

Operator Norm of Random Matrices

Operator norm $\|M\|_{\text{op}} = \sup_{x \in \mathbb{C}^d, |x|=1} |Mx|$

Goal: Good upper tail bounds $\mathbb{P}(\|M\|_{\text{op}} \geq \lambda) \leq \dots$

Outline

1. ϵ -net argument
2. Concentration of measure
3. Moment method

ϵ -net Argument

Define $S = \{x \in \mathbb{C}^d : |x| = 1\}$.

Note $\mathbb{P}(\|M\|_{\text{op}} > \lambda) = \mathbb{P}(\sup_{x \in S} |Mx| > \lambda)$.

Assumption A1: Let M be an $n \times n$ matrix. Suppose coefficients $\{\xi_{ij}\}$ of M are independent, have mean 0, and are uniformly bounded in magnitude by 1.

Lemma 1: Assume A1. Choose $x \in S$. There exist absolute constants $C, c > 0$ s.t. for A sufficiently large,

$$\mathbb{P}(|Mx| \geq A\sqrt{n}) \leq Ce^{-cAn}.$$

Note x is a fixed number, not a random variable.

\hookrightarrow Can't immediately apply to $x_0 = \operatorname{argmax}_S |Mx|$.

Def: A maximal ϵ -net of S is a set of points in S separated from each other by a distance of at least ϵ , which is maximal w.r.t. set inclusion.

Lemma 2: Let \mathcal{Z} be a maximal ϵ -net, with $0 < \epsilon < 1$. For some absolute constant p , $|\mathcal{Z}| \leq \left(\frac{p}{\epsilon}\right)^n$.

Lemma 3: Assume A1. Let \mathcal{Z} be a maximal ε -net, with $0 < \varepsilon < 1$. For $M \in \mathbb{C}^{n \times n}$, and $\lambda > 0$,

$$\mathbb{P}(\|M\|_{\text{op}} > \lambda) \leq \mathbb{P}\left(\max_{z \in \mathcal{Z}} |My| > \frac{\lambda}{1-\varepsilon}\right).$$

Cor: Assume A1 Then there exist absolute constants $C, c > 0$ s.t.

$$\mathbb{P}(\|M\|_{\text{op}} \geq A\sqrt{n}) \leq C e^{-cAn},$$

for all $A \geq C$.

Proof: By lemma 3 with $\varepsilon = \frac{1}{2}$,

$$\begin{aligned} \mathbb{P}(\|M\|_{\text{op}} > \lambda) &\leq \mathbb{P}\left(\max_{z \in \mathcal{Z}} |My| > \frac{\lambda}{2}\right) \\ &= \sum_{y \in \mathcal{Z}} \mathbb{P}\left(|My| \geq \frac{\lambda}{2}\right) \\ &\stackrel{\text{Lemma 1}}{\leq} \sum_{y \in \mathcal{Z}} C e^{-cAn} \\ &\stackrel{\text{Lemma 2}}{\leq} (2p)^n C e^{-cAn}. \end{aligned}$$

If A is sufficiently large, we can choose c large enough so that $(2p)^n$ is absorbed into the exponential. □

Concentration of Measure

Thm (Talagrand): Let $\{X_i\}_{i=1}^n$ be independent, complex RVs, bounded above by 1. Let $F: \mathbb{C}^n \rightarrow \mathbb{R}$ 1-Lipschitz convex function. Then \exists absolute constants $c, C > 0$ s.t. $\forall \lambda > 0$,

$$\mathbb{P}\left(|F(X) - \text{med}| \geq \lambda\right) \leq C e^{-c\lambda^2}.$$

Cor: Assume A1. There exist absolute constants $c, C > 0$ s.t.
 $\forall \lambda > 0,$

$$\mathbb{P}\left(\left| \|M\|_{op} - \text{med} \right| \geq \lambda\right) \leq Ce^{-c\lambda^2}.$$

"med" is a median value for M . The same result holds if we replace med with $\mathbb{E}(\|M\|_{op})$.

Proof: Define $F: \mathbb{C}^n \rightarrow \mathbb{R}$ by $F(\{\xi_{ij}\}_{i,j=1}^n) = \|M\|_{op}$.

Since the operator norm is convex, F is convex.

If $\|M\|$ is the Frobenius norm, then $\|M\|_{op} \leq \|M\|_F$, so F is Lipschitz with constant 1.

The claim follows from Talagrand's. □

The Moment Method

Assume M symmetric ($M = M^T$) or Hermitian ($M = \overline{M^T}$).

