MTH 995-003: Intro to CS and Big Data	Spring 2014

Lecture 14 — Feb 20, 2014

Inst. Mark Iwen *Scribe: Amineh Farzannia*

Contents

1 The Johnson-LindenStrauss Lemma

Theorem 1. *Let $P \subseteq \mathbb{C}^N$, where P is a finite set and $|P| = M \in \mathbb{N}$, be an arbitrary set of points. Let $p, \epsilon \in (0,1)$. Finally, let $A \in \mathbb{R}^{m \times N}$ have i.i.d (independent identically distributed)mean 0,variance 1,subgaussian entries with parameter c. Then*

$$(1-\epsilon)\|\vec{x}-\vec{y}\|_2^2 \leq \left\| \frac{1}{\sqrt{m}} A(\vec{x}-\vec{y}) \right\|_2^2 \leq (1+\epsilon)\|\vec{x}-\vec{y}\|_2^2 \qquad (*)$$

for all $\vec{x}, \vec{y} \in P$ with $\vec{x} \neq \vec{y}$ with probability at least p, provided that

$$m \geq \frac{8c(16c+1)}{\epsilon^2} \ln\left(\frac{4M^2}{1-p}\right)$$

Here A is entirely independent of P, yet it preserves its “intrinsic geometry” with high probability!

1.1 Homework Problem 5- due Feb 25th(Tue.)

Homework 5: Prove theorem 1.
Hint: Use theorem 2 from lecture 13 and union bound.
Explicit version for a” best c” can be obtained by using lemma 1 from lecture 10.

Definition 1. *A matrix satisfying (*) is called a Johnson-Lindenstrauss (J-L)embedding for P.*

Corollary 1. *A “strict” Johnson-Lindenstrauss embedding of P that satisfies*

$$(1-\epsilon)\|\vec{x}\|_2^2 \leq \left\| \frac{1}{\sqrt{m}} A\vec{x} \right\|_2^2 \leq (1+\epsilon)\|\vec{x}\|_2^2$$

for all $\vec{x} \in P$ exists by theorem 1 for any $P \subset \mathbb{R}^N$. In the lower bound on m we can have $4M^2 \to 2M$.

1.2 J-L embeddings for subspaces of $\mathbb{R}^N$

Definition 2. *Let $T \subset \mathbb{R}^N$. A δ-cover of T is an $S_\delta \subset T$ such that*

$$\forall \vec{x} \in T, \exists \vec{y} \in S_\delta \ s.t. \ \|\vec{x} - \vec{y}\|_2 \leq \delta \Longleftrightarrow T \subseteq \bigcup_{\vec{y} \in S_\delta} B(\vec{y}, \delta).$$

Definition 3. *The δ-covering number of $T \subset \mathbb{R}^N$, $C_\delta(T)$, is the smallest number such that a δ-cover of T with cardinality $C_\delta(T)$ exists.*

Definition 4. *Let $T \subset \mathbb{R}^N$. A δ-packing of T is a set $P_\delta \subset T$with the property that*

$$\|\vec{x} - \vec{y}\| \geq \delta \qquad \forall \vec{x}, \vec{y} \in P_\delta \subset T, \ \text{with } \vec{x} \neq \vec{y}$$

Definition 5. *The δ-packing number of T, denoted by $P_\delta(T)$, is the largest achievable cardinality of any δ-packing P_δ of T.*

Lemma 1. *Let $T \subset \mathbb{R}^N$,and $\delta \in \mathbb{R}^+$,then*

$$P_{2\delta}(T) \leq C_\delta(T) \leq P_\delta(T).$$

Proof: Let $P_{2\delta} \subset T$ be a maximal 2δ-packing of T. Each $\vec{x} \in P_{2\delta}$ is closest to a different $\vec{y}$ in any minimal δ-cover of T. This defines an injection from $P_{2\delta} \subset T$ into ant minimal δ-cover of T. This proves the first inequality. For the second inequality, we note that every maximal δ-packing of T is also a δ-cover of T, lest we can increase its size (a contradiction). □

Lemma 2. *Let $B = B(\vec{0}, 1)$be the unit ball in $\mathbb{R}^N$.Then*

$$C_\delta(B) \leq \left(1 + \frac{2}{\delta}\right)^N \qquad \forall \delta \in \mathbb{R}^+.$$

Proof: Let $\{\vec{x}_1, \vec{x}_2, ..., \vec{x}_P\}$be a maximal δ-packing of B. Note that $P = P_\delta(B)$. Then,

$$\text{Vol}\left(\cup_{l=1}^p B(\vec{x}_l, \delta/2)\right) = P \cdot \text{Vol}\left(\frac{\delta}{2} B\right) \leq \text{Vol}\left(\left(1 + \frac{\delta}{2}\right) B\right)$$

