Metric Embeddings:
\[ d(x, y) = 0 \quad \forall x, y \]
\[ d(x, y) + d(y, z) \geq d(x, z) \]

Examples & metrics:
1. Points on a plane
\[ d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2| \]
   \[ \Rightarrow d_1 \text{- metric} \]
2. Extend to higher dimensions (d > 2)
   \[ d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \]
   \[ \Rightarrow d_2 \text{- metric} \]
   \[ d((x_1, y_1), (x_2, y_2)) = \sum (|x_1 - x_2| + |y_1 - y_2|) \]
   \[ \Rightarrow d_p \text{- metric} \]
   \[ d((x_1, y_1), (x_2, y_2)) = \max (|x_1 - x_2|, |y_1 - y_2|) \]
   \[ \Rightarrow d_{\infty} \text{- metric} \]

Natural Example of norm metric
\[ d((x_1, y_1), (x_2, y_2)) = (x_1 - x_2)^2 + (y_1 - y_2)^2 \]

Graph Metric
\[ G = (V, E) \text{ have weights on edges} \]
\[ d(u, v) = \text{shortest path w.r.t. weights} \]

Where do metrics appear in algorithm design?
1. Metrics arise as part of the problem!
2. Metrics appear as part of the solution technique

What are Metric Embeddings?

How well do one family of metrics embed into another family of metrics?

Example:
If \( d \) be any 1-point metric \((X, d)\)
embed isometrically into \( d_0 \) metrics
\[ d \subset \mathbb{R} \]
\[ \epsilon \in \{ 1, 2, \ldots \} \]
\[ \begin{align*}
\text{Let } &\quad \mathbf{v}_a = (0, 3, 2, 6, 1) \\
&\quad \mathbf{v}_b = (3, 0, 9, 7, 2) \\
&\quad \mathbf{v}_c = (8, 7, 0, 2, 4) \\
&\quad \mathbf{v}_d = (6, 7, 2, 0, 5) \\
&\quad \mathbf{v}_e = (1, 2, 7, 5, 0)
\end{align*} \]

Example

\[ \mathbf{v}_a - \mathbf{v}_d = (-6, -4, 6, 6, -4) \]

\[ ||\mathbf{v}_a - \mathbf{v}_d||_\infty = 6 \]

\[ \begin{align*}
\text{for } &\quad p \neq q \\
\text{(A)} &\quad ||\mathbf{v}_p - \mathbf{v}_q||_\infty \geq d(p, q) \leq \text{look at } p^{th} \text{ column of } V_p \cap V_q, \\
\text{(B)} &\quad ||\mathbf{v}_p - \mathbf{v}_q||_\infty \leq d(p, q) \\
\text{look at any coordinate } i' \\
|d(p, i') - d(q, i')| &\leq d(p, q) \\
follows because
\end{align*} \]

\[ \begin{align*}
\text{\Delta \text{ bew.}} \begin{cases}
|d(p, i') - d(q, i')| &\leq d(p, q) \\
|d(q, i') - d(p, i')| &\leq d(p, q)
\end{cases}
\end{align*} \]

Provided enough dimensions, \( L_\infty \) metrics are all powerful to capture any finite metric !!
Next question can we 'approximatively' embed all metrics into $L_1$ (or) $L_2$?

Thus [Bourgain]

Any metric can be embedded into an $L_1$ metric with distortion $O(\log n)$.

\[ \exists \ell : x \to \mathbb{R}^d \text{ s.t.} \]
\[ d(x,y) \leq \|f(x) - f(y)\|_1 \leq O(\log n) \cdot d(x,y) \quad \forall x, y \in X. \]

Similarly, any metric also embeds into $L_2$ with distortion $O(\log n)$.

In both cases, $d$ will be reasonably small.

Cool property

- Originally needed $n^2$ entries to compute the metric.
- If we are OK with approximately remembering distances, can do it with $n^2$ entries.

\[ n \cdot \log^2 n \]

**LB:** Can't improve on $O(\log n)$ distortion.

(ie) Each metric in best possible embedding into $L_1$ has distortion $\Omega(\log n)$.

**Dimensionality Reduction**

Given $n$ points in $\mathbb{R}^d$, 

\[ d(v_i, v_j) = \|v_i - v_j\|_2 \]

& suppose $d$ is very very large.

