An Algorithmist's Toolkit — Brown CSCI 2952T – Spring 2026

Lecture 9: Metric Embeddings, Random Projections, JL, Bourgain's Embedding

April 1, 2026 — Lecturer: D Ellis Hershkowitz — Scribe: Gabriel Chen and Jake Rosenberg

1 Metric Embeddings and Johnson-Lindenstrauss (JL)

An embedding of one metric space (V, d) into another metric space (V', d') is just a function $f : V \to V'$. We say that (V, d) embeds isometrically into (V', d') if there is some f such that

$$d(u, v) = d'(f(u), f(v)) \qquad \forall u, v \in V.$$

This is not always possible. For example, an equilateral triangle in $\mathbb{R}^2$ cannot be embedded isometrically into $\mathbb{R}$ with the Euclidean metric. An embedding f has distortion $\alpha \geq 1$ if

$$d(u, v) \leq d'(f(u), f(v)) \leq \alpha\, d(u, v) \qquad \forall u, v \in V.$$

Theorem: given m points $V \subseteq \mathbb{R}^n$, $\epsilon > 0$, there exists a linear map f embedding (V, d) into (V', d) for $V' \subseteq \mathbb{R}^k$ where $k = O(\frac{\log(m)}{\epsilon^2})$ with distortion $\alpha = \sqrt{\frac{1+\epsilon}{1-\epsilon}} \approx 1 + \epsilon$.

JL-Lemma: There exists a random linear function $\tilde{F} : \mathbb{R}^n \to \mathbb{R}^k$ for $k = 24\frac{\ln m}{\epsilon^2}$ such that for any set of m^2 unit vectors, call it W, we have $||\tilde{F}(w)||^2 \in [1 + \epsilon, 1 - \epsilon]$ except with probability $\frac{2}{m}$.

The lemma implies the theorem, and the general proof idea is to let $W := \{\frac{x-y}{||x-y||} : x, y \in V\}$.

2 Brief interlude on χ^2 distributions

We say $X \sim \chi_k^2$ if

$$X = \sum_{i=1}^{k} Z_i^2$$

where the Z_i's are independent $N(0, 1)$ random variables.

By linearity of expectation,

$$\mathbb{E}[X] = k \qquad \text{if } X \sim \chi_k^2.$$

The following is a result that will be useful later:

$$\Pr\big(|X - \mathbb{E}[X]| \geq \varepsilon k\big) \leq 2 \exp\left(-\frac{\varepsilon^2 k}{8}\right) \qquad \text{for } X \sim \chi_k^2.$$

The proof is in the lecture notes.

3 A result about random projections

Let

$$X = (X_1, \ldots, X_n)$$

where each $X_i \sim N(0,1)$ independently, and define

$$f : \mathbb{R}^n \to \mathbb{R}, \qquad f(w) = \langle X, w \rangle.$$

Claim: if w is a unit vector, then

$$f(w)^2 \sim \chi_1^2.$$

Indeed, if $w = (w_1, \ldots, w_n)$ and $\|w\| = 1$, then

$$f(w) = \sum_{i=1}^{n} w_i X_i.$$

Each $w_i X_i \sim N(0, w_i^2)$, and since sums of independent Gaussians are Gaussian,

$$f(w) \sim N\left(0, \sum_{i=1}^{n} w_i^2\right) = N(0,1)$$

because

$$\sum_{i=1}^{n} w_i^2 = \|w\|^2 = 1.$$

So $f(w)^2 \sim \chi_1^2$, as claimed.

Corollaries: for every unit vector w,

$$\mathbb{E}[f(w)^2] = 1,$$

and

$$\Pr\big(|f(w)^2 - 1| \geq \varepsilon\big) \leq 2 \exp\left(-\frac{\varepsilon^2}{8}\right).$$

4 Proving the JL Lemma

Now let

$$X^{(1)}, X^{(2)}, \ldots, X^{(k)} \in \mathbb{R}^n$$

be independent Gaussian vectors, where

$$X^{(i)} = (X_1^{(i)}, \ldots, X_n^{(i)})$$

and every coordinate is $N(0,1)$.

