Notes for Random Matrix Theory

Vishwak Srinivasan

Contents

1 Introduction

A useful fact is the following: *tall matrices are approximate isometries.* Let's parse this statement.

- Tall matrices are those matrices $A \in \mathbb{R}^{N \times n}$ where $N \gg n$.
- Approximate isometries: consider the vector space $\mathbb{R}^n$ and $\mathbb{R}^N$. A tall matrix transforms a vector $x \in \mathbb{R}^n$ to $Ax \in \mathbb{R}^N$. Mathematically:

  $$(1-\delta)K\|x\|_2 \leq \|Ax\|_2 \leq (1+\delta)K\|x\|_2$$

  where K is a normalization factor and $\delta \ll 1$. This looks like a version of Johnson-Lindenstrauss.

Now divide by $\|x\|_2$ to get:

$$(1-\delta)K \leq \frac{\|Ax\|_2}{\|x\|_2} \leq (1+\delta)K \Rightarrow (1-\delta)K \leq \sigma_{\min}(A) \leq \sigma_{\max}(A) \leq (1+\delta)K$$

and this tells us that the range of singular values is small. Furthermore, the condition number be bounded as: $\kappa(A) \leq \frac{1+\delta}{1-\delta} \approx 1$ for the specified bound on δ. Therefore, tall matrices are well conditioned *always.*

1.1 Some miscellaneous results

ϵ-nets are a neat way of computing quantities that can be expressed over balls. An ϵ-net of a set S is a set of points $\mathcal{N}_\epsilon(S)$ that approximates a point in the original set. That is to say, for every $x \in S$, there exists $y \in \mathcal{N}_\epsilon(S)$ such that $\|y - x\| \leq \epsilon$. As you can see, it is define w.r.t a norm. The cardinality of the $\mathcal{N}_\epsilon(S)$ is called the ϵ-covering number of S (w.r.t. a norm).

Now, let's look at an application of nets in computing the ℓ_2-norm of a vector.

Lemma 1.1. *Let $x \in \mathbb{R}^p$. Then:*

$$\|x\|_2 \leq \frac{1}{1-\epsilon} \sup_{v \in \mathcal{N}_\epsilon(S^{p-1})} |v^T x|$$

Proof. Let v be the unit vector that results in $|v^T x| = \|x\|_2$ (this is just unit vector in the direction of x). Now, let $w \in \mathcal{N}_\epsilon(S^{p-1})$ be the closest element to v. By Cauchy-Schwarz:

$$|(v-w)^T x| \leq \|v - w\|_2 \|x\|_2 \leq \epsilon \|x\|_2$$

Hence, by triangle inequality:

$$|w^T x| \geq |v^T x| - |(w-v)^T x| \geq \|x\|_2 - \epsilon\|x\|_2 \Rightarrow \|x\|_2 \leq \frac{1}{1-\epsilon}|w^T x| \leq \frac{1}{1-\epsilon} \sup_{w \in \mathcal{N}_\epsilon(S^{p-1})} |w^T x|$$

□

Remark. An important takeaway is that an optimization problem over an uncountable set, has now be reduced to a countable set, at the cost of some sub-optimality. The ϵ-covering number of S^{p-1} is $\left(1 + \frac{2}{\epsilon}\right)^p$.

Back to random matrix theory, let's use an idea from the above lemma to compute the spectral norm of a matrix A.

Lemma 1.2. *Let $A \in \mathbb{R}^{N \times n}$. Then:*

$$\|A\|_2 \leq \frac{1}{1-\epsilon} \sup_{v \in \mathcal{N}_\epsilon(S^{n-1})} \|Av\|_2$$

Proof. The proof is the same from earlier. Let v be unit vector leading to $\|Av\|_2 = \|A\|_2$. Choose $w \in \mathcal{N}_\epsilon(S^{n-1})$ closest to v. Using the fact that $\|Ax\|_2 \leq \|A\|_2\|x\|_2$:

$$\|A(w-v)\| \leq \|A\|_2\|w-v\|_2 \leq \|A\|_2\epsilon$$

By triangle inequality:

$$\|Aw\|_2 \geq \|Av\|_2 - \|A(w-v)\|_2 \geq \|A\|_2 - \epsilon\|A\|_2 \Rightarrow \|Aw\|_2 \leq \frac{1}{1-\epsilon}\|A\|_2 \leq \frac{1}{1-\epsilon}\sup_{w\in\mathcal{N}_\epsilon(S^{n-1})}\|Aw\|_2$$

$\square$