

15-750

4/10/19

Johnson-Lindenstrauss (JL)

Def $f: X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^d$ is an isoperometric embedding
if $\forall p, q \in X \quad \|p - q\|_2 = \|f(p) - f(q)\|_2$

Def $f: X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^d$ is k-bi-Lipschitz

if $\exists c$ s.t. $c \|p - q\|_2 \leq \|f(p) - f(q)\|_2 \leq c \cdot k \cdot \|p - q\|_2$

Such f is called a k-embedding

Thm (Johnson-Lindenstrauss)

Let $X \subseteq \mathbb{R}^m$, $|X| = n$, $0 < \epsilon \leq 1$, $\exists (1 + \epsilon)$ -embedding
of X into \mathbb{R}^d for $d = O(\epsilon^{-2} \log n)$.

Note May assume $m \leq n-1$

Consider affine subspace S spanned by

$$p_1, \dots, p_n \in \mathbb{R}^m.$$

$$\dim(S) \geq n-1.$$

Claim: Project a random $x \in S^{n-1}$ onto first $\log n$ coordinates we get concentration.

Def $f(x) = \sqrt{x_1^2 + \dots + x_k^2}$ for $x \in S^{n-1}$

Lemma For $x \in S^{n-1}$ picked uniformly then

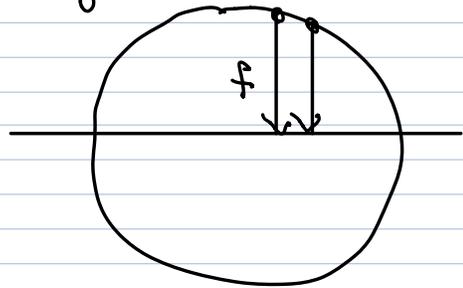
$\exists m = (n, k)$ st.

$$1) P_r[f(x) \geq m+t], P[f(x) \leq m-t] \leq 2e^{-\frac{t^2 n}{2}} \quad 0 < t < 1$$

$$2) \text{ If } k \geq 10 \ln n \text{ then } m \geq \frac{1}{2} \sqrt{k/n}$$

Note f is the projection of x onto first k coord. & projections are 1-Lipschitz.

eg



Thus by Levy $m = \text{med}(f)$ prove 1). To show 2).

Let $X = (X_1, \dots, X_n)$ random point on S^{n-1}

Consider random variables $X_1^2, X_2^2, \dots, X_n^2$.

$$\begin{aligned} \text{Now } X^2 &= \sum X_i^2 \quad \& \quad 1 = \mathbb{E}(X^2) = \sum \mathbb{E}(X_i^2) \\ &= \frac{1}{n} \mathbb{E}(X_j^2) \Rightarrow \mathbb{E}(X_j^2) = 1/n \quad \forall j \end{aligned}$$

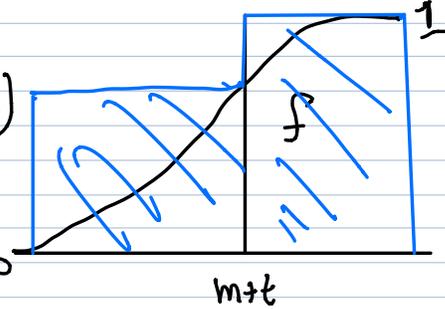
For $f(x) = \sqrt{x_1^2 + \dots + x_k^2}$

$$\mathbb{E}(f^2(x)) = \mathbb{E}\left[\sum_{i=1}^k X_i^2\right] = \sum \mathbb{E}(X_i^2) = k/n$$

To show $\text{med}(f) \geq \frac{1}{2} \sqrt{k/n}$ for $k \geq 10 \ln n$ 3

Claim
 $\frac{k}{n} = E(f^2) \leq \Pr[f \leq m+t] (m+t)^2 + \Pr[f \geq m+t] \cdot 1$

$\leq 1 \cdot (m+t)^2 + 2e^{-t^2 n/2}$ (bery)



Set $t = \sqrt{k/5n}$ & $k \geq 10 \ln n$

We get

$$2e^{-t^2 n/2} = 2e^{-\frac{k}{5n \cdot 2}} \leq 2e^{-\frac{10 \ln n}{10}} = 2/n$$

thus $\frac{k}{n} \leq (m + \sqrt{k/5n})^2 + (2/n)$

Check that $\Rightarrow m \geq \frac{1}{2} \sqrt{k/n}$

Lemma says random $X \in S^{n-1}$ projects onto b^k with we get concentration.

We now flip this around: Fix $X \in S^{n-1}$ and pick random k -dim subspace.

Case $k=1$ we need to generate random direction

Use use mult-dim normal

$$\begin{aligned} \text{Pdf}(N^n(0,1)) &= (2\pi)^{-n/2} e^{-x_1^2 - \dots - x_n^2} \\ &= (2\pi)^{-n/2} e^{-r^2} \end{aligned}$$

Thm (Johnson-Lindenstrauss)

Let $X \subseteq \mathbb{R}^m$, $|X|=n$, $0 < \epsilon \leq 1$, $\exists (1+\epsilon)$ -embedding of X into \mathbb{R}^d for $d = O(\epsilon^{-2} \log n)$.

pf

Claim with prob $\geq 1 - 1/n^2$

$$(1 - \epsilon/3) \text{med } \|x-y\| \leq \|P_f(x) - P_f(y)\| \leq (1 + \epsilon/3) \text{med } \|x-y\| \quad (*)$$

Claim \Rightarrow Thm

Since $|X|=n \exists \binom{n}{2}$ pairs to satisfy. Thus by

Union bd Prob failure $\leq 1/n^2 \Rightarrow$

Prob failure for some pair $\leq 1/2$.

Let $u = x-y$ then (*) is

$$m(1 - \epsilon/3) \|u\| \leq \|P_f(u)\| \leq m(1 + \epsilon/3) \|u\|$$

$$\text{wlog } \|u\|=1$$

$$\text{to show } \left| \|P_f(u)\| - m \right| \leq \frac{\epsilon}{3} m \quad (**)$$

Now $\|P_f(u)\| = f(u)$ from Lemma

Thus prob of (**) failure is

$$P_f \left[|f(u) - m| \geq \frac{\epsilon}{3} m \right]$$

By Lemma we get $\Pr[|f(u) - m| \geq t] \leq 4e^{-t^2 n / 2}$ for $0 < t < 1$

We pick $t = \varepsilon m / 3$

$$\Pr\left[|f(u) - m| \geq \frac{\varepsilon}{3} m\right] \leq 4 e^{-\frac{\varepsilon^2}{18} mn} \quad (***)$$

Since $m \geq \frac{1}{2} \sqrt{k/n}$ for $k = 200 \varepsilon^{-2} \ln n$

$$(***) \leq 4 e^{-\varepsilon^2 k / 72} \leq 4 e^{-200 / 72 \ln n} \leq 1/n^2$$

Stronger versions of JL.

6

Recall Our Alg:

Input $X \subseteq \mathbb{R}^m$ $|X|=n$ viewed as matrix $X^{m \times n}$ columns

We constructed random orthonormal matrix $M^{k \times m}$

by 1) Constructing a $M^{k \times m}$ matrix of iid Gaussian's

2) Apply Gram-Schmit to M making it orthonormal.

Output MX (Possibly scaling the vectors)

Simpler constructions also work! (Give JL)

1) Step 2) not necessary (Matrix of iid Gaussian's)

Intuition: $M^{k \times m}$ is very close to orthonormal

2) A matrix M of iid ± 1 's also works.

3) Many other dist also work (Sub Gaussians)