Then M has n real eigenvalues, $\{\lambda_1, \dots, \lambda_n\}$.

Notice for k even,

$$\|M\|_{op}^k = \left(\max_{1 \leq i \leq n} |\lambda_i|\right)^k$$

$$= \max_{1 \leq i \leq n} \lambda_i^k$$

$$\leq \sum_{i=1}^n \lambda_i^k$$

$$= \text{tr}(M^k) \quad \textcircled{1}$$

$$\leq \sum_{i=1}^n \max_{1 \leq j \leq n} \lambda_j^k$$

$$= \sum_{i=1}^n \|M^k\|_{op}$$

$$= n \|M^k\|_{op}. \quad \textcircled{2}$$

In sum, $\|M\|_{op}^k \leq \text{tr}(M^k) \leq n \|M\|_{op}^k$

$k = 2$:

Assume A1; then $\mathbb{E}[|\xi_{ij}|^2] = 1$.

By the weak law of large numbers,

$$\text{tr}(M^2) = \sum_{i,j=1}^n |\xi_{ij}|^2 = (1 + o(1)) n^2 \text{ a.s.}$$

Thus, $\|M\|_{op} \leq (1 + o(1)) n \leq \sqrt{n} \|M\|_{op}$, so

$$\Rightarrow (1 + o(1)) \sqrt{n} \leq \|M\|_{op} \leq (1 + o(1)) n$$

$k = 4$:

Assumption A2: M is symmetric or Hermitian, and for each ξ_{ij} ,

1. $\mathbb{E}[\xi_{ij}] = 0$, $\mathbb{E}[\xi_{ij}^2] = 1$.

2. $|\xi_{ij}| \leq K$.

Assume A2.

$$\text{tr}(M^4) = \sum_{1 \leq i_1, i_2, i_3, i_4 \leq n} \xi_{i_1 i_2} \xi_{i_2 i_3} \xi_{i_3 i_4} \xi_{i_4 i_1}$$

$$\Rightarrow \mathbb{E}[\text{tr}(M^4)] = \sum_{1 \leq i_1, i_2, i_3, i_4 \leq n} \mathbb{E}[\xi_{i_1 i_2} \xi_{i_2 i_3} \xi_{i_3 i_4} \xi_{i_4 i_1}]$$

Since entries are independent with mean 0, we only need to consider the cases:

1. $i_1 = i_3$, $i_2 \neq i_4$.

2. $i_1 = i_3$, $i_2 = i_4$.

3. $i_1 = i_2 = i_3$, $i_4 \neq i_1$.

4. $i_1 = i_2 = i_3 = i_4$.

- ① $\Rightarrow \mathbb{E}[\] = 1, O(n^3)$ terms
- ② $\Rightarrow \mathbb{E}[\] \stackrel{\text{symmetry}}{=} \mathbb{E}[\xi_{ij}^2] = K^2, O(n^2)$ terms
- ③ $\Rightarrow \mathbb{E}[\] = 1, O(n^2)$ terms
- ④ $\Rightarrow \mathbb{E}[\] = K^2, O(n)$ terms

So $\mathbb{E}[\text{tr}(M^k)] \leq O(n^3) + O(n^2 K^2)$.

If $K = O(\sqrt{n})$, then $\mathbb{E}[\text{tr}(M^k)] \leq Cn^3$ for some C .

By Markov's inequality, for $\varepsilon > 0$,

$$\mathbb{P}(\text{tr}(M^k) > \frac{Cn^3}{\varepsilon}) \leq \frac{\mathbb{E}[\text{tr}(M^k)]}{Cn^3} \varepsilon \leq \varepsilon.$$

$$\Rightarrow \mathbb{P}(\text{tr}(M^k) \leq \frac{Cn^3}{\varepsilon}) \geq 1 - \varepsilon.$$

$$\Rightarrow \mathbb{P}(\|M\|_{\text{op}} \leq \frac{Cn^{3/4}}{\varepsilon}) \geq 1 - \varepsilon, \text{ i.e. } \|M\|_{\text{op}} \leq O_{\varepsilon}(n^{3/4}).$$

k even (general)

Prop: Assume A2. Then

$$\|M\|_{\text{op}} = O\left(\sqrt{n} \ln(n) \max\left(1, \frac{k}{\sqrt{n}}\right)\right)$$

with probability $1 - O(n^{-c})$, i.e. "high probability."