MTH 995-003: Intro to CS and Big Data	Spring 2015
Lecture 22 — April 7, 2015	
Inst. Mark Iwen	*Scribe: Sami Merhi*

1 Overview

In the previous two lectures we showed how the Johnson-Lindenstrauss lemma implies the RIP (**restricted isometery property**), and discussed the lower bound on the number of rows of RIP matrices.

In this lecture, and the one that follows, we consider the converse, i.e. we show that the RIP yields **J-L embeddings** of sets.

2 Main Theorem

The goal is to prove the following theorem.

Theorem 1. *Let $P \subseteq \mathbb{R}^N$ with $|P| = M$. Suppose $A \in \mathbb{R}^{m\times N}$ has a RIP constant $\epsilon_{2k}(A) \leq \frac{\eta}{4}$ for some $\eta, \delta \in (0,1)$ and $k \geq 16\ln\left(\frac{4M}{\delta}\right)$. Let $\vec{\psi} \in \mathbb{R}^N$ have independent, identically distributed Bernoulli random entries ($\psi_i = \pm 1$ each with probability $\frac{1}{2}$). Then*

$$(1-\eta)\|\vec{x}\|_2^2 \leq \left\| A Diag\left(\vec{\psi}\right)\vec{x}\right\|_2^2 \leq (1+\eta)\|\vec{x}\|_2^2$$

for all $\vec{x} \in P$, with probability at least $1-\delta$.

Remark 1. By $\text{Diag}(\vec{x})$ we mean the $N \times N$ diagonal matrix X, with $X_{i,i} = x_i$, the i^{th} component of $\vec{x}$.

We will need to prove a couple of lemmas first, before proving the main theorem above. The difficulty with a direct approach is illustrated as follows.

Let $A \in \mathbb{R}^{m\times N}$, and $\tilde{A} = A \cdot \text{Diag}\left(\vec{\psi}\right)$. Then

$$\tilde{A}^*\tilde{A} = \text{Diag}\left(\vec{\psi}\right) A^* A \text{Diag}\left(\vec{\psi}\right).$$

The main difficulty is that $\vec{\psi}$ shows up on both sides of this product, making it difficult to use "independence" results.

3 A couple of Lemmas

Lemma 1. ***(Decoupling)*** *Let $B \in \mathbb{R}^{N\times N}$ be symmetric with zeros on its diagonal. Let $\vec{\psi} \in \mathbb{R}^N$ be a vector of independent, mean zero random variables. If $f : \mathbb{R} \to \mathbb{R}$ is a convex function, then*

$$E := \mathbb{E}\left[f\left(\vec{\psi}^* B\vec{\psi}\right)\right] \le \mathbb{E}\left[f\left(4\vec{\psi}^*\left(B + Diag\left(\vec{y}\right)\right)\vec{\psi}'\right)\right]$$

where $\vec{\psi}'$ is an independent copy of $\vec{\psi}$, and $\vec{y} \in \mathbb{R}^N$ is arbitrary.

The proof of this wonderful lemma is due to the French mathematician **Michel Talagrand**.

Proof. Let $\vec{\delta} \in [0,1]^N$ be a random vector with independent, identically distributed entries, such that

$$\delta_j = \begin{cases} 1 & \text{with probability } \frac{1}{2}, \\ 0 & \text{with probability } \frac{1}{2}. \end{cases}$$

Then

$$\mathbb{E}\left[\delta_k\left(1-\delta_i\right)\right] = \frac{1}{4},\ j \ne k.$$

We have

$$\begin{aligned} E &:= \mathbb{E}\left[f\left(\vec{\psi}^* B\vec{\psi}\right)\right] \\ &= \mathbb{E}_{\vec{\psi}}\left[f\left(4\sum_{j\ne k}\mathbb{E}_{\vec{\delta}}\left[\delta_j\left(1-\delta_k\right)\right]\psi_j B_{jk}\psi_k\right)\right], \end{aligned}$$

where we have used the fact that $B_{k,k} = 0$ for all $k = 1, \dots, N$.

By convexity of f and Jensen's Inequality, we get

$$\begin{aligned} E &\le \mathbb{E}_{\vec{\psi}}\mathbb{E}_{\vec{\delta}}\left[f\left(4\sum_{j\ne k}\delta_j\left(1-\delta_k\right)\psi_j B_{jk}\psi_k\right)\right] \\ &= \mathbb{E}_{\vec{\delta}}\mathbb{E}_{\vec{\psi}}\left[f\left(4\sum_{j\ne k}\delta_j\left(1-\delta_k\right)\psi_j B_{jk}\psi_k\right)\right], \end{aligned}$$

where the last equality follows by Fubini's theorem.

