EE378B Inference, Estimation, and Information Processing

A better bound on the norm of random matrices

Andrea Montanari *Lecture 4-5 - Due on 2/1/2021*

Homework should be submitted via Gradescope, by Monday afternoon: the code will be communicated by an announcement on Canvas. This homework requires some

For getting credit for the class, you are required to present solutions of some of these homeworks during the first 15 minutes of class starting on 1/20. Please, sign up for (at least) one slot, and be sure that your explanation lasts 15 minutes (or less). For these presentations, you are free to choose whatever format you prefer (slides, typed notes, handwriting, ...).

This week, the presentations will be:

- Monday 2/1: Questions (a), (b), (c)
- Wednesday 2/3: Questions (d), (e), (f), (g).

Problem

This exercise aims at developing a more refined version of the ϵ-net method to bound the operator norm of random matrices.

We say that a centered random variable X (with $\mathbb{E}X = 0$) is b-sub-exponential if, for all λ with $|\lambda| \le 1/b$,

$$\mathbb{E}\{e^{\lambda X}\} \le e^{\lambda^2 b^2/2}\,. \tag{1}$$

There are other equivalent ways to define sub-exponential random variables. It is useful to recall following Bernstein inequality for sub-exponential random variables.

Theorem 1 (Bernstein's inequality)**.** *Let $(X_i)_{i\le N}$ be a sequence of independent centered random variables, where X_i is b_i-sub-exponential, and define $\boldsymbol{b} = (b_1, \dots, b_N)$. Then there exists a universal constant c_0 such that, for all $t \ge 0$,*

$$\mathbb{P}\left\{\Big|\sum_{i=1}^N X_i\Big| \ge t\right\} \le 2\,\exp\left\{-c_0 \min\Big(\frac{t}{\|\boldsymbol{b}\|_\infty}, \frac{t^2}{\|\boldsymbol{b}\|_2^2}\Big)\right\}\,. \tag{2}$$

Let $\boldsymbol{X} \in \mathbb{R}^{d_1\times d_2}$ be a random matrix with independent centered b-sub-exponential entries. We are interested in bounding the operator norm $\|\boldsymbol{X}\|_{\rm op}$. Using the naive ϵ-net method may not give a desirable bound.

To give a better bound on the operator norm $\|\boldsymbol{X}\|_{\rm op}$, we will construct a special ϵ-net as follows. For L an integer, define the set

$$S_L = \Big\{0, 1, \frac{1}{2}, \frac{1}{2^2}, \dots, \frac{1}{2^L}\Big\}\,. \tag{3}$$

We then define

$$N^n(L) \equiv \Big\{\boldsymbol{x} \in \mathbb{R}^n:\ \|\boldsymbol{x}\|_2 \le 1,\ \ x_i^2 \in S_L\Big\}\,. \tag{4}$$

We further define $\pi_{<\ell} : N^n(L) \to N^n(L)$ and $\pi_{=\ell} : N^n(L) \to N^n(L)$ by

$$\pi_{<\ell}(\boldsymbol{x})_i = x_i\, \mathbf{1}_{x_i^2 > 2^{-\ell}}\,, \tag{5}$$

$$\pi_{=\ell}(\boldsymbol{x})_i = x_i\, \mathbf{1}_{x_i^2 = 2^{-\ell}}\,. \tag{6}$$

We also let $N^n_{=\ell} = \pi_{=\ell}(N^n(L))$, $N^n_{<\ell} = \pi_{<\ell}(N^n(L))$.

(*a*) Give an example of a random variable that is sub-exponential but not sub-Gaussian (and prove your claim).

(*b*) Prove that, if $L = \log_2 n + c_0$ for c_0 a suitable constant, $N^n(L)$ is an ϵ_0-net of the unit ball $\mathsf{B}_2^n(0,1)$ for some $\epsilon_0 < 1/2$. As a consequence, for suitably chosen L,

$$\mathbb{P}\Big(\|\boldsymbol{X}\|_{\rm op} \geq t\Big) \leq \mathbb{P}\Big(\max_{\boldsymbol{u}\in N^{d_1}(L)} \max_{\boldsymbol{v}\in N^{d_2}(L)} |\langle \boldsymbol{u}, \boldsymbol{X}\boldsymbol{v}\rangle| \geq C(\epsilon_0)\, t\Big)\,. \tag{7}$$

(*c*) Prove that, for c_1 a suitable constant[1] (recall that $a \vee b \equiv \max(a,b)$)

$$|N^n_{=\ell}| \vee |N^n_{<\ell}| \leq \Big(\frac{c_1 n}{2^\ell}\Big)^{2^\ell}\,. \tag{8}$$

(*d*) Prove that, for any $t \geq 0$, $\boldsymbol{u} \in \mathsf{B}_2^{d_1}(0,1)$ and $\boldsymbol{v} \in \mathsf{B}_2^{d_2}(0,1)$,

$$\mathbb{P}\Big(\big|\langle \boldsymbol{u}, \boldsymbol{X}\boldsymbol{v}\rangle\big| \geq t\Big) \leq 2\,\exp\Big\{-c_0 \min\Big(\frac{t}{b\cdot\|\boldsymbol{u}\|_\infty\|\boldsymbol{v}\|_\infty}, \frac{t^2}{b^2}\Big)\Big\}\,. \tag{9}$$

(*e*) Show that, for any matrix $\boldsymbol{M} \in \mathbb{R}^{d_1\times d_2}$ and $\boldsymbol{u} \in N^{d_1}(L)$, $\boldsymbol{v} \in N^{d_2}(L)$

$$\langle \boldsymbol{u}, \boldsymbol{M}\boldsymbol{v}\rangle = \sum_{\ell=0}^{L}\langle \pi_{=\ell}(\boldsymbol{u}), \boldsymbol{M}\pi_{=\ell}(\boldsymbol{v})\rangle + \sum_{\ell=0}^{L}\langle \pi_{=\ell}(\boldsymbol{u}), \boldsymbol{M}\pi_{<\ell}(\boldsymbol{v})\rangle + \sum_{\ell=0}^{L}\langle \pi_{<\ell}(\boldsymbol{u}), \boldsymbol{M}\pi_{=\ell}(\boldsymbol{v})\rangle\,. \tag{10}$$

(*f*) Use the above results to upper bound the following probabilities, for $\ell \in \{0,\dots,L\}$:

$$\mathbb{P}\Big(\max_{\boldsymbol{u}\in N^{d_1}(L),\boldsymbol{v}\in N^{d_2}(L)} \langle \pi_{=\ell}(\boldsymbol{u}), \boldsymbol{X}\pi_{=\ell}(\boldsymbol{v})\rangle \geq t_\ell\Big)\,, \tag{11}$$

$$\mathbb{P}\Big(\max_{\boldsymbol{u}\in N^{d_1}(L),\boldsymbol{v}\in N^{d_2}(L)} \langle \pi_{<\ell}(\boldsymbol{u}), \boldsymbol{X}\pi_{=\ell}(\boldsymbol{v})\rangle \geq t_\ell\Big)\,, \tag{12}$$

$$\mathbb{P}\Big(\max_{\boldsymbol{u}\in N^{d_1}(L),\boldsymbol{v}\in N^{d_2}(L)} \langle \pi_{=\ell}(\boldsymbol{u}), \boldsymbol{X}\pi_{<\ell}(\boldsymbol{v})\rangle \geq t_\ell\Big). \tag{13}$$

(*g*) Combine the above elements to obtain a bound on $\|\boldsymbol{X}\|_{\rm op}$.

[1] It might be useful to remember that $\binom{n}{k} \leq (ne/k)^k$.