**Theorem:**

If $d \geq 30$, then a mapping $f : \mathbb{R}^d \to \mathbb{R}$
\textbf{Theorem 1} \hspace{1cm} 

\[ f \in \mathcal{C}, \exists \text{ a mapping } f : \mathbb{R}^d \to \mathbb{R}^d \]

\[ \forall i \neq j \]

\[ \| f(x_i) - f(x_j) \|_2 \leq \| x_i - x_j \|_2 \]

\[ \text{where } L = \frac{O(\log n)}{\epsilon^2} \]

\# entries per point comes down from \( d \) to \( L = O(\log n) \)

\textbf{Johnson-Lindenstrauss Lemma}

\textbf{Bourgain Theorem}

Given a metric \((X, d)\) on \( n \) points, we can always embed it into \( \mathbb{R}^d \) with distortion \( \leq \frac{O(\log n)}{\epsilon^2} \) when \( d = \Theta(\log n) \)

\[ \forall x \in X \]

\[ f : X \to \mathbb{R}^d \]

\[ d(x, y) \leq \| f(x) - f(y) \| \leq O(\log n) \cdot d(x, y) \]

\[ \forall x, y \in X \]

We'll carefully choose many subsets \( S \subset X \), & create one coordinate in the embedding

\[ f_1(x) = d(x, s) \]

\[ f(x) = (f_1(x), f_2(x), \ldots, f_d(x)) \]

\[ d(x, s) = \min_{u \in S} d(x, u) \]

\textbf{Claim (1)}

\[ \| f(x) - f(y) \|_1 \leq D \cdot d(x, y) \]

\textbf{Proof}:

\[ \forall \mathcal{S} \subset X \]

\[ \| d(x, s) - d(y, s) \| \leq d(x, y) \]

\[ \Rightarrow \]

\[ d(x, s) - d(y, s) \leq d(x, y) \quad \text{and} \quad d(y, s) - d(x, s) \leq d(x, y) \]

\[ \text{we'll set } D = \Theta(\log n) \]

\textbf{Theorem (2)}

we can carefully choose \( f_1, f_2, \ldots, f_d \) such

\[ \| f(x) - f(y) \|_1 \geq \Theta(\log n) \cdot d(x, y) \]

\textbf{How to choose } \( f_1, f_2, \ldots, f_d \)?

Idea/Intuition
How to compute $x$.

**Intuition**

Consider $B(u, r_u \Delta)$ and $B(v, r_v)$. If $f_3(u) \leq r_u$

& support $S \subseteq X = \{l \mid S \cap B(u, r_u) \neq \phi \}

& S \cap B(v, r_v) = \phi \Rightarrow f_3(v) = r_v$

In this case, $\|f_3(u) - f_3(v)\| \geq \alpha$

**Algorithm**

For $i = 1, 2, \ldots, 100\cdot \Theta \left(\log n\right)$

For $b = 0, 1, \ldots, \log n$

Add $u$ to $S_i^b$ w.p. $\frac{1}{2^b}$ independent for all $u$

For $v : f_3(v) = d(v, S_i^b)$

$\forall v : f(v) = \left\{ \left( f_3(v) \right)_{1 \leq i \leq \Theta \left(\log n\right)} \right\}$

$0 \leq \Theta \left(\log n\right)$

**WTS**

For each $u \in X$

$\|f(u) - f(v)\|_1 \geq \Theta \left(\log n\right) \cdot d(u, v)$

For every $(u, v)$, let

$\gamma_u^l = \text{radius of } B(u, \gamma_u^l)$

$\approx 2^l$ points with
\[
\text{Let } t \text{ be the level at which the 2 tall falls overlap.}
\]
\[
\Rightarrow \quad r_t^u + r_t^l > d(u,v)
\]
\[
\text{WLOG, let } r_t^u > \frac{d(u,v)}{2}
\]

For all \( t \leq b \), we'll show the following:

\[
\| \frac{f(u) - f(v)}{2} \|_1 \geq 0.1 \left( r_u^l - r_u^t \right)
\]

with probability \( \geq 0.1 \)

**Claim 1**

Call \( S_k \) good if \( | f_{S_k}(w) - f_{S_k}(v) | > \Omega(1) (r_u^l - r_u^t) \)

har otherwise.

