Lecture 34

Recall that we want to do dim. reduction to a set of vec. while preserving their pairwise distance.

Thm (JL lem) For $0<\varepsilon<1$, and a set X of n points in $\mathbb{R}^d$, then there exists a map $f: \mathbb{R}^d \to \mathbb{R}^m$, where $m = O\left(\frac{\log n}{\varepsilon^2}\right)$, s.t. $\forall\, u, v \in X$,

$$\| f(u) - f(v) \|_2^2 = (1 \pm \varepsilon) \| u - v \|_2^2$$

Rmk: The $O\left(\frac{\log n}{\varepsilon^2}\right)$ bound in JL. lem. is optimal

Larsen-Nelson, Alon-Klartag [FOCS 2017].

Thm (Distributional JL)

For $0<\delta, \varepsilon<1$, there exists a linear map $\Pi \in \mathbb{R}^{m \times d}$ for $m = O\left(\frac{\log 1/\delta}{\varepsilon^2}\right)$, such that for any $x \in \mathbb{R}^d$

$$\mathbb{P}\left(\| \Pi x \|_2^2 = (1 \pm \varepsilon) \| x \|_2^2 \right) \geq 1 - \delta.$$

— Application to least square regression —

$$\boxed{\text{Given } \begin{cases} A \in \mathbb{R}^{n\times d} \\ y \in \mathbb{R}^{n} \end{cases}}$$

$$A = \begin{pmatrix} — & a_1 & — \\ & \vdots & \\ — & a_n & — \end{pmatrix} \qquad \begin{array}{c} y \\ \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \end{array}$$

$a_1, \dots, a_n \in \mathbb{R}^d$ data pts

Find x s.t. $Ax \approx y$

$$\underset{\in\mathbb{R}^d}{a_i} \longrightarrow \overset{\text{response}}{y_i \in \mathbb{R}}$$

$$\boxed{\min_{x\in\mathbb{R}^d} \|Ax - y\|_2^2}$$

← original problem

Let x^* be an optimal

$$\boxed{\min_{x\in\mathbb{R}^d} \|\Pi A x - \Pi y\|_2^2}$$

← sketched problem

Let $\tilde{x}^*$ be an optimal

<u>Claim</u>: If for any vector of the form $Ax - y$, we have $\|\Pi Ax - \Pi y\|_2^2 = (1\pm\varepsilon)\|Ax-y\|_2^2$

then $\tilde{x}^*$ gives a good approximate for x^*.
the sketched optimal

<u>pf</u>: <u>NTS</u> $\|A\tilde{x}^* - y\|_2^2 \leq (1+o(1))\|Ax^* - y\|_2^2$

- Optimality of $\tilde{x}^*$: $\forall x \in \mathbb{R}^d$

$$\|\Pi A\tilde{x}^* - \Pi y\|_2^2 \leq \|\Pi Ax - \Pi y\|_2^2$$

$$\|A\tilde{x}^* - y\|_2^2 \overset{\Pi}{\leq} (1+\varepsilon)\|\Pi A\tilde{x}^* - \Pi y\|_2^2$$

$$\overset{\text{optimality of } \tilde{x}^*}{\leq} (1+\varepsilon)\|\Pi A \underline{x^*} - \Pi y\|_2^2$$

$$\overset{\Pi}{\leq} (1+\varepsilon)^2 \|Ax^* - y\|_2^2 \qquad \blacksquare$$

The distributional JL can preserve a single vector's length. But there are so many vectors of the form $Ax-y$, the union bound cannot work.

<u>Idea</u>: We construct an ε-net for the subspace spanned by col^n of A & y, and apply union bound over the ε-net.

Subspace embedding

<u>Thm 1</u> Let $U \subseteq \mathbb{R}^n$ be a d-dim. linear subsp. of $\mathbb{R}^n$ and let $\Pi \in \mathbb{R}^{m\times n}$ be the matrix from distributional JL. Then w/ probability $\geq 1-\delta$,

$\forall v \in U, \quad \|\Pi v\|_2^2 = (1 \pm \varepsilon)\|v\|_2^2$

$x \in \mathbb{R}^d$

$Ax, y \in \mathbb{R}^n$

provided $\quad m = O\left(\frac{d \log \frac{1}{\varepsilon} + \log \frac{1}{\delta}}{\varepsilon^2}\right)$

$\underset{m \times n}{\Pi} \underset{n \times d}{A} x - \Pi y$

Thm 1 implies that we can find a projection $\Pi \in \mathbb{R}^{m \times n}$ w/ $m = O\left(\frac{d \log \frac{1}{\varepsilon} + \log \frac{1}{\delta}}{\varepsilon^2}\right)$ s.t. $\forall$ vector $Ax - y$,

