1 The curse (and blessing) of dimensionality

The “curse” of dimensionality refers to the fact that many algorithmic approaches to problems in $\mathbb{R}^k$ become exponentially more difficult as $k$ grows. This is essentially due to volume growth: The $k$-dimensional Euclidean ball of radius $r$ grows exponentially in $k$: $\text{vol}_k(B(0,2r)) = 2^k \text{vol}_k(B(0,r))$.

On the other hand, in high dimensions, the concentration of measure phenomenon can come to our aid: Sufficiently smooth functionals on $\mathbb{R}^k$ are tightly concentrated around their expected value. A prototypical example is the Johnson-Lindenstrauss dimension reduction lemma.

1.1 Dimension reduction

**Lemma 1.1** (Johnson-Lindenstrauss). For every $n \geq 1$ and every $n$-point subset $X \subseteq \mathbb{R}^n$, the following holds. For every $\epsilon > 0$, there is a linear map $A : \mathbb{R}^n \to \mathbb{R}^k$ such that

$$(1 - \epsilon)^2 \|x - y\|_2^2 \leq \|A(x) - A(y)\|_2^2 \leq (1 + \epsilon)^2 \|x - y\|_2^2 \quad \forall x, y \in X,$$

and $k \leq \frac{24 \ln n}{\epsilon^2}$.

**Proof.** We may assume that $0 < \epsilon < 1/2$. We will define a random linear map $A : \mathbb{R}^n \to \mathbb{R}^k$ and then argue that $A$ satisfies (1.1) with high probability. Let $\{X_{ij} : i = 1, \ldots, k, j = 1, \ldots, n\}$ be a family of i.i.d. $N(0,1)$ random variables, and define an $m \times n$ matrix $A$ by $A_{ij} = \frac{1}{\sqrt{k}} X_{ij}$.

**Claim 1.2.** For every $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$, we have

$$\Pr \left[ \|Au\|_2^2 \notin [1 - \epsilon, 1 + \epsilon] \right] \leq 2e^{-ck/8}.$$

Let us finish the proof of Lemma 1.1 and then prove the claim. By setting $u = \frac{x-y}{\|x-y\|_2}$ for every $x, y \in X$ and taking a union bound over these $n^2$ different pairs, we have

$$\Pr \left( (1 - \epsilon)\|x - y\|_2^2 \leq \|A(x - y)\|_2^2 \leq (1 + \epsilon)\|x - y\|_2^2 \right) \geq 1 - 2n^2 e^{-ck/8}.$$ 

Taking $k = \frac{24 \ln n}{\epsilon^2}$ the latter probability is at least $1 - \frac{2}{n}$, showing that the desired map $A$ exists (and can be found by a randomized algorithm).

**Proof of Claim 1.2.** Note first that

$$\|Au\|_2^2 = \frac{1}{k} \sum_{i=1}^k \left( \sum_{j=1}^n u_j X_{ij}^{(j)} \right)^2.$$

By the 2-stability property of independent $N(0,1)$ random variables, we have $\sum_{j=1}^n u_j X_{ij}^{(j)}$ again has distribution $N(0,1)$. Thus we have

$$\|Au\|_2^2 = \frac{1}{k} \sum_{i=1}^k Y_i^2,$$

where $\{Y_i\}$ are i.i.d. $N(0,1)$.
Since $\|Au\|_2^2$ is a sum of independent random variables, it makes sense to use the method of Laplace transforms. So far we have only done this for bounded random variables, but since $N(0, 1)$ variables have a quickly diminishing tail, we can hope that the method will work here as well.

