# Dimension Reduction

Computational and Metric Geometry

Instructor: Yury Makarychev

## 1 Dimension Reduction

**Theorem 1.1** (Johnson—Lindenstruass Lemma)**.** *Consider a finite metric subspace $X \subset \ell_2^N$. Let $\varepsilon \in (0,1)$, $n = |X|$, and $d > C \ln n/\varepsilon^2$ (where $C$ is a sufficiently large absolute constant). Then there exists an embedding $\varphi$ of $X$ into $\ell_2^d$ s.t.*

$$(1-\varepsilon) \leq \frac{\|\varphi(x) - \varphi(y)\|}{\|x-y\|_2} \leq (1+\varepsilon). \tag{1}$$

*(that is, the embedding $\varphi$ is "almost" isometric). Moreover, we can find such embedding in randomized polynomial time.*

*Proof.* We show that the algorithm presented below finds the desired embedding with probability that tends to 1 as $n$ tends to $\infty$.

---

**Dimension Reduction Algorithm**
**Input:** A metric space $X \subset \ell_2^N$.
**Output:** An embedding $\varphi$ of $X$ into $\ell_2^d$.

1. Choose a random $d \times N$ matrix $\Gamma = (\gamma_{ij})$, whose entries $\gamma_{ij}$ are i.i.d. standard Gaussian random variables, $\gamma_{ij} \sim \mathcal{N}(0,1)$.
2. Define $\varphi(x) = \frac{1}{\sqrt{d}}\Gamma x$ for every $x \in X$.
3. Return embedding $\varphi$.

---

Consider a pair of points $x$ and $y$ in $X$. Our plan is to prove that

$$p_{xy} \equiv \Pr\left(\text{Inequality (1) does not hold for } x \text{ and } y\right) \leq 1/n^4.$$

Once we establish this bound, the theorem will follow since the probability that Inequality (1) does not hold for some pair $x, y \in X$ is at most $\sum_{x,y\in X} p_{xy} \leq n^2 \cdot (1/n^4) = 1/n^2$ by the union bound.

We now prove that $p_{xy} \le 1/n^4$. Denote $z = (x-y)/\|x-y\|_2$. We have,

$$\|\varphi(x)-\varphi(y)\|_2^2 = \frac{\|\Gamma x - \Gamma y\|_2^2}{d} = \frac{\|\Gamma(x-y)\|_2^2}{d} = \frac{\|x-y\|^2\|\Gamma z\|_2^2}{d}$$
$$= \frac{\|x-y\|^2}{d}\sum_{i=1}^{d}\left(\sum_{j=1}^{N}\gamma_{ij}z_j\right)^2 = \frac{\sum_{i=1}^d g_i^2}{d}\|x-y\|^2,$$

where $g_i = \sum_{j=1}^N z_j\gamma_{ij}$. Therefore,

$$\frac{\|\varphi(x)-\varphi(y)\|_2^2}{\|x-y\|^2} = \frac{\sum_{i=1}^d g_i^2}{d}.$$

Note that each $g_i$ is a sum of scaled Gaussian random variables, and hence $g_i$ is a Gaussian random variable. Let us compute the mean and variance of $g_i$.

$$\mathbb{E}g_i = \mathbb{E}\left[\sum_{j=1}^N z_j\gamma_{ij}\right] = \sum_{j=1}^N z_j\mathbb{E}\left[\gamma_{ij}\right] = 0,$$
$$\operatorname{Var}\left[g_i\right] = \operatorname{Var}\left[\sum_{j=1}^N z_j\gamma_{ij}\right] = \sum_{j=1}^N z_j^2\operatorname{Var}\left[\gamma_{ij}\right] = \sum_{j=1}^N z_j^2 = \|z\|_2^2 = 1.$$

That is, $g_1,\dots,g_d$ are i.i.d. random variables distributed as $\mathcal{N}(0,1)$.

It remains to prove the following lemma (note that $1-\varepsilon > (1-\varepsilon)^2$ and $1+\varepsilon < (1+\varepsilon)^2$).

