Johnson-Lindenstrauss Lemma and Restricted Isometry Property

Grigor Keropyan

Department of Mathematics
Mathematics in Data Science
Technische Universität München (TUM)

July, 2021

Outline

Johnson-Lindenstrauss Lemma 1

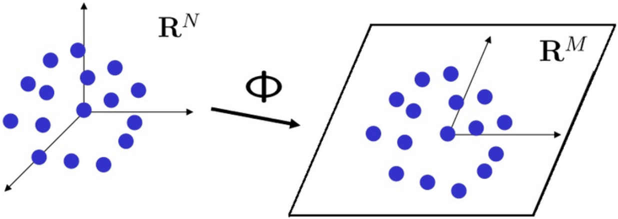

Figure: Mapping from high dimensional space on lower dimensional space

Johnson-Lindenstrauss Lemma 2

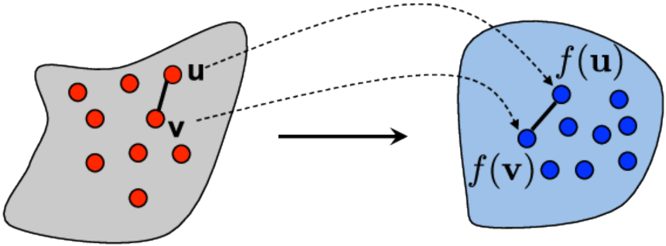

Figure: Mapping from high dimensional space on lower dimension by keeping relative distances

Johnson-Lindenstrauss Lemma 3

Lemma 4 (Johnson-Lindenstrauss) (Lemma 4.1 in (Baraniuk et al., 2008))
Let $\epsilon \in (0,1)$ is given and $N \in \mathbb{N}$. For every set $\mathcal{Q} \in \mathbb{R}^N$, if n is a positive integer such that $n \geq n_0 := O(ln(\#\mathcal{Q})/\epsilon^2)$, there exists a Lipschtz mapping $f : \mathbb{R}^N \to \mathbb{R}^n$ such that

$$(1-\epsilon)\left\|u-v\right\|_2^2 \leq \left\|f(u)-f(v)\right\|_2^2 \leq (1+\epsilon)\left\|u-v\right\|_2^2, \qquad (1)$$

for all $u, v \in \mathcal{Q}$.

Concentration Inequality

Definition 2
Let $\Phi \in \mathbb{R}^{n\times N}$ *be a random matrix and for any* $x \in \mathbb{R}^N$ *expectation of random variable* $\|\Phi(x)\|_2^2$ *is* $\|x\|_2^2$, *that is*

$$\mathbb{E}[\|\Phi(x)\|_2^2] = \|x\|_2^2\,,$$

where $n, N \in \mathbb{N}$. *We say that random matrix* Φ *satisfies concentration inequality if the random variable* $\|\Phi(x)\|_2^2$ *is strongly concentrated about its expected value, that is*

$$\mathbb{P}(|\,\|\Phi(x)\|_2^2 - \|x\|_2^2\,| \geq \epsilon\,\|x\|_2^2) \leq 2e^{-nc(\epsilon)}, \quad 0 \leq \epsilon \leq 1,$$

where ϵ *is a constant and* $c(\epsilon)$ *is also a constant depends only on* ϵ *such that for all* $\epsilon \in (0,1)$, $c(\epsilon) \geq 0$.

Concentration Inequality: Examples 1

Entries $\Phi_{i,j}$ of Φ are independent samples from Normal distribution

$$\{\Phi_{i,j}\}_{i\in[n],j\in[N]} \overset{i.i.d}{\sim} \mathcal{N}(0, \frac{1}{n}).$$

Concentration Inequality: Examples 2

Entries $\Phi_{i,j}$ of Φ are independent samples from following distribution

$$X := \begin{cases} +1/\sqrt{n} & \text{with probability } 1/2, \\ -1/\sqrt{n} & \text{with probability } 1/2. \end{cases}$$

Concentration Inequality: Examples 3

Entries $\Phi_{i,j}$ of Φ are independent samples from following distribution

$$X := \begin{cases} +\sqrt{3}/\sqrt{n} & \text{with probability } 1/6, \\ 0 & \text{with probability } 2/3, \\ -\sqrt{3}/\sqrt{n} & \text{with probability } 1/6. \end{cases}$$

In all the above examples the constant $c(\epsilon)$ can be chosen as $\epsilon^2/4 - \epsilon^3/6$.

