

## Today

- $\ell_1$  Metrics + Bourgain for  $\ell_1$
- $\ell_1$  vs  $\ell_2$
- Metrics as vectors + convex cones
- $\ell_1(V) = \text{Cone}(\text{cut}(V))$

## Recall don't need

A metric space consists of a set of "points"  $V$  and function  $f: V \times V \rightarrow \mathbb{R}_{\geq 0}$  satisfying

- "metric" for short
- 1)  $f(x, x) = 0 \quad \forall x \in V \rightarrow$  recall  $f(x, y) = 0$  iff  $x = y$  (or a "pseudo metric")
  - 2)  $f(x, y) = f(y, x) \quad \forall x, y \in V$  (Symmetric)
  - 3)  $f(x, y) \leq f(x, z) + f(z, y) \quad \forall x, y, z \in V$  (triangle inequality)

An embedding of metric space  $(V, d)$  into metric space  $(V', d')$  is a function  $f: V \rightarrow V'$ .  
 $(V, d)$  embeds isometrically into  $(V', d')$  iff  $\exists$  embedding  $f$  s.t.  $d(u, v) = d'(f(u), f(v)) \quad \forall u, v \in V$

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

don't need  
 For  $j \in [\log n]$  (|B| possibilities)

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

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

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

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

don't need  
Exc. distance can reduce to  $O(\log n)$  w/ JL

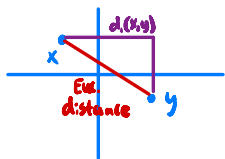
Expansion Claim:  $\sum_{i,j} |\tilde{X}_{ij} - \tilde{Y}_{ij}| \geq \frac{c}{40} \cdot \log n \cdot \delta(x, y) \quad \forall x, y$  except w/  $\text{Pr} \leq \frac{1}{n}$

So far, have worked w/ "Straight line" distances in  $\mathbb{R}^n$

Today:  $l_1$  distances in  $\mathbb{R}^n$  and their connection to cuts

$l_1$  Metrics

Given  $x, y \in \mathbb{R}^n$  define  $d_1(x, y) := \sum_i |x_i - y_i|$



Claim:  $(V, d_1)$  is a metric space  $\forall V \subseteq \mathbb{R}^n$

1)  $d_1(x, x) = \sum_i |x_i - x_i| = 0$

2)  $d_1(x, y) = \sum_i |x_i - y_i| = \sum_i |y_i - x_i| = d_1(y, x)$

3)  $d_1(x, y) = \sum_i |x_i - y_i| = \sum_i |x_i - z_i + z_i - y_i| \leq \sum_i |x_i - z_i| + |z_i - y_i| = d_1(x, z) + d_1(z, y)$

$|a+b| \leq |a|+|b| \forall a, b$

Skippable

Bourgain's Theorem: given any  $n$ -point metric  $(V, \delta)$ ,  $\exists$  (poly-time computable) embedding  $f$  w/ distortion  $O(\log n)$  of  $(V, \delta)$  into  $(\tilde{V}, d_1)$  for  $\tilde{V} \subseteq \mathbb{R}^{O(\log^2 n)}$

Let  $\hat{x} := \frac{\tilde{x}}{\log n}$  where  $\tilde{x}$  from  $l_2$  embedding

WTS  $\delta(x, y) \leq d_1(\hat{x}, \hat{y}) \leq O(\log n) \cdot \delta(x, y)$

So  $d_1(\hat{x}, \hat{y}) = \sum_{i,j} |\hat{x}_{i,j} - \hat{y}_{i,j}| = \frac{1}{\log n} \sum_{i,j} |(\delta(x, s_{i,j}) - \delta(y, s_{i,j}))| \leq \frac{1}{\log n} \sum_{i,j} \delta(x, y) = c \cdot \log n \cdot \delta(x, y)$

OTOH  $d_1(\hat{x}, \hat{y}) = \frac{1}{\log n} \cdot d_1(\tilde{x}, \tilde{y}) \geq \frac{1}{\log n} \cdot \frac{c}{40} \cdot \log n \cdot \delta(x, y) \geq \delta(x, y)$

expansion claim

c large

skipped, die  $l_2$  instead

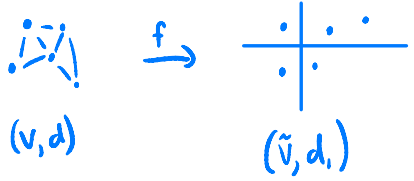
Set triangle inequality

# $l_1$ vs $l_2$

$(V, \delta)$  is an  $l_1$  metric if it embeds isometrically into  $(\tilde{V}, d_1)$  for  $\tilde{V} \subseteq \mathbb{R}^n$  for some  $n$