$$\|\Pi Ax - \Pi y\|_2^2 = (1 \pm \varepsilon)\|Ax - y\|_2^2$$

Take U = spanned by col^{n} A & y

$\dim U = d + 1$

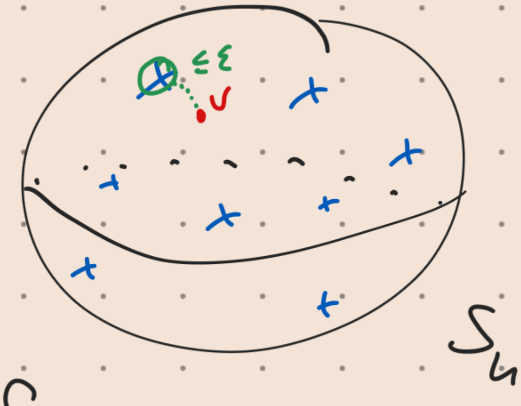
To prove Thm 1, first observe that as Π is linear, it suffices to consider unit vectors $v \in U$. Write S_U for the unit sphere

$$S_U = \{v \in U : \|v\|_2^2 = 1\}.$$

We shall find an ε-net N_ε for S_U. that is, $\forall v \in S_U$, $\exists x \in N_\varepsilon$ s.t.

$$\|v - x\|_2^2 \le \varepsilon$$

Lem 2 $\forall\, 0<\varepsilon<1$, $\exists$ an ε-net $N_\varepsilon \subseteq S_u$ w/

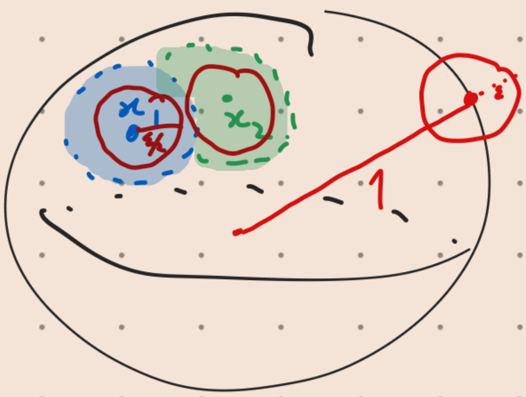


$$|N_\varepsilon| \le \left(\frac{4}{\varepsilon}\right)^d \quad \text{s.t. } \forall\, u \in S_u$$

$$\min_{x \in N_\varepsilon} \|x-u\|_2^2 \le \varepsilon.$$

Pf Sketch: • Iteratively pick $x_1, x_2, \dots$ s.t. x_i's pairwise distance $\geq \varepsilon$.

Let N_ε be the maximal set of such x_i's.



all $B(x_i, \varepsilon/2)$ disjoint

$$\subseteq B(0, 1+\varepsilon/2)$$

$$\Rightarrow \quad |N_\varepsilon| \cdot \mathrm{vol}_d(\varepsilon/2) \le \mathrm{vol}_d(1+\varepsilon/2)$$

$$\mathrm{vol}_d(r) = c \cdot r^d$$

$$|N_\varepsilon| \le \frac{\mathrm{vol}_d(1+\varepsilon/2)}{\mathrm{vol}_d(\varepsilon/2)} = \left(\frac{1+\frac{\varepsilon}{2}}{\frac{\varepsilon}{2}}\right)^d \le \left(\frac{4}{\varepsilon}\right)^d \quad \blacksquare$$

Pf (Thm 1) Let N_ε be an ε-net for S_u. By distribution JL lemma and union bound, for

$$m = O\left(\frac{\log |N_\varepsilon|/\delta}{\varepsilon^2}\right) = O\left(\frac{d \log\frac{1}{\varepsilon} + \log\frac{1}{\delta}}{\varepsilon^2}\right),$$

there is a linear $\Pi \in \mathbb{R}^{m\times n}$ s.t. w/ $\mathbb{P} \geq 1-\delta$

$$\forall x \in N_\varepsilon, \quad \|\Pi x\|_2^2 = (1\pm\varepsilon)\|x\|_2^2$$

$N_\varepsilon \subseteq S_u$

NTS $\forall v \in S_u$, $\|\Pi v\|_2^2 = (1\pm O(\varepsilon))\underbrace{\|v\|_2^2}_{=1} = 1 \pm O(\varepsilon)$

$\boxed{\text{HW:}}$ $\forall v \in S_u$, $\exists$ a seq of pts $x_0, x_1, x_2, \dots \in N_\varepsilon$ s.t.

$$v = x_0 + c_1 x_1 + c_2 x_2 + \cdots$$

for $|c_i| \leq \varepsilon^i$.

Recall $x_i \in N_\varepsilon \subseteq S_u$, $\|x_i\|_2^2 = 1$

$$\|\Pi v\|_2 = \|\Pi x_0 + c_1 \Pi x_1 + c_2 \Pi x_2 + \cdots\|_2$$

$$\overset{\Delta\text{-ineq}}{\leq} \|\Pi x_0\|_2 + c_1\|\Pi x_1\|_2 + c_2\|\Pi x_2\|_2 + \cdots$$

$$\overset{\text{choice of }\Pi}{\leq} (1+\varepsilon)\overset{1}{\cancel{\|x_0\|_2^2}} + \varepsilon(1+\varepsilon) + \varepsilon^2(1+\varepsilon) + \cdots$$

$$\leq 1 + O(\varepsilon).$$

Similarly: $\|\Pi v\|_2^2 \geqslant 1 - O(\varepsilon)$