For some parameter $\lambda > 0$, we have

$$
\mathbb{P} (\|Au\|_2^2 > 1 + \varepsilon) = \mathbb{P} \left( \sum_{i=1}^{k} (Y_i^2 - 1) > \varepsilon k \right) = \mathbb{P} \left( e^{\lambda \sum_{i=1}^{k} (Y_i^2 - 1)} > e^{\varepsilon k} \right) 
= \mathbb{E} \left[ e^{\lambda \sum_{i=1}^{k} (Y_i^2 - 1)} \right] 
\leq \frac{e^{\varepsilon k}}{e^{\varepsilon k}} \prod_{i=1}^{n} \mathbb{E} \left[ e^{\lambda (Y_i^2 - 1)} \right].
$$

Thus our remaining goal is to bound $\mathbb{E} \left[ e^{\lambda (Y^2 - 1)} \right]$ when $Y$ is $N(0, 1)$. First, observe that

$$
\mathbb{E} \left[ e^{\lambda Y^2} \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} e^{\lambda t^2} dt = \frac{1}{\sqrt{1 - 2\lambda}}
$$

for $0 < \lambda < 1/2$. (Clearly the integral is divergent for $\lambda \geq 1/2$.) Thus in this range of $\lambda$,

$$
\mathbb{E} \left[ e^{\lambda (Y^2 - 1)} \right] = \frac{e^{-\lambda}}{\sqrt{1 - 2\lambda}}.
$$

Let us finally calculate (using the Taylor expansion $\log(1 - x) = -\sum_{k \geq 1} \frac{x^k}{k}$),

$$
\left| \log \mathbb{E} \left[ e^{\lambda (Y^2 - 1)} \right] \right| = \left| -\lambda - \frac{1}{2} \log(1 - 2\lambda) \right|
\leq 2\lambda^2(1 + (2\lambda) + (2\lambda)^2 + \cdots)
= \frac{2\lambda^2}{1 - 2\lambda}.
$$

Therefore:

$$
\mathbb{P} (\|Au\|_2^2 > 1 + \varepsilon) \leq e^{2\lambda^2/(1 - 2\lambda) - \varepsilon k}.
$$

Choosing $\lambda = \varepsilon/4$ and using $0 < \varepsilon < 1/2$, we conclude that $\mathbb{P} (\|Au\|_2^2 > 1 + \varepsilon) \leq e^{-\varepsilon k/8}$, completing the proof. A similar argument shows that $\mathbb{P} (\|Au\|_2^2 < 1 - \varepsilon) \leq e^{-\varepsilon k/8}$. □

**Remark 1.3.** A simple volume bound shows that the $\Theta(\log n)$ dependence in the dimension is necessary. A linear algebraic argument of Alon shows that the dimension must be at least $\Omega(\log n / \log(1/\varepsilon))$. Very recently (FOCS 2017), Larsen and Nelson established that there is a lower bound of $\Omega(\log n / \varepsilon^2)$, showing that Lemma 1.1 is tight up to the constant factor.

**Remark 1.4.** Lemma 1.1 actually works for any independent family $X_{ij}$ where each random variable satisfies a sub-Gaussian tail bound: $\mathbb{E}[e^{\alpha X}] \leq e^{C\alpha^2}$ for some constant $C \geq 1$. For instance, if $X$ is a uniform $\pm 1$ random variable, then $\mathbb{E}[e^{\alpha X}] = \frac{1}{2}(e^\alpha + e^{-\alpha}) \leq e^{\alpha^2}/2$ (recall that to prove this, you should Taylor expand both sides).
The proof uses a clever trick. Note that if $Z$ is a $N(0,1)$ random variable (independent of $X$) then $\mathbb{E}[e^{\alpha Z}] = e^{\alpha^2/2}$ for any $\alpha > 0$. Now write

$$\mathbb{E}[e^{\lambda X^2}] = \mathbb{E}[e^{(\sqrt{\lambda}X)^2/2}] = \mathbb{E}[e^{\sqrt{\lambda}ZX}] \leq \mathbb{E}[e^{2C\lambda Z^2}] = \frac{1}{\sqrt{1 - 2\lambda C}},$$

for $0 < \lambda < \frac{1}{2C}$, where the last equality is exactly the equality we proved in (1.2). Given this bound, we can finish the proof just as in Claim 1.2.