Euclidean distance

$$(\tilde{V}, d_1)$$



$$l_2(V) \subseteq \mathbb{R}^{\binom{n}{2}}$$

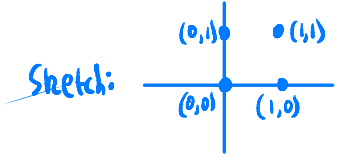
For arbitrary  $V$ , w/  $|V|=n$ , let  $l_1(V) \subseteq \mathbb{R}^{\binom{n}{2}}$  be all  $l_1$  metrics on  $V$

↓ below suspicious

Fact:  $\forall V$  have  $l_2(V) \subseteq l_1(V)$

sketch: random projections

Fact:  $\exists V$  s.t.  $l_1(V) \not\subseteq l_2(V)$   $\rightarrow$   $l_1$  has more "representational power" than  $l_2$



So what can we represent? Short answer: cuts

## Metrics as Vectors

Claim: If  $(V, d)$  is a metric then  $(V, \alpha \cdot d)$  is a metric  $\forall \alpha \geq 0$

- 1)  $d(x, x) = 0 \quad \forall x \in V \rightarrow \alpha \cdot d(x, x) = 0 \quad \forall x \in V$
  - 2)  $d(x, y) = d(y, x) \quad \forall x, y \rightarrow \alpha \cdot d(x, y) = \alpha \cdot d(y, x) \quad \forall x, y$
  - 3)  $\alpha \cdot d(x, y) \leq \alpha \cdot (d(x, z) + d(z, y)) = \alpha \cdot d(x, z) + \alpha \cdot d(z, y) \quad \forall x, y, z$
- $\hat{d}$  triangle inequality
- } Skippable

Claim: If  $(V, d_1)$  and  $(V, d_2)$  are metrics then so is  $(V, d_1 + d_2)$

- 1)  $d_1(x, x) + d_2(x, x) = 0 \quad \forall x$
  - 2)  $d_1(x, y) + d_2(x, y) = d_1(y, x) + d_2(y, x) \quad \forall x, y$
  - 3)  $d_1(x, y) + d_2(x, y) \leq d_1(x, z) + d_1(z, y) + d_2(x, z) + d_2(z, y) \quad \forall x, y, z$
- $d_1, d_2$  triangle inequality
- } Skippable

Life Lesson: if can add + scale, good to interpret as vectors

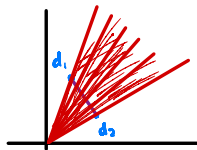
## 1-1 Correspondance Between Metrics and Vectors

$$(V, d) \quad \text{s.t.} \quad V = \{v_1, v_2, \dots, v_n\} \quad \leftrightarrow \quad \begin{pmatrix} d(v_1, v_2) \\ d(v_1, v_3) \\ \vdots \\ d(v_1, v_n) \\ d(v_2, v_3) \\ \vdots \end{pmatrix} \in \mathbb{R}^{\binom{n}{2}} \quad \text{E.g.} \quad \begin{matrix} v_1 \\ \swarrow \quad \searrow \\ v_2 \quad \quad v_3 \\ \text{---} \quad \text{---} \\ (V, d_G) \end{matrix} \quad \leftrightarrow \quad \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}$$

Notation abuse: use  $d \in \mathbb{R}^{\binom{n}{2}}$  for vector of  $(V, d)$  and call  $d$  a metric on  $V$

## Convex Cones

The convex cone of  $D \subseteq \mathbb{R}^{\binom{n}{2}}$  is  $\text{Cone}(D) := \left\{ \sum_{d \in D} \alpha_d \cdot d : \alpha_d \geq 0 \quad \forall d \right\}$



$$D = \{d_1, d_2\}$$

$$[d_1, d_2] \in \text{Cone}(D)$$

$$x \in \text{Cone}(D) \rightarrow \alpha x \in \text{Cone}(D) \quad \forall \alpha \geq 0$$

Corollary: If  $D$  is a set of metrics on  $V$ , then  $\tilde{d} \in \text{Cone}(D)$  is a metric on  $V$

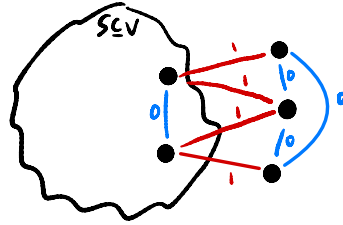
Observe: If  $d_1, d_2 \in \text{Cone}(D)$  then  $d_1 + d_2 \in \text{Cone}(D)$  ← needed?

$$d_1 + d_2 = \sum_{d \in D} \alpha_d \cdot d + \sum_{d \in D} \beta_d \cdot d = \sum_{d \in D} \underbrace{(\alpha_d + \beta_d)}_{\geq 0} \cdot d$$