**Lemma 1.2.** *Let $g_1,\dots,g_d$ be i.i.d. standard Gaussian random variables, where $d > C\ln n/\varepsilon^2$. Then*

$$\Pr\left(-\varepsilon d \le \sum_{i=1}^d g_i^2 - d \le \varepsilon d\right) \ge 1 - 1/n^4.$$

**Exercise 1.** *The random variable $\sum_{i=1}^d g_i^2$ has the chi-square distribution with $d$ degrees of freedom, with density $\frac{1}{2^{d/2}\Gamma(d/2)}\, x^{d/2-1}e^{-x/2}$ (where $\Gamma(t)$ is the gamma function). Use this fact to directly estimate the desired probability and prove the lemma.*

*Proof of Lemma 1.2.* Denote $T = \sum_{i=1}^d g_i^2 - d$. Consider the random variable $e^{(1-\alpha^2)T/2}$ (where $\alpha > 0$ is some number). Note that

$$\mathbb{E}\left[e^{(1-\alpha^2)g_i^2/2}\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{(1-\alpha^2)t^2/2}e^{-t^2/2}dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\alpha^2t^2/2}dt$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{(\alpha t)^2/2}\frac{d(\alpha t)}{\alpha} = \frac{1}{\alpha}.$$

Therefore,

$$\mathbb{E}\left[e^{(1-\alpha^2)T/2}\right] = \mathbb{E}\left[e^{\frac{1-\alpha^2}{2}\sum_{i=1}^d g_i^2}\right]\cdot e^{-d(1-\alpha^2)/2} = e^{-d(1-\alpha^2)/2}\prod_{i=1}^d \mathbb{E}\left[e^{(1-\alpha^2)g_i^2/2}\right] = \frac{e^{-d(1-\alpha^2)/2}}{\alpha^d}.$$

Let $\alpha = 1 + \delta$, where $\delta \in (-1/2, 1/2)$ (we will fix $\delta$ later). Then

$$\mathbb{E}\left[e^{(1-\alpha^2)T/2}\right] = \frac{e^{-d(1-\alpha^2)/2}}{\alpha^d} = e^{\delta d + d\delta^2/2 - d\ln(1+\delta)}.$$

By Taylor's theorem, $\ln(1+\delta) = \delta + R_1(\delta)$, where $|R_1(\delta)| \leq \frac{|\delta|^2}{2} \cdot \max_{t\in(-1/2,1/2)|} |(\ln(1+t))''| = \frac{|\delta|^2}{2} \max_{t\in(-1/2,1/2)} \frac{1}{(1+t)^2} = 2\delta^2$ for all $\delta \in (-1/2, 1/2)$. Therefore,

$$\mathbb{E}\left[e^{(1-\alpha^2)T/2}\right] \leq e^{3|\delta|^2 d}.$$

We now use the Chebyshev inequality to bound $\Pr(T > \varepsilon d)$. For $\delta < 0$ and $\alpha < 1$, we have

$$\mathbb{E}\left[e^{(1-\alpha^2)T/2}\right] \geq e^{(1-\alpha^2)\varepsilon d/2} \cdot \Pr(T > \varepsilon d).$$

Therefore,

$$\Pr(T > \varepsilon d) \leq e^{3\delta^2 d} e^{\delta\varepsilon d + \delta^2 \varepsilon d/2} = e^{\varepsilon\delta d(1+\delta/2+3\delta/\varepsilon)}$$

We let $\delta = -\varepsilon/6$ and get $\Pr(T > \varepsilon d) \leq e^{-\varepsilon^2 d/18} < 1/(2n^4)$ if $C > 90$ (recall that $d > \frac{C\ln n}{\varepsilon^2}$). Similarly, we bound $\Pr(T < -\varepsilon d)$. For $\delta > 0$ and $\alpha > 1$, we have

$$\mathbb{E}\left[e^{(1-\alpha^2)T/2}\right] \geq e^{(\alpha^2-1)\varepsilon d/2} \cdot \Pr(T < -\varepsilon d).$$

Therefore,

$$\Pr(T < -\varepsilon d) \leq e^{3\delta^2 d} e^{-\delta\varepsilon d - \delta^2 \varepsilon d/2} \leq e^{-\varepsilon\delta d(1-3\delta/\varepsilon)}.$$

We let $\delta = \varepsilon/6$ and get $\Pr(T < -\varepsilon d) \leq e^{-\varepsilon^2 d/12} < 1/(2n^4)$ if $C > 60$. We conclude that $\Pr(|T| > \varepsilon d) < 1/n^4$ if $C > 90$. □

□

**Remark 1.3.** The algorithm we presented in this note runs in polynomial time but is relatively slow. In fact, the embedding $\varphi$ can be computed very efficiently using the *Fast Johnson–Lindenstrauss Transform*, which was introduced recently by Ailon and Chazelle. For more information, see *N. Ailon and B. Chazelle. Faster dimension reduction. Commun. ACM 53(2): 97-104 (2010)* and *N. Ailon, E. Liberty. Almost optimal unrestricted fast Johnson-Lindenstrauss transform. CoRR, abs/1005.5513.*