Define

$$f_i(w) = \langle X^{(i)}, w \rangle$$

and then

$$F(w) = \big(f_1(w), f_2(w), \ldots, f_k(w)\big).$$

This is linear, as if A is the matrix whose i-th row is $X^{(i)}$, then

$$F(w) = Aw.$$

Now if w is a unit vector, then each $f_i(w) \sim N(0,1)$, independently, so

$$\|F(w)\|^2 = \sum_{i=1}^{k} f_i(w)^2 \sim \chi_k^2.$$

Therefore:

$$\mathbb{E}[\|F(w)\|^2] = k,$$

and

$$\Pr\big(|\|F(w)\|^2 - k| \geq \varepsilon k\big) \leq 2\exp\left(-\frac{\varepsilon^2 k}{8}\right).$$

Now define

$$\widetilde{F}(w) = \frac{1}{\sqrt{k}} F(w).$$

Then

$$\mathbb{E}[\|\widetilde{F}(w)\|^2] = \frac{1}{k}\mathbb{E}[\|F(w)\|^2] = 1.$$

Also,

$$\|\widetilde{F}(w)\|^2 \notin [1-\varepsilon, 1+\varepsilon]$$

only if

$$\|F(w)\|^2 \notin [k(1-\varepsilon), k(1+\varepsilon)],$$

which only if

$$|\|F(w)\|^2 - k| \geq \varepsilon k.$$

So

$$\Pr\big(\|\widetilde{F}(w)\|^2 \notin [1-\varepsilon, 1+\varepsilon]\big) \leq 2\exp\left(-\frac{\varepsilon^2 k}{8}\right).$$

Do you think Ellis actually reads these? Now choose

$$k = \frac{24 \ln m}{\varepsilon^2}.$$

Then

$$2\exp\left(-\frac{\varepsilon^2 k}{8}\right) = 2\exp(-3\ln m) = \frac{2}{m^3}.$$

So for any fixed unit vector w,

$$\Pr\big(\|\widetilde{F}(w)\|^2 \notin [1-\varepsilon, 1+\varepsilon]\big) \leq \frac{2}{m^3}.$$

If W is a set of at most m^2 unit vectors, then by the union bound,

$$\Pr\Big(\exists w \in W : \|\widetilde{F}(w)\|^2 \notin [1-\varepsilon, 1+\varepsilon]\Big) \leq m^2 \cdot \frac{2}{m^3} = \frac{2}{m}.$$

So with probability at least $1 - \frac{2}{m}$,

$$\|\widetilde{F}(w)\|^2 \in [1-\varepsilon, 1+\varepsilon] \qquad \forall w \in W.$$

That proves the JL lemma.

5 Bourgain's Embedding (The Setup)

Theorem: Given **any** n-point metric (V, δ), there exists an embedding f with distortion $O(\log n)$ of (V, δ) into $(\tilde{V}, d)$ for $\tilde{V} \subseteq \mathbb{R}^{O(\log^2 n)}$.

Before attempt to prove this Theorem, we need some preliminary definitions.

For $S \subseteq V$, $x \in V$, we define $S(x) = \arg\min_{y \in S} \delta(x, y)$, and $\delta(x, S) = \delta(x, S(x))$.

Suppose we fix $S \subseteq V$ and define $f(x) = \delta(x, S)$ for all $x \in V$. Let $x, y \in V$, $\tilde{x} = f(x)$ and $\tilde{y} = f(y)$.

Claim: $d(\tilde{x}, \tilde{y}) \leq \delta(x, y)$ $\forall x, y \in V$**.**

Proof. We see that

$$\delta(x, s) = \delta(x, S(x)) \leq \delta(x, S(y)) \leq \delta(x, y) + \delta(y, S(y)) = \delta(x, y) + \delta(y, S)$$

Thus we get

$$\delta(\tilde{x}, \tilde{y})| = |\delta(x, S) - \delta(y, S)| \leq \delta(x, y)$$

. $\square$

This looks promising, since we could end up with distortion 1! However, it is not necessarily true that $\delta(\tilde{x}, \tilde{y}) \geq \delta(x, y)$. (For example, pick x and y such that $x \neq y$, but $\delta(x, S) = \delta(y, S)$).

We ask ourselves the following question: *When do distances not go down by too* much?

For example, let $B = \delta(x, y)/2$ and $G = \delta(x, y)/4$. If there is a point in $S \cap G$, and there is no point in $S \cap B$, then $d(\tilde{x}, \tilde{y}) \geq \frac{\delta(x,y)}{4}$.

If $|B| = |G|$ then we can independently for each $v \in V$ let $v \in S$ with probability $\frac{1}{|B|}$, and this will achieve the desired distortion with $\Omega(1)$ probably. However, we can do **better**.