Now let

$$\sigma\left(\vec{\delta}\right) := \left\{j \in [N] \mid \delta_j = 1\right\}.$$

Then

$$E \le \mathbb{E}_{\vec{\delta}}\mathbb{E}_{\vec{\psi}}\left[f\left(4\sum_{j\in\sigma\left(\vec{\delta}\right)}\sum_{k\notin\sigma\left(\vec{\delta}\right)}\psi_j B_{jk}\psi_k\right)\right].$$

Notice that ψ_j and ψ_k are now (in the sum) independent of one another! So we may write

$$E \le \mathbb{E}_{\vec{\delta}}\mathbb{E}_{\vec{\psi}}\mathbb{E}_{\vec{\psi}'}\left[f\left(4\sum_{j\in\sigma\left(\vec{\delta}\right)}\sum_{k\notin\sigma\left(\vec{\delta}\right)}\psi_j B_{jk}\psi_k'\right)\right]$$

where $\vec{\psi}'$ is an independent copy of $\vec{\psi}$.

The last inequality implies the existence of some $\vec{\delta}_0 \in [0,1]^N$ such that $\sigma := \sigma\left(\vec{\delta}_0\right)$ satisfies

$$\begin{aligned} E &\leq \mathbb{E}_{\vec{\psi}}\mathbb{E}_{\vec{\psi}'}\left[f\left(4\sum_{j\in\sigma(\vec{\delta})}\sum_{k\notin\sigma(\vec{\delta})}\psi_j B_{jk}\psi'_k\right)\right] \\ &= \mathbb{E}_{\vec{\psi}|_\sigma}\mathbb{E}_{\vec{\psi}'|_{\bar{\sigma}}}\left[f\left(4\sum_{j\in\sigma}\left\{\sum_{k\notin\sigma}\psi_j B_{jk}\psi'_k + \sum_{k\in\sigma}\psi_j B_{jk}\mathbb{E}_{\vec{\psi}'|_\sigma}\left[\psi'_k\right]\right\}\right.\right. \\ &\quad \left.\left. +4\sum_{j\notin\sigma}\mathbb{E}_{\vec{\psi}|_{\bar{\sigma}}}\left[\psi_j\right]\sum_{k=1}^{N}\psi'_k B_{jk}\right)\right]. \end{aligned}$$

This step is justified since $\mathbb{E}_{\vec{\psi}|_{\bar{\sigma}}}\left[\psi_j\right] = \mathbb{E}_{\vec{\psi}'|_\sigma}\left[\psi'_k\right] = 0$ by the assumption on $\vec{\psi}$ and $\vec{\psi}'$.

We are thus allowed to add y_k to every $B_{k,k}$ term that appears in the last two sums; that is

$$\begin{aligned} E &\leq \mathbb{E}\left[f\left(4\sum_{j,k=1}^{N}\psi_j\left(B_{jk} + \mathrm{Diag}\left(\vec{y}\right)_{jk}\right)\psi'_k\right)\right] \\ &= \mathbb{E}\left[f\left(4\vec{\psi}^*\left(B + \mathrm{Diag}\left(\vec{y}\right)\right)\vec{\psi}'\right)\right], \end{aligned}$$

where, once again, the inequality follows from the convexity of f and Jensen's Inequality. □

Lemma 2. ***(Rademacher Chaos)*** *Let $B \in \mathbb{R}^{N\times N}$ be symmetric with zeros on its diagonal, and let $\vec{\psi} \in \mathbb{R}^N$ have independent, identically distributed Bernoulli entries. Then for $t > 0$, we have*

$$\begin{aligned} \mathbb{P}\left(\left|\vec{\psi}^* B\vec{\psi}\right| \geq t\right) &\leq 2\exp\left(-\min\left\{\frac{3t^2}{128\left\|B\right\|_F^2}, \frac{t}{32\left\|B\right\|_2}\right\}\right) \\ &= \begin{cases} 2\exp\left(\frac{-3t^2}{128\|B\|_F^2}\right), & \text{if } 0 < t \leq \frac{4}{3}\frac{\|B\|_F^2}{\|B\|_2}; \\ 2\exp\left(-\frac{t}{32\|B\|_2}\right), & \text{if } t > \frac{4}{3}\frac{\|B\|_F^2}{\|B\|_2}. \end{cases} \end{aligned}$$

Remark 2. Recall that

$$\begin{aligned} \left\|B\right\|_2 &= \max_{\vec{x}\in\mathbb{R}^N}\frac{\left\|B\vec{x}\right\|_2}{\left\|\vec{x}\right\|_2} = \sigma_1\left(B\right), \\ \left\|B\right\|_F^2 &= \sum_{i,j=1}^{N}\left|B_{ij}\right|^2 = \sum_{j=1}^{N}\sigma_j^2\left(B\right), \end{aligned}$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N$ are the singular values of B.