Restricted Isometry Property 1

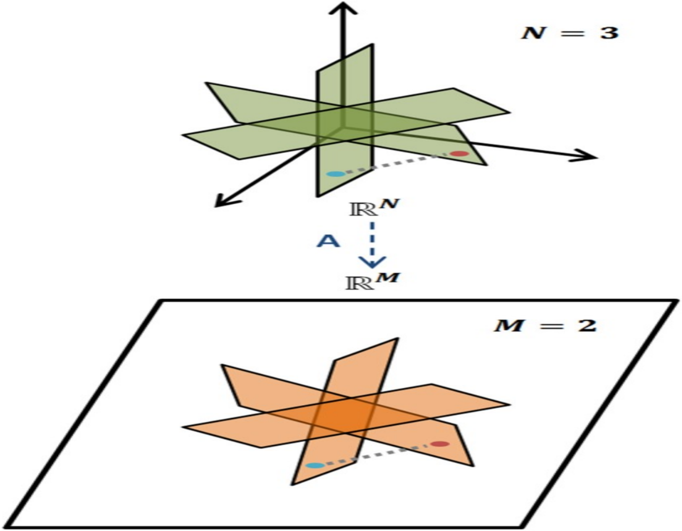

Figure: RIP mapping

Restricted Isometry Property 2

Definition 3
We say that a matrix Φ satisfies Restricted Isometry Property (RIP) of order k and level $\delta \in (0,1)$ (concisely, $(k,\delta) - RIP$) if

$$(1-\delta)\,\|x\|_2^2 \leq \|\Phi(x)\|_2^2 \leq (1+\delta)\,\|x\|_2^2 \quad \forall k-sparse\ x \in \mathbb{R}^N. \quad (2)$$

The restricted isometry constant δ_k is defined as the smallest value of δ for which (2) holds.

JL Lemma implies RIP

Theorem 4 (Theorem 5.2 in (Baraniuk et al., 2008))
Let $n, N \in \mathbb{N}, \delta \in (0,1)$ *and* $\Phi \in \mathbb{R}^{n \times N}$ *be a random matrix which satisfies concentration inequality (see Definition 2), then* $\exists c_1, c_2 > 0$ *depending only on* δ *such that* Φ *satisfies Restricted Isometry Property of order* $k \leq c_1 n / log(N/k)$ *and level* δ *with probability* $\geq 1 - 2e^{-c_2 n}$.

Proof idea of Theorem 4

1. For each index set T with $\#T = k$ there is finite set of points $\mathcal{Q}_T \subseteq X_T$ such that $\|q\|_2 = 1, \forall q \in \mathcal{Q}_T$ and $\forall x \in X_T$ with $\|x\|_2 = 1$ we have

$$min_{q \in \mathcal{Q}_T} \|x - q\|_2 \leq \delta/4. \tag{3}$$

   such that $\#\mathcal{Q}_T \leq (12/\delta)^k$.

2. Using 1. prove

$$(1-\delta)\|x\|_2 \leq \|\Phi(x)\|_2 \leq (1+\delta)\|x\|_2, \quad \forall x \in X_T, \tag{4}$$

   with high probability.

3. Using union bound with 2. prove RIP

$$(1-\delta)\|x\|_2 \leq \|\Phi(x)\|_2 \leq (1+\delta)\|x\|_2 \quad \forall k-\text{sparse } x \in \mathbb{R}^N, \tag{5}$$

   with high probability.

Proof idea of Theorem 4

1. For each index set T with $\#T = k$ there is finite set of points $\mathcal{Q}_T \subseteq X_T$ such that $\|q\|_2 = 1, \forall q \in \mathcal{Q}_T$ and $\forall x \in X_T$ with $\|x\|_2 = 1$ we have

$$min_{q \in \mathcal{Q}_T} \|x - q\|_2 \leq \delta/4. \tag{3}$$

   such that $\#\mathcal{Q}_T \leq (12/\delta)^k$.

2. Using 1. prove

$$(1-\delta)\|x\|_2 \leq \|\Phi(x)\|_2 \leq (1+\delta)\|x\|_2, \quad \forall x \in X_T, \tag{4}$$

   with high probability.

3. Using union bound with 2. prove RIP

$$(1-\delta)\|x\|_2 \leq \|\Phi(x)\|_2 \leq (1+\delta)\|x\|_2 \quad \forall k-\text{sparse } x \in \mathbb{R}^N, \tag{5}$$

   with high probability.