6 Bourgain's Embedding: The Procedure

Let $c \geq 1$ be large

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

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

S_{ij} contains $v \in V$ independently with probability $\frac{1}{2^j}$

$\tilde{x}_{ij} = \delta(x, S_{ij})$ (So, $\tilde{x} \in \mathbb{R}^{O(\log^2 n)}$)

We need to prove that this embedding has distortion $O(\log n)$ with probablity > 0. This hinges on proving something called the Expansion Claim.

Expansion Claim: $\sum_{i,j} |\tilde{x}_{ij} - \tilde{y}_{ij}| \geq (\frac{c}{40} \cdot \log n) \cdot \delta(x, y)$ **for all** x**,** y **expect with** $Pr \leq \frac{1}{n}$.

(Note: Can show Bourgain from Expansion Claim using Cauchy-Schwartz (in the lecture notes)

7 Proving The Expansion Claim

Proof. Fix $x, y \in V$.

For $j = 1, 2, 3, \dots$, let $r_j = \min r$ such that $|B(x, r)|, |B(y, r)| \geq 2^j$.

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

We're going to assume that $2r_t = \delta(x, y)$ *for convenience. The proof doesn't change significantly.*

We claim the following, which will be useful later:

1. $\frac{\delta(x,y)}{2} \leq r_t$
2. $r_j + r_{j-1} \leq 2r_j \leq \delta(x, y)$ for all $j \leq t$
3. $|B^O(x, r_j)|$ or $|B^O(y, r_j)| < 2^j$ for all $j \leq t$
4. $|B(x, r_j)|$ and $|B(y, r_j)| \geq 2^j$ for all $j \leq t$

WLOG suppose $|B^O(x, r_j)| < 2^j$. Let

$$B_j = B^O(x, r_j), G_j = B(y, r_{j-1})$$

We define the following event:

$$E_i = (G_i \cap S_{ij} \neq \emptyset \text{ and } B_j \cap S_{ij} = \emptyset)$$

Note that if E_i, then

$$|\tilde{x}_{ij} - \tilde{y}_{ij}| \geq r_j - r_{j-1}$$

Claim: $Pr(E_i) \geq \frac{1}{10}$

Proof. We know that $|B_j| < 2^j$ and $|G_j| \geq 2^{j-1}$. This means that

$$\begin{aligned} Pr(G_j \cap S_{ij} \neq \emptyset) &\geq 1 - \left(1 - \frac{1}{2^j}\right)^{2^j} \\ &\geq 1 - \exp\left(-\frac{2_{j-1}}{2^j}\right) \\ &= 1 - \exp(-1/2) \\ &= 1 - \frac{1}{\sqrt{e}} \end{aligned}$$

and

$$\begin{aligned} Pr(B_j \cap S_{ij} = \emptyset) &\geq \left(1 - \frac{1}{2^j}\right)^{2^j} \\ &= \left(1 - 0.63\frac{1}{0.63 \cdot 2^j}\right)^{2^j} \\ &\geq \exp\left(-\frac{1}{0.63}\right) \end{aligned}$$

Thus,

$$Pr(E_i) \geq \left(1 - \frac{1}{\sqrt{e}}\right) \cdot \exp\left(-\frac{1}{0.63}\right) \geq \frac{1}{10}$$

□

Let $x_i = \mathbf{1}(E_i)$. This means that

$$x = \sum_{i=1}^{c \log n} x_i$$

Then

$$\mathbb{E}[x] \geq \frac{c \log n}{10}$$

so by a Chernoff bound,

$$Pr\left(x \leq \frac{c}{20} \cdot \log n\right) \leq \frac{1}{n^4}$$

By Union Bound, this holds for all $j \leq t$ except with $Pr \leq \frac{t}{n^4} \leq \frac{1}{n^3}$. So,

$$\begin{aligned} \sum_{i,j} |\tilde{x}_{i,j} - \tilde{y}_{i,j}| &\geq \sum_j \frac{c}{20} \cdot \log n \cdot (r_j - r_{j-1}) \\ &= \frac{c}{20} \cdot \log n \cdot \sum_j (r_j - r_{j-1}) \end{aligned}$$

$$= \frac{c}{20} \cdot \log n \cdot r_t$$
$$\geq \frac{c}{40} \cdot \log n \cdot \delta(x, y)$$

except with $Pr \leq \frac{1}{n^3}$.

So by another Union Bound,

$$\sum_{i,j} |\tilde{x}_{ij} - \tilde{y}_{i.j}| \geq \frac{c}{40} \cdot \log n \cdot \delta(x, y)$$

for all x, y except with $Pr \leq \frac{1}{n}$, which proves the expansion claim. $\square$