Last time:

$\forall$ n-point (X, d_X) $\forall \varepsilon > 0$ $\exists$ a data structure of size $O(n^{1+\varepsilon})$ s.t. one can estimate $d_X(x_1, x_2)$ up to $O(1/\varepsilon)$ within time $O(1)$.

Today: better approximation, slower query time.

Th $\forall k \geq 1$ can get space $O(kn^{1+1/k})$, approximation $2k-1$, query time $O(k)$

$3 \quad n^{3/2}$
$5 \quad n^{4/3}$
$\dots$

Let's forget about query time and assume that X is induced by an unweighted undirected graph $G = (V, E)$.

Girth: g, the length of the shortest cycle.

Cl 1 Any graph of girth $\geq 2k$ has $O(n^{1+1/k})$ edges

Cl 2 $\exists$ graphs of girth $\geq 2k$ with $n^{1+\Omega(1/k)}$ edges

Conj (known for small k). $\Omega(n^{1+1/k})$

Pf (Cl 1) remove all nodes of degree $\leq n^{1/k}$

The remaining graph if non-empty has girth $\geq 2k$.

Pf (Cl 2)

Probabilistic method.

$G(n,p)$ $\quad p = n^{\lambda - 1}$ $\quad \lambda < 1.$

$$E[\#\text{short cycles}] \le \sum_{j=3}^{2k} n^j p^j \le O(n^{2k\lambda}).$$

if $2k\lambda < 1 \Rightarrow \ll \frac{n}{100}$.

By Markov at most $\frac{n}{50}$ cycles whp.

All degrees are $\Theta(n^\lambda)$ whp.

Remove a vertex from each short cycle.

$O(n^{1+\lambda})$ edges $\quad \lambda < \frac{1}{2k}$

$O(n^{1+\frac{1}{2k}})$ edges

Greedy spanner:

add edges that don't create cycles of length $2k$.

$\le O(n^{1+1/k})$ edges

$(2k-1)$-approx.

Any approx. $\le 2k-1-\varepsilon$ requires $\Omega(n^{1+1/k})$ space

What about fast query time?

General metrics

$k=2$ (3-approx, $n^{3/2}$-ish space).

Sample $O(\sqrt{n})$ landmarks:

$$\Pr[x - \text{landmark}] = \frac{1}{\sqrt{n}}$$

For points:

store distances to all landmarks

store distances to all points closer than the closest landmark. $B(v)$.

store closest landmark

Query:

if $u \in B(v) \Rightarrow$ know the answer

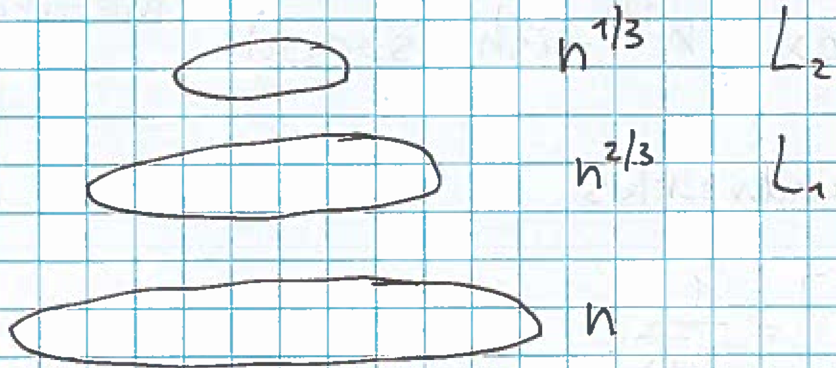
$d(u,w) + d(w,v)$, where w - landmark closest to v.

$$d(u,v) \le d(u,w) + d(w,v) \le 3d(u,v).$$

$d(u,w) \le d(u,v) + d(v,w) \le 2d(u,v)$, $d(w,v) \le d(u,v)$.

Space: $n^{3/2}$.

$k=3$ 5-approx $n^{4/3}$



$$B(x) = \{\tilde{x} \in X \mid d(x,\tilde{x}) < d(x, L_1)\} \cup$$
$$\cup \{\tilde{x} \in \cancel{X}_1 \mid d(x,\tilde{x}) < d(x, L_2)\} \cup$$
$$\cup L_2.$$

<u>Query:</u>

if $u \in B(v) \Rightarrow d(u,v)$

if $L_1(v) \in B(u) \Rightarrow d(L_1(v), v) + d(u, L_1(v))$.

$$d(u, L_2(u)) + d(L_2(u), v).$$

$$d(u, L_2(u)) + d(L_2(u), v) \cancel{\leq}$$

$\cancel{\leq d(u, L_1(v)) + d(L_2(u)}$ $\qquad$ $\geq d(u, L_2(u)) +$

$\leq d(u, L_1(v)) \leq$ $\qquad$ $+ d(u,v) \leq$

$\leq d(u,v) + d(v, L_1(v)) \leq$ $\qquad$ $\leq 3d(u,v)$.

$\leq 2d(u,v)$