Thus,

$$P \left(\frac{\delta}{2}\right)^N \text{Vol}(B) \leq \left(1 + \frac{\delta}{2}\right)^N \text{Vol}(B),$$

implying

$$P \leq \left(1 + \frac{2}{\delta}\right)^N.$$

Lemma 1 now tells us that

$$C_\delta(B) \leq \left(1 + \frac{2}{\delta}\right)^N.$$

□

Lemma 3 (See [1]). *Let H be an arbitrary k-dimensional subspace of $\mathbb{R}^N$. Let $p, \epsilon \in (0,1)$, and let $A \in \mathbb{R}^{m\times N}$ be a matrix with i.i.d, mean 0, variance 1,subgaussian entries with parameter c. Choose*

$$m \geq \frac{32c(16c+1)}{\epsilon^2} \cdot k \cdot \ln\left(\frac{(1+\frac{16}{\epsilon})}{\sqrt[k]{(1-p)/2}}\right).$$

Then,

$$(\dagger) \qquad (1-\epsilon)\|\vec{x}\|_2 \leq \left\|\frac{1}{\sqrt{m}}A\vec{x}\right\|_2 \leq (1+\epsilon)\|\vec{x}\|_2$$

will hold $\forall \vec{x} \in H$ with probability at least p.

Proof: Let $S \subseteq H$ be the $(k-1)$-dimension unit sphere in H,and B be the k-dimension unit ball in H. Choose a minimal $\frac{\epsilon}{8}$-cover of B,

$$C_{\frac{\epsilon}{8}} \subset B.$$

We now apply theorem 1 to produce a strict J-L embedding of $C_{\frac{\epsilon}{8}}$ with $\epsilon \to \epsilon/2$ (see Corollary 1). The bound for m above now holds by Corollary 1 and Lemma 2 (here $\epsilon \to \frac{\epsilon}{2}$ and $4M^2 \to 2|C_{\frac{\epsilon}{8}}| \leq 2\left(1+\frac{16}{\epsilon}\right)^k$ in theorem 1). We have with probability more than p that

$$(**) \qquad (1-\epsilon/2)\|\vec{x}\|_2^2 \leq \left\|\frac{1}{\sqrt{m}}A\vec{x}\right\|_2^2 \leq (1+\epsilon/2)\|\vec{x}\|_2^2.$$

holds for all $\vec{x} \in C_{\frac{\epsilon}{8}}$. We want to show that this weaker result of (**)implies (†) $\forall \vec{x} \in H$.

Note. It is enough to establish (†) for all $\vec{x} \in S$ since A and H are linear.

Let δ be the smallest number so that

$$\left\|\frac{1}{\sqrt{m}}A\vec{x}\right\|_2 \leq (1+\delta)\|\vec{x}\|_2 = 1+\delta$$

holds $\forall \vec{x} \in S$

Choose any $\vec{x} \in S$, and let $\vec{y} \in C_{\frac{\epsilon}{8}}$be such that $\|\vec{x}-\vec{y}\| \leq \frac{\epsilon}{8}$.
We have that

$$\left\|\frac{1}{\sqrt{m}}A\vec{x}\right\|_2 \leq \left\|\frac{1}{\sqrt{m}}A\vec{y}\right\|_2 + \left\|\frac{1}{\sqrt{m}}A(\vec{x}-\vec{y})\right\|_2 \leq \left(\sqrt{1+\frac{\epsilon}{2}}\right)\|\vec{y}\|_2 + (1+\delta)\cdot\frac{\epsilon}{8} \leq 1+\frac{\epsilon}{2}+(1+\delta)\cdot\frac{\epsilon}{8}.$$

Now by definition of δ, we have

$$(1+\delta) \leq 1+\frac{\epsilon}{2}+(1+\delta)\cdot\frac{\epsilon}{8}.$$

This implies that $\delta \leq \epsilon$, and gives us the upper bound in (†).

You can obtain the lower bound by noting that any $\vec{x} \in S$ will have

$$\left\|\frac{1}{\sqrt{m}}A\vec{x}\right\|_2 \geq \left\|\frac{1}{\sqrt{m}}A\vec{y}\right\|_2 - \left\|\frac{1}{\sqrt{m}}A(\vec{x}-\vec{y})\right\|_2 \geq (1-\epsilon/2)\left(1-\frac{\epsilon}{8}\right) - (1+\epsilon)\cdot\frac{\epsilon}{8} \geq 1-\epsilon.$$

□

References

[1] Richard Baraniuk, and Mark Davenport, and Ronald DeVore, and Michael Wakin. *A Simple Proof of the Restricted Isometry Property for Random Matrices.* Constructive Approximation, vol. 28, num. 3, pp. 253–263, 2008.