\[
\mathbb{E} \left[ \# \text{ good sets } S_k \text{ over } i = 1, 2, \ldots, 10^8 \log n \right] \geq 0.1 \times 100 \log n = 10 \log n
\]

Consider bounds

\[
\Rightarrow \quad \exists \text{ with probability } \geq 1 - \frac{1}{n^6}
\]

the \# good sets \( S_k > 5 \log n \)

For each good set, coordinate value

\[
| f_{S_k}(u) - f_{S_k}(v) | > 0.1 (r_u^l - r_u^t)
\]

\[
\Rightarrow \quad \text{total coordinate value over all good sets } S_k
\]

\[
> 0.5 \log n (r_u^l - r_u^t)
\]

sum over all \( 0 \leq i < b \)

\[
\text{total dist } > 0.5 \log n (r_u^l - r_u^0)
\]

\[
> 0.5 \log n \frac{d(u,v)}{2}
\]

Suffices to show Claim 1

\[
\ldots
\]
Call a set $S_k$ good if

$$|f_{S_k}(u) - f_{S_k}(v)| \geq 0.1 (R_u - R_v)$$

Need to show that picking $S_k$ randomly

$$\frac{1}{2e}$$

and giving a good $u$ with prob.

$$\nu \geq 0.1$$

Note that $\forall u \in [0, t)$

$$U_k = \text{Ball}(u, R_u)$$

Case 1

$$r_u < r_v^2$$

Consider $U_k = \text{Ball}(u, R_u)$

and $V_k = \text{Ball}(v, R_v)$

Suppose $S_k$ intersects $U_k$ but completely misses $V_k$

$$d(u, S_k) \leq R_u$$

$$d(v, S_k) > R_v \geq R_u$$

$$\Rightarrow \|f_{S_k}(u) - f_{S_k}(v)\| \geq R_u - R_v$$

$$= \mathbb{P}(S_k \text{ intersects } U_k \& S_k \text{ misses } V_k)
= \mathbb{P}(S_k \text{ misses } V_k) \left(1 - \mathbb{P}(S_k \text{ misses } U_k)\right)
= \left(1 - \frac{1}{2e}\right)^{2e} \left[1 - \left(1 - \frac{1}{2e}\right)^{2e - 1}\right]
\geq \frac{1}{e} \left[1 - \frac{1}{e}\right] \geq 0.1$$

Case 2

$$r_u \geq r_v^2$$

$$\Rightarrow \nu \geq 0.1$$
Case 2a
\[ r_v \leq \frac{r_v^k + r_{v^{k-1}}}{2} \]

Set
\[ U_k = \text{Ball}(u, r_{v^{k-1}}) \] & \[ V_k = \text{Ball}(v, r_{v^k}) \]

We'll get that if
\[ S_k \cap U_k = \emptyset \text{ and } S_k \cap U_k = \emptyset \]
\[ |d(u, S_k) - d(v, S_k)| > \frac{r_v^k - r_v^{k-1}}{2} \]

Case 2b
\[ r_v \leq \frac{r_v^k + r_{v^{k-1}}}{2} \]

In this case, set
\[ U_k = \text{Ball}(u, r_v^k) \] & \[ V_k = \text{Ball}(v, r_v^k) \]

A same proof to show that
\[ S_k \cap U_k = \emptyset \text{ and } S_k \cap U_k = \emptyset \]
\[ |d(u, S_k) - d(v, S_k)| > \frac{r_v^k - r_v^k}{2} \]

Similar proof technique can be used for embedding.

Proof of Claim 1 OED

There is a Frechet Embedding.