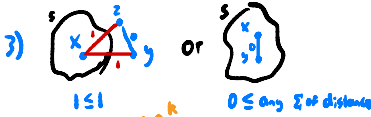
# The Cut Cone

$(V, d_S)$  is a cut metric for  $S \subseteq V$  s.t.

$$d_S(u, v) = \begin{cases} 0 & u, v \in S \text{ or } u, v \notin S \\ 1 & \text{otherwise} \end{cases}$$



A metric b/c (1), (2) trivial



For arbitrary  $V$ , w/  $|V|=n$  let  $\text{Cut}(V) \subseteq \mathbb{R}^{\binom{n}{2}}$  be all cut metrics of  $V$

Theorem:  $\mathcal{L}_1(V) = \text{Cone}(\text{Cut}(V)) \forall V$ .

Also, given  $d \in \mathcal{L}_1(V)$ , <sup>+L1 embedding</sup> can poly-time compute  $d_i \in \text{Cut}(V), \alpha_i \in \mathbb{R} \forall i$  s.t.  $d = \sum_i \alpha_i d_i$

Start w/ simplest case of line metrics

$(V, d)$  is a line metric if it embeds isometrically into  $(\tilde{V}, \tilde{d})$  for  $\tilde{V} \subseteq \mathbb{R}$  and  $\tilde{d}(x, y) = |x - y| \forall x, y \in \tilde{V}$

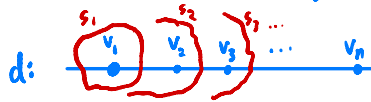


Claim: If  $d$  is a line metric on  $V$  then  $d \in \text{Cone}(\text{cut}(V))$

Also, can poly-time compute  $d_i \in \text{Cut}(V), \alpha_i \in \mathbb{R} \forall i$  s.t.  $d = \sum_i \alpha_i d_i$

Notation abuse: use  $v \in V$  and position on line interchangeably

Let  $V = \{v_1, v_2, \dots, v_n\}$  be  $V$  sorted left  $\rightarrow$  right



Let  $S_i := \{v_1, v_2, \dots, v_i\}$  and let  $d_{S_i}$  be corresponding cut metric for  $i < n$

Let  $\alpha_i := v_{i+1} - v_i \geq 0$  by definition

Have  $d = \sum_i \alpha_i d_{S_i}$  b/c  $\forall y \geq x \sum_i \alpha_i d_{S_i}(v_x, v_y) = \sum_i (v_{i+1} - v_i) = \sum_{i \in [x, y)} (v_{i+1} - v_i) = v_y - v_x = d(v_x, v_y)$

$\because S_i$  separates  $v_x, v_y$   $\uparrow$  telescoping

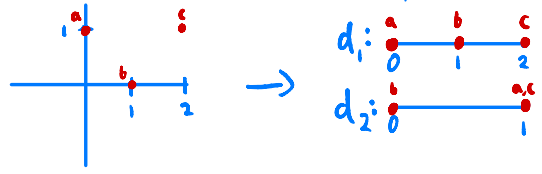
So  $d \in \text{Cone}(\text{cut}(V))$

Poly-time computation trivial

Proof of Theorem

$\mathcal{L}_1(V) \subseteq \text{Cone}(\text{cut}(V))$

If  $d$  is an  $\mathcal{L}_1$  metric then  $d = \sum_i d_i$  where  $d_i$  is a line metric

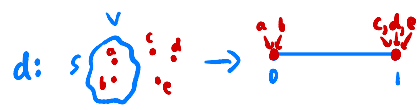


But each  $d_i \in \text{Cone}(\text{cut}(V))$  by claim so by observation  $\sum_i d_i = d \in \text{Cone}(\text{cut}(V))$

Poly-time computation trivial assuming  $d$  embedded into  $\mathcal{L}_1$  w/ poly dimension

$\text{Cone}(\text{cut}(V)) \subseteq \mathcal{L}_1(V)$

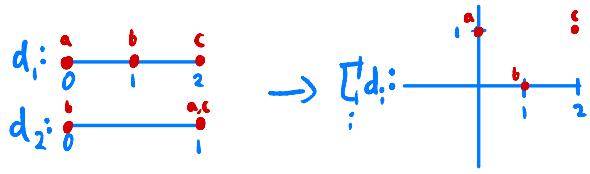
Observe if  $d$  is a cut metric, it is also a line metric



Also if  $d$  is a line metric then so is  $\alpha \cdot d \quad \forall \alpha \geq 0$



Also  $\sum_i d_i$  for  $d_i$  a line metric is an  $\mathcal{L}_1$  metric



So  $d \in \text{Cone}(\text{cut}(V)) = \sum_i \alpha_i d_i$  so  $d \in \mathcal{L}_1(V)$

