Today

- Bourgain intuition
- Expansion claim
- Proof of ↑
- Proof of Bourgain w/

An embedding of metric space (V,d) into metric space (V',d') is a function $f: V \to V'$

f has distortion α if $d(u,v) \le d'(f(u), f(v)) \le \alpha \cdot d(u,v)$ $\forall u,v \in V$

Given metric (V, δ)

The (closed) radius r ball centered at $x \in V$ is $B(x,r) := \{y \in V : \delta(x,y) \le r\}$

The open radius r ball centered at $x \in V$ is $B^o(x,r) := \{y \in V : \delta(x,y) < r\}$

may not need

Strong Cauchy-Schwarz: $\sum_i |u_i v_i| \le \sqrt{\sum_i u_i} \sqrt{\sum_i v_i} = \|u\| \cdot \|v\|$ $\forall u,v \in \mathbb{R}^n$

The Concentration Framework

a) Show fact true if all RVs near $\mathbb{E}$

b) Concentration: each RV at $\mathbb{E}$ $(\pm \log n)$ w/ good probability

c) Union bound: all " "

Don't really need

Chernoff Bound

Let $X_1, X_2, \ldots, X_n$ be independent RVs

s.t. $X_i = \begin{cases} 1 & \text{w/ Pr } p_i \\ 0 & \text{o/w} \end{cases}$

Let $X := \sum_i X_i$, $\mu := \mathbb{E}[X]$

Then $\forall$ $\delta \in (0,1)$

$$\Pr(X \le (1-\delta) \cdot \mu) \le \exp(-\delta^2 \mu / 2)$$

Probabilistic Method Approach

1) Randomly make choices

2) Show Pr(desired outcome) > 0

(so desired outcome is possible)

Last time: (V,d) for $V \subseteq \mathbb{R}^n$ embeds into (V',d) for $V' \subseteq \mathbb{R}^k$ and $k << n$

But not all metrics are from subsets of $\mathbb{R}^n$ w/ Euc. distance, e.g. shortest paths on graphs; see hw.

This time: *any* metric embeds into (V,d) for $V \subseteq \mathbb{R}^n$ w/ low distortion

Bourgain's Theorem: given *any* n-point metric (V,δ), $\exists$ (poly-time computable) embedding f w/ distortion $O(\log n)$ of (V,δ) into $(\tilde{V}, d)$ for $\tilde{V} \subseteq \mathbb{R}^{O(\log^2 n)}$

Euc. distance

Can reduce to $O(\log n)$ w/ JL

Construction of f:

Intuition: estimate distances by difference in distances to "waypoint" sets

Distances to sets

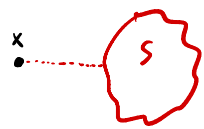
For $S \subseteq V$, $x \in V$, let $S(x) := \underset{y \in S}{\text{argmin}}\, \delta(x,y)$ and $\delta(x,S) := \delta(x, S(x))$

Claim: $\delta(x,S) \leq \delta(x,y) + \delta(y,S)$ $\forall x,y \in V$, $S \subseteq V$ (set triangle inequality)

Have $\delta(x,S) = \delta(x,S(x)) \leq \delta(x,S(y)) \underset{\text{triangle inequality}}{\leq} \delta(x,y) + \delta(y,S(y)) = \delta(x,y) + \delta(y,S)$

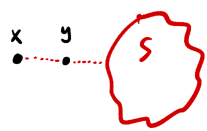
Will let $\tilde{x} := f(x)$

First Attempt: Fix some $S \subseteq V$, $x \overset{f}{\rightsquigarrow} (\delta(x,S))$

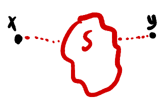


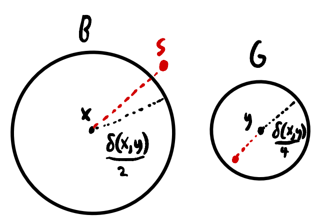
$\overset{f}{\rightsquigarrow} (\delta(x,S))$

Distances never go up: $d(\tilde{x},\tilde{y}) = |\delta(x,S) - \delta(y,S)| \underset{\text{Set Triangle inequality}}{\leq} \delta(x,y)$ $\forall x,y \in V$



Distances can go down: e.g. when $\delta(x,S) = \delta(y,S)$ $\forall x,y \in V$





When distances don't go down much:

Let $B := B\left(x, \frac{\delta(x,y)}{2}\right)$, $G := \left(y, \frac{\delta(x,y)}{4}\right)$

trivial if only care about x,y

"the dream" If $S \in G$ and $S \notin B$ then $d(\tilde{x},\tilde{y}) = |\delta(x,S) - \delta(y,S)| \geq \frac{\delta(x,y)}{4}$

Suppose $|B| \approx |G|$ and say x',y' is <u>like</u> x,y if

$$\left|B\left(x', \tfrac{\delta(x',y')}{2}\right)\right| \approx \left|B\left(y', \tfrac{\delta(x',y')}{4}\right)\right| \approx |B|$$

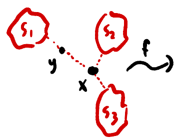
If $S \in V$ w/ Pr $\frac{1}{|B|}$ then $\forall x',y'$ like x,y dream happens w/ $\Omega(1)$ prob.