Every metric can embed into \( l_1 \) with distortion
\[ \text{dist} \leq O(\log n) \]
Every undirected graph \( G = (V, E) \) with |V| \( \leq O(\log n) \) has a balanced cut. Given a graph \( G = (V, E) \), \( n \) is the size of the graph, and \( E \) is the set of edges. The balanced cut condition is that for any two sets \( S, \bar{S} \) partitioning \( V \), \( |E(S, \bar{S})| \leq \frac{\epsilon}{2} |E| \). This theorem states that for every undirected graph with a certain size, there exists a balanced cut.
\[
\begin{align*}
\overset{\text{Special type of metric}}{= & \quad \alpha(u, v)} \\
\overset{\text{Observation}}{E(3, 3)} & = \sum_{(u, v) \in E} \delta_5(u, v) \\
|S| \delta(S) & = \sum_{u \in V} \delta_5(u, v) \\
\overset{\text{Sparsest cut equivalent to}}{\min} \quad \sum_{(u, v) \in E} \delta(u, v) & \leq \sum_{u \in V} \delta(u, v) \\
\overset{\text{Same algorithm}}{\delta : \delta \text{ is a cut-metric}} & \leq \text{So still NP-hard!} \\
\overset{\text{Same as } l_1 - \text{metric}}{\min} \quad \sum_{(u, v) \in E} \delta(u, v) & \leq \sum_{u \in V} \delta(u, v) \\
\overset{\text{Relax } l_1 \text{ requirement}}{\text{OPT (metric)}} & \leq \text{DPT (metric)} \\
\overset{\text{Special Cut}}{= \quad 1} & \leq \text{Special Cut} \\
\overset{\text{LP II}}{\min} \quad \sum_{(u, v) \in E} \delta(u, v) = 1 \\
\delta(u, v) & \geq 0 \\
\delta(u, v) + \delta(v, w) & \geq \delta(u, w) \\
\delta(u, v) & = \delta(v, u) \\
\delta(u, v) & > 0
\end{align*}
\]
Bourgain's Theorem

eMBED THIS EQUATION HERE $\delta \rightarrow \ell_1$

\[ l_1 \rightarrow \text{cut matrix} \rightarrow \text{good set} \]

Overall Result:

If the embedding has distortion $\alpha$,

then we can get $\alpha$-approximation to Sparsest Cut.

\[ \Rightarrow O(\log n) \text{- approximation to Sparsest Cut.} \]

\[
\min_{\delta: \text{within}} \sum_{(u,v) \in E} \delta(u,v) \quad \text{subject to} \quad \sum_{(u,v) \in \mathcal{E}} \delta(u,v) = 1
\]

\[ \sum_{(u,v) \in \mathcal{E}} \delta(u,v) \geq \delta(u,w) \quad \delta(u,v) \geq 0 \]

P3

P4

$\text{OPT}(P4) = \text{OPT}(P3) \leq \text{OPT}(P1)$

Solve P4 optimally and get $\delta^*$

Use Bourgain's Theorem & embed $\delta^* \rightarrow \ell_1$

with distortion $\log n$.

\[ \|f(u,v) - f(v)\|_1 \leq O(\log n) \delta^*(u,v) \]

\[ \sum_{(u,v) \in \mathcal{E}} \|f(u,v) - f(v)\|_1 \leq O(\log n) \sum_{(u,v) \in \mathcal{E}} \delta^*(u,v) \]

& \[ \sum_{u,v \in V} \|f(u,v) - f(v)\|_1 \geq \sum_{u,v \in \mathcal{E}} \delta^*(u,v) = 1 \]

\[ \Rightarrow \text{Sparsity of my } \ell_1 \text{- solution} \]
\[ \Rightarrow \] Sparsity of my $L_1$-solution

\[ \sum_{(u,v) \in E} \frac{\| f(u) - f(v) \|_2}{\| f(u) - f(v) \|_2} \leq O(\log n) \text{OPT}(14) \]

\[ \Rightarrow \] I've found a set $S$ to $P_2$ with cost

\[ \leq O(\log n) \text{OPT}(14) \]

\[ \Rightarrow \] Can find set to $P_1$ with cost

\[ \leq O(\log n) \text{OPT}(14) \]

---

Equivalence between $L_1$ and cut within:

\[ \sum_{(u,v) \in E} \delta(u,v) = \sum_{(u,v) \in E} \| f(u) - f(v) \|_2 \]

Write $d$ as

\[ \delta = 0.2 \delta_{s_1} + 0.8 \delta_{s_2} + 0.5 \delta_{s_3} + 5 \delta_{s_4} \]

\[ \varphi(u,v), \quad \delta(u,v) = 0.2 \delta_{s_2}(u,v) + 0.8 \delta_{s_2}(u,v) + 0.5 \delta_{s_2}(u,v) + 5 \delta_{s_4}(u,v) \]