Proof idea of Theorem 4

1. For each index set T with $\#T = k$ there is finite set of points $\mathcal{Q}_T \subseteq X_T$ such that $\|q\|_2 = 1, \forall q \in \mathcal{Q}_T$ and $\forall x \in X_T$ with $\|x\|_2 = 1$ we have

$$min_{q \in \mathcal{Q}_T} \|x - q\|_2 \leq \delta/4. \tag{3}$$

   such that $\#\mathcal{Q}_T \leq (12/\delta)^k$.

2. Using 1. prove

$$(1-\delta)\|x\|_2 \leq \|\Phi(x)\|_2 \leq (1+\delta)\|x\|_2, \quad \forall x \in X_T, \tag{4}$$

   with high probability.

3. Using union bound with 2. prove RIP

$$(1-\delta)\|x\|_2 \leq \|\Phi(x)\|_2 \leq (1+\delta)\|x\|_2 \quad \forall k-\text{sparse } x \in \mathbb{R}^N, \tag{5}$$

   with high probability.

RIP implies JL Lemma (1)

1. Matrices that satisfy RIP are not random!
2. How can we randomize a matrix?

RIP implies JL Lemma (2)

Definition 5
We call $\xi \in \mathbb{R}^N$ *as a Rademacher sequence if it is uniformly distributed on* $\{-1, +1\}^N$*, where* $N \in \mathbb{N}$*.*

For a matrix Φ we might consider random matrix ΦD_ξ, where ξ is a Rademacher sequence and D_ξ is a diagonal matrix with entries from ξ

RIP implies JL Lemma (2)

Definition 5
We call $\xi \in \mathbb{R}^N$ as a Rademacher sequence if it is uniformly distributed on $\{-1, +1\}^N$, where $N \in \mathbb{N}$.

For a matrix Φ we might consider random matrix ΦD_ξ, where ξ is a Rademacher sequence and D_ξ is a diagonal matrix with entries from ξ

RIP implies JL Lemma (3)

Theorem 6 (Theorem 3.1 in (Krahmer and Ward, 2010))
Let $\eta > 0, \epsilon \in (0,1), n, N \in \mathbb{N}$ be given and the set $E \subset \mathbb{R}^N$ has cardinality $p := \#E$. Suppose the matrix $\Phi \in \mathbb{R}^{n \times N}$ satisfies Restricted Isometry Property of order $k \geq 40 log \frac{4p}{\eta}$ and level $\delta \leq \frac{\epsilon}{4}$. Taking $\xi \in \mathbb{R}^N$ as a Rademacher sequence we have

$$(1-\epsilon)\|x\|_2^2 \leq \|\Phi D_\xi x\|_2^2 \leq (1+\epsilon)\|x\|_2^2, \qquad (6)$$

for all $x \in E$, with probability

$$\geq 1 - \eta.$$

Proof idea of Theorem 6 (1)

1 Split matrices by column wise blocks

$$
\begin{aligned}
\|\Phi D_\xi x\|_2^2 &= \|\Phi D_x \xi\|_2^2 \\
&= \sum_{J=1}^{R} \left\| \Phi_{(J)} D_{(x_{(J)})} \xi_{(J)} \right\|_2^2 \\
&+ 2\xi_{(1)}^* D_{x_{(1)}} \Phi_{(1)}^* \Phi_{(-1)} D_{x_{(-1)}} \xi_{(-1)} \\
&+ \sum_{J,L=2,J\neq L}^{R} \left\langle \Phi_{(J)} D_{x_{(J)}} \xi_{(J)}, \Phi_{(L)} D_{x_{(L)}} \xi_{(L)} \right\rangle .
\end{aligned} \tag{7}
$$

Proof idea of Theorem 6 (2)

2 Get bounds for every item in the right hand side of equation (7)

2.1 Deterministic bound on the first item

$$(1-\epsilon/4)\,\|x\|_2^2 \le \sum_{J=1}^{R} \left\|\Phi_{(J)} D_{x_{(J)}} \xi_{(J)}\right\|_2^2 \le (1+\epsilon/4)\,\|x\|_2^2 . \quad (8)$$

2.2 Probabilistic bound on the second item

$$\mathbb{P}(\exists x \in E : |2\xi_{(1)}^* D_{x_{(1)}} \Phi_{(1)}^* \Phi_{(-1)} D_{x_{(-1)}} \xi_{(-1)}| \ge 0.2\epsilon) \le \frac{\eta}{2} \quad (9)$$