Problem: need distances to always not increase, not just w/ $\Omega(1)$ probality

Solution: $\Theta(\log n)$ repetitions

Second Attempt: Let $S_i \leftarrow V$ w/ Pr $\frac{1}{|B|}$ independently for $i \in [\Theta(\log n)]$

Let $\tilde{X} = (\delta(x, S_1), \delta(x, 2), \ldots)$



$\tilde{x} = (\delta(x, S_1), \delta(x, S_2), \delta(x, S_3))$

$\tilde{y} = (\delta(y, S_1), \delta(y, S_2), \delta(y, S_3))$

WhP any pair like x, y achieves dream for $\Theta(\log n)$ S_i

↳ distances don't decrease for these pairs

Problem: need f to work for all pairs not just those like x, y (balls of size $\approx |B|$)

Solution: try all possible values of $|B|$ as powers of 2

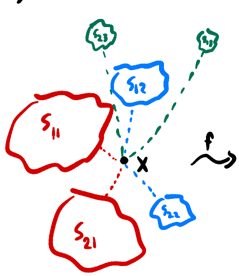
Final Attempt: Let $c \geq 1$ be a large constant chosen later

For $j \in [\log n]$ ($|B|$ possibilities)

For $i \in [c \cdot \log n]$ (repetitions)

S_{ij} contains each $v \in V$ independently w/ Pr $\frac{1}{2^j}$

$\tilde{X}_{ij} = \delta(x, S_{ij})$ so $\tilde{x} \in \mathbb{R}^{c \cdot \log^2 n}$



$\overset{f}{\rightsquigarrow}$ $\tilde{x} = (\delta(x, S_{11}), \delta(x, S_{12}), \delta(x, S_{13}), \ldots)$

Expansion Claim: $\sum_{i,j} |\tilde{x}_{ij} - \tilde{y}_{ij}| \geq \frac{c}{40} \cdot \log n \cdot \delta(x,y)$ $\quad \forall x,y$ except w/ $Pr \leq \frac{1}{n}$

Skip+ come back

Proof of Bourgain Using Expansion Claim

WTS $\delta(x,y) \leq d(\tilde{x},\tilde{y}) \leq O(\log^2 n) \cdot \delta(x,y)$

Set triangle inequality

Observe $\delta(x,S_{ij}) - \delta(y,S_{ij}) \leq \delta(x,y)$ $\forall i,j$ b/c $\delta(x,S_{ij}) \leq \delta(x,y) + \delta(y,S_{ij})$

So $d(\tilde{x},\tilde{y}) = \sqrt{\sum_{i,j}(\tilde{x}_{ij} - \tilde{y}_{ij})^2} = \sqrt{\sum_{i,j}(\delta(x,S_{ij}) - \delta(y,S_{ij}))^2} \leq \sqrt{\sum_{i,j}(\delta(x,y))^2} = c \cdot \log^2 n \cdot \delta(x,y)$

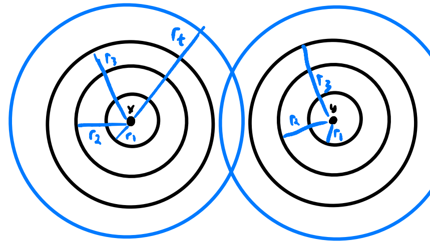
$\in \mathbb{R}^{c \cdot \log^2 n}$

OTOH $d(\tilde{x},\tilde{y}) = \frac{\|\tilde{x}-\tilde{y}\| \cdot \|\mathbb{1}\|}{\|\mathbb{1}\|} \geq \frac{\sum_{i,j} |\tilde{x}_{ij} - \tilde{y}_{ij}|}{\sqrt{c} \cdot \log n} \geq \frac{\sqrt{c}}{40} \cdot \delta(x,y) \geq \delta(x,y)$

Strong Cauchy-Schwarz + $\|\mathbb{1}\| = \sqrt{c} \cdot \log n$ $\quad$ Expansion Claim $\quad$ c large enough

Bourgain follows by Probabalistic method

Proof of Expansion Claim



Fix $x, y \in V$

Let $r_j := \min r$ s.t. $|B(x,r)|, |B(y,r)| \geq 2^j$

Let $t := \min t$ s.t. $2r_t \geq \delta(x,y)$,

For simplicity will <u>assume</u> $2r_t = \delta(x,y)$

So $\boxed{\frac{\delta(x,y)}{2} \leq r_t}$ (1)

Also for $j \leq t$ since $r_{j-1} \leq r_j$ have $\boxed{r_j + r_{j-1} \leq 2r_j \leq \delta(x,y)}$ (2)

Also, for every $j \leq t$ $\boxed{|B^o(x, r_j)| \text{ or } |B^o(y, r_j)| < 2^j}$ (3)

for every $j < t$ $\boxed{|B(x, r_j)| \text{ and } |B(y, r_j)| \geq 2^j}$ (4)

To dispense w/ assumption:

Redefine $r_t := \frac{\delta(x,y)}{2}$ let $\tilde{r}_t$ be old r_t

(a), (d) trivial

Still have $r_{t-1} \leq r_t$ since o/w $2r_{t-1} \geq 2r_t = \delta(x,y)$ contradicting t choice

So $r_t + r_{t-1} \leq 2r_t = \delta(x,y) \to$ gives (6)

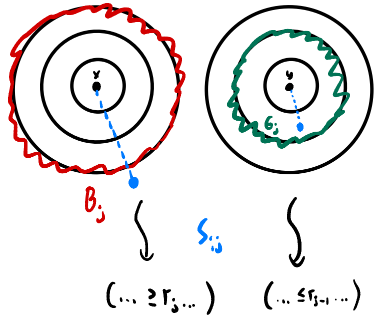
Also $r_t \leq \hat{r}_t$ by definition $\to$ gives (c)

For fixed $j \le t$, have $\sum_i^t |\tilde{X}_{ij} - \tilde{y}_{ij}| \ge \frac{c}{20} \cdot \log n \, (r_j - r_{j-1})$ w/ $\Pr \ge 1 - \frac{1}{n^4}$

WLOG suppose $|B^0(x, r_j)| < 2^j$ (3)

Let $B_j = B^0(x, r_j)$ → Want empty

Let G_j be $B(y, r_{j-1})$ → want non-empty



Let $E_i := G_j \cap S_{ij} \ne \emptyset$ and $B_j \cap S_{ij} = \emptyset$

If E_i then $|\tilde{X}_{ij} - \tilde{y}_{ij}| \ge r_j - r_{j-1}$

If $B_j \cap S_{ij} = \emptyset$ then $\tilde{X}_{ij} = \delta(x, S_{ij}) \ge r_j$

If $G_j \cap S_{ij} \ne \emptyset$ then $\tilde{y}_{ij} = \delta(y, S_{ij}) \le r_{j-1}$

so if $B_j \cap S_{ij} = \emptyset$ and $G_j \cap S_{ij} \ne \emptyset$ then $|\tilde{X}_{ij} - \tilde{y}_{ij}| \ge r_j - r_{j-1}$

$\Pr(E_i) \ge \frac{1}{10}$

Notice $|B_j| < 2^j$ and $|G_j| \ge 2^{j-1}$ (4)

$$\Pr(G_j \cap S_{ij} \ne \emptyset) \ge 1 - \left(1 - \tfrac{1}{2^j}\right)^{2^{j-1}} \underset{\substack{\uparrow \\ 1-x \le e^{-x} \ \forall x}}{\ge} 1 - \exp(-2^{j-1}/2^j) = 1 - \frac{1}{\sqrt{e}}$$

$$\Pr(B_j \cap S_{ij} = \emptyset) \ge \left(1 - \tfrac{1}{2^j}\right)^{2^j} = \left(1 - .63 \tfrac{1}{.63 \cdot 2^j}\right)^{2^j} \underset{\substack{1 - .63x \ge e^{-x} \\ \downarrow}}{\ge} \exp(-.63) = \frac{1}{e^{.63}}$$

But $j \le t$ so $B_j \cap G_j = \emptyset$ (2)

so $B_j \cap S_{ij} = \emptyset$ and $G_j \cap S_{ij} \ne \emptyset$ are independent and $\left(1 - \frac{1}{\sqrt{e}}\right)\left(\frac{1}{e^{.63}}\right) \ge \frac{1}{10}$

Let $X_i := \mathbb{1}(E_i)$ and $X := \sum_i^t X_i$ so $\mathbb{E}[X] \ge \frac{c}{10} \cdot \log n$

$\Pr(X \le \frac{c}{20} \cdot \log n) \le \frac{1}{n^4}$ by Chernoff for c large constant → $\sum_i^t |\tilde{X}_{ij} - \tilde{y}_{ij}| \ge \frac{c}{20} \cdot \log n \, (r_j - r_{j-1})$ w/ $\Pr \ge 1 - \frac{1}{n^4}$

By union bound over $c \cdot \log n$ j: $\sum_i^t |\tilde{X}_{ij} - \tilde{y}_{ij}| \ge \frac{c}{20} \cdot \log n \, (r_j - r_{j-1})$ w/ $\Pr \ge 1 - \frac{1}{n^3}$ $\forall j \le t$

Thus
$$\sum_{i,j} |\tilde{X}_{ij} - \tilde{y}_{ij}| \ge \sum_{i, j \le t} |\tilde{X}_{ij} - \tilde{y}_{ij}| \ge \frac{c}{20} \cdot \log n \sum_{j \le t} (r_j - r_{j-1}) \underset{\substack{\uparrow \\ \text{telescoping} \\ + \\ r_0 = 0}}{=} \frac{c}{20} \cdot \log n \cdot r_t \underset{\substack{\uparrow \\ r_t \ge \frac{\delta(x,y)}{2} \ (1)}}{\ge} \frac{c}{20} \cdot \log n \cdot \frac{\delta(x,y)}{2}$$

except w/ $\Pr \le \frac{1}{n^3}$

By a union bound over $\le n^2$ x, y have the expansion claim