ctd.

- Let x, y arbitrary, and r a resolution where $r/2 < \mathrm{dist}(x,y)/2 < r$. $\Rightarrow x$ and y are in different clusters, and with prob. $1/2$ the ball $B(x, O(1/\log n))$ is contained in the cluster of x
- Chernoff $\Rightarrow$ w.h.p. a constant proportion of the coordiantes j will differ by $\Omega(r/\log n)$ (when x, y get different coin flips)
- if they differ on k flips, then these cords contribute distance at least $\Omega(\sqrt{k}/\log n)$.

- *spread* Φ: ratio of largest/smallest distance in X. By 'snapping' distances less than r/n or much more than r, we get new metrics on X with spread $\Phi = O(n^2)$, and there are $O(n^2)$ distinct metrics, get coords from each.

Embedding special metrics into ℓ_2

Tree metric: induced by possitively edge-weighted tree.

Theorem (Matoušek 1999). Any tree metric can be embedded into ℓ_2 with distortion $O(\sqrt{\log \log n})$.

Distortion bound is tight (up to constant factors.)

Embedding special metrics into ℓ_2

Tree metric: induced by possitively edge-weighted tree.

Theorem (Matoušek 1999). Any tree metric can be embedded into ℓ_2 with distortion $O(\sqrt{\log \log n})$.

Distortion bound is tight (up to constant factors.)

Theorem (Rao 1999). Let $\mathcal{G}$ be graph class that excludes some forbidden minor H (e.g. planar graphs.). Then any $\mathcal{G}$-metric can be mebedded into ℓ_2 with distortion $O(\sqrt{\log n})$.

Distortion bound is tight (up to constant factors.)