HW II

EE 546: Mathematics of High-dimensional Data

University of Southern California

Assigned on: October 2, 2018 Due date: beginning of class on Monday October 15, 2018

The purpose of this homework is two fold:

(1) Give you some practice with concentration inequalities.

(2) Help you refresh some results on operator norms.

1- **Concentration.** Let $X_1, X_2, \ldots, X_m$ be i.i.d. Gaussian random variables with $\mathcal{N}(0,1)$ distribution. Also let $\boldsymbol{a}_r$ be i.i.d. random Gaussian vectors distributed as $\mathcal{N}(\mathbf{0}, \boldsymbol{I})$ and independent of $X_1, X_2, \ldots, X_m$. Let $\boldsymbol{y}$ be a fixed vector. We are interested in understanding the random variable

$$Z = \frac{1}{m}\sum_{r=1}^{m} X_r \big|X_r\big| \textbf{sign}\left(X_r + \mathbf{a}_r^T \boldsymbol{y}\right)$$

(a) What is the expected value of Z?
Hint: You are allowed to use the following identity without proof. For X and Y independent $\mathcal{N}(0,1)$ random variables we have

$$\mathbb{E}\big[X^2 \text{sgn}(X)\text{sgn}(\alpha X + \beta Y)\big] = \frac{2}{\pi}\tan^{-1}\left(\frac{\alpha}{\beta}\right) + \frac{2}{\pi}\frac{\alpha\beta}{\alpha^2+\beta^2}.$$

(b) How well is the random variable Z concentrated? That is, can you bound $\mathbb{P}\{\big|Z - \mathbb{E}[Z]\big| \geq t\}$?
Hint: What kind of a random variable is $X_r \big|X_r\big| \textbf{sign}\left(X_r + \mathbf{a}_r^T \boldsymbol{y}\right)$?

2- **Operator norms.** Remember that the ℓ_2 operator norm of $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is defined as

$$\|\boldsymbol{A}\| := \max\left\{ \|\boldsymbol{A}\boldsymbol{x}\|_{\ell_2} : \ \|\boldsymbol{x}\|_{\ell_2} = 1\right\}.$$

(i) Prove that the operator norm can alternatively be written in the following two forms

$$\begin{aligned} \|\boldsymbol{A}\| =& \max\left\{\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{y}\rangle : \quad \boldsymbol{x} \in \mathbb{R}^m,\ \boldsymbol{y} \in \mathbb{R}^n,\ \text{and}\ \|\boldsymbol{x}\|_{\ell_2} = \|\boldsymbol{y}\|_{\ell_2} = 1\right\}, \\ =& \sigma_1(\boldsymbol{A}), \end{aligned}$$

where $\sigma_1(\boldsymbol{A}) \geq \sigma_2(\boldsymbol{A}) \geq \ldots \geq \sigma_n(\boldsymbol{A})$ are the singular values of $\boldsymbol{A}$.

(ii) Prove that

$$\sum_{s=1}^{r} \sigma_s(\boldsymbol{A}) = \max\left\{\text{trace}(\boldsymbol{U}^T\boldsymbol{A}\boldsymbol{V}) : \quad \boldsymbol{U} \in \mathbb{R}^{m\times r},\ \boldsymbol{V} \in \mathbb{R}^{n \times r},\ \text{and}\ \boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}_r\right\}.$$

where $\sigma_1(\boldsymbol{A}) \geq \sigma_2(\boldsymbol{A}) \geq \ldots \geq \sigma_n(\boldsymbol{A})$ are the singular values of $\boldsymbol{A}$ and $\boldsymbol{I}_r$ is the $r \times r$ identity matrix.

3- Prove the following upper and lower bounds

$$\|\boldsymbol{A}\| \leq \sqrt{m} \max_{i\in\{1,2,\dots,m\}} \left(\sum_{j=1}^{n} A_{ij}^2\right)^{1/2},$$
$$\|\boldsymbol{A}\| \geq \frac{1}{\sqrt{mn}} \sum_{i=1}^{m} \left|\sum_{j=1}^{n} A_{ij}\right|.$$

Give examples of matrices such that these bounds are satisfied with equality.

4- Prove that

$$\|\boldsymbol{A}\|_F = \left(\sum_{i=1}^{\min(m,n)} \sigma_i^2(\boldsymbol{A})\right)^{1/2},$$

where $\|\boldsymbol{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}^2}$ is the Frobenius norm of the matrix $\boldsymbol{A}$. Deduce that

$$\|\boldsymbol{A}\| \leq \|\boldsymbol{A}\|_F \leq \sqrt{\operatorname{rank}(\boldsymbol{A})}\|\boldsymbol{A}\|.$$