2.3 Probabilistic bound on the third item

$$\mathbb{P}(\exists x \in E : |\sum_{J,L=2,J\ne L}^{R} \langle \Phi_{(J)} D_{x_{(J)}} \xi_{(J)}, \Phi_{(L)} D_{x_{(L)}} \xi_{(L)} \rangle| \ge 0.55\epsilon) \le \eta/2 \quad (10)$$

Proof idea of Theorem 6 (2)

2 Get bounds for every item in the right hand side of equation (7)

2.1 Deterministic bound on the first item

$$(1-\epsilon/4)\,\|x\|_2^2 \leq \sum_{J=1}^{R} \left\|\Phi_{(J)} D_{x_{(J)}} \xi_{(J)}\right\|_2^2 \leq (1+\epsilon/4)\,\|x\|_2^2 . \quad (8)$$

2.2 Probabilistic bound on the second item

$$\mathbb{P}(\exists x \in E : |2\xi_{(1)}^* D_{x_{(1)}} \Phi_{(1)}^* \Phi_{(-1)} D_{x_{(-1)}} \xi_{(-1)}| \geq 0.2\epsilon) \leq \frac{\eta}{2} \quad (9)$$

2.3 Probabilistic bound on the third item

$$\mathbb{P}(\exists x \in E : |\sum_{J,L=2, J\neq L}^{R} \langle \Phi_{(J)} D_{x_{(J)}} \xi_{(J)}, \Phi_{(L)} D_{x_{(L)}} \xi_{(L)} \rangle| \geq 0.55\epsilon) \leq \eta/2 \quad (10)$$

Proof idea of Theorem 6 (2)

2 Get bounds for every item in the right hand side of equation (7)

2.1 Deterministic bound on the first item

$$(1-\epsilon/4)\left\|x\right\|_2^2 \leq \sum_{J=1}^{R} \left\|\Phi_{(J)} D_{x_{(J)}} \xi_{(J)}\right\|_2^2 \leq (1+\epsilon/4)\left\|x\right\|_2^2 . \quad (8)$$

2.2 Probabilistic bound on the second item

$$\mathbb{P}(\exists x \in E : |2\xi_{(1)}^{*} D_{x_{(1)}} \Phi_{(1)}^{*} \Phi_{(-1)} D_{x_{(-1)}} \xi_{(-1)}| \geq 0.2\epsilon) \leq \frac{\eta}{2} \quad (9)$$

2.3 Probabilistic bound on the third item

$$\mathbb{P}(\exists x \in E : |\sum_{J,L=2,J\neq L}^{R} \langle \Phi_{(J)} D_{x_{(J)}} \xi_{(J)}, \Phi_{(L)} D_{x_{(L)}} \xi_{(L)} \rangle| \geq 0.55\epsilon) \leq \eta/2$$

(10)

Proof idea of Theorem 6 (2)

2 Get bounds for every item in the right hand side of equation (7)

2.1 Deterministic bound on the first item

$$(1-\epsilon/4)\,\|x\|_2^2 \leq \sum_{J=1}^{R} \left\|\Phi_{(J)} D_{x_{(J)}} \xi_{(J)}\right\|_2^2 \leq (1+\epsilon/4)\,\|x\|_2^2 . \quad (8)$$

2.2 Probabilistic bound on the second item

$$\mathbb{P}(\exists x \in E : |2\xi_{(1)}^* D_{x_{(1)}} \Phi_{(1)}^* \Phi_{(-1)} D_{x_{(-1)}} \xi_{(-1)}| \geq 0.2\epsilon) \leq \frac{\eta}{2} \quad (9)$$

2.3 Probabilistic bound on the third item

$$\mathbb{P}(\exists x \in E : |\sum_{J,L=2,J\neq L}^{R} \langle \Phi_{(J)} D_{x_{(J)}} \xi_{(J)}, \Phi_{(L)} D_{x_{(L)}} \xi_{(L)} \rangle| \geq 0.55\epsilon) \leq \eta/2 \quad (10)$$

Proof idea of Theorem 6 (2)

3 Putting everything together

$$\mathbb{P}(\exists x \in E : |\, \|\Phi D_\xi x\|_2^2 \,/\, \|x\|_2^2 - 1| \geq \epsilon) \leq \eta. \qquad (11)$$

Thank you!

references

Baraniuk, R., Davenport, M., DeVore, R., and Wakin, M. (2008). A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263.

Krahmer, F. and Ward, R. (2010). New and improved Johnson-Lindenstrauss embeddings via the Restricted Isometry Property. *arXiv e-prints*, page arXiv:1009.0744.