EE 546 HW2

1. Solve: (a) Since $|X_r| = \begin{cases} X_r & \text{when } X_r \geq 0 \\ -X_r & \text{when } X_r < 0 \end{cases}$

$\therefore |X_r| = X_r \cdot \text{sign}(X_r)$

we can know $Z_r = X_r \cdot |X_r| \, \text{sign}(X_r + a_r^T y)$

$= X_r^2 \, \text{sign}(X_r) \cdot \text{sign}(X_r + a_r^T y)$

$\because Z = \frac{1}{m} \sum_{r=1}^{m} Z_r$

$\therefore Z = \frac{1}{m} \sum_{r=1}^{m} [X_r^2 \cdot \text{sign}(X_r) \cdot \text{sign}(X_r + a_r^T y)]$

Since random variables $X_1, X_2, \dots, X_m$ are i.i.d

$\therefore E(Z) = \frac{1}{m} E(\sum_{r=1}^{m} Z_r) = \frac{1}{m} \sum_{r=1}^{m} E(Z_r)$

$= \frac{1}{m} \sum_{r=1}^{m} E[X_r^2 \cdot \text{sign}(X_r) \cdot \text{sign}(X_r + a_r^T y)]$

We know that a_r are i.i.d Gaussian random vectors distributed as $N(0, I)$

$\therefore a_r^T y \sim N(0, y^T y)$

$\Rightarrow \frac{a_r^T y}{\sqrt{y^T y}} \sim N(0,1) \Rightarrow E(Z_r) = E[X_r^2 \cdot \text{sign}(X_r) \cdot \text{sign}(X_r + \sqrt{y^T y} \cdot \frac{a_r^T y}{\sqrt{y^T y}})]$

Due to $E[X^2 \text{sign}(X) \cdot \text{sign}(\alpha X + \beta Y)] = \frac{2}{\pi} \tan^{-1}(\frac{\alpha}{\beta}) + \frac{2}{\pi} \frac{\alpha\beta}{\alpha^2 + \beta^2}$ for X and Y are independent $N(0,1)$

$\therefore E(Z_r) = \frac{2}{\pi} \tan^{-1}(\frac{1}{\sqrt{y^T y}}) + \frac{2}{\pi} \frac{\sqrt{y^T y}}{1 + y^T y}$, which is a fixed constant since y is a fixed vector

Therefore $E(Z) = \frac{1}{m} \sum_{r=1}^{m} E(Z_r) = E(Z_r) = \frac{2}{\pi} \tan^{-1}(\frac{1}{\sqrt{y^T y}}) + \frac{2}{\pi} \frac{\sqrt{y^T y}}{1 + y^T y}$

(b) Since $P(|Z_r| \geq t) = P(|X_r^2 \cdot \text{sign}(X_r) \cdot \text{sign}(X_r + a_r^T y)| \geq t)$

$\xrightarrow{\text{ignore sign}} P(|X_r^2| \geq t) = P(X_1^2 \geq t)$, where X_1 is a Gaussian random variable $N(0,1)$

$\therefore X_1^2$ is sub-exponential so that we can easily obtain $P(X_1^2 \geq t)$

$\because Z = \frac{1}{m} \sum_{i=1}^{m} Z_r$

$\therefore Z - E(Z)$ is also a sub-exponential random variable

Hence $P\{|Z - E(Z)| \geq t\} \leq \exp(1 - \frac{t}{k_1})$

2. Proof:

(i) From Cauchy Schwartz inequality, I have $|\langle x, y\rangle| \le \|x\|_{\ell_2} \cdot \|y\|_{\ell_2}$

Thus $\langle x, Ay\rangle \le \|x\|_{\ell_2} \cdot \|Ay\|_{\ell_2}$, where $A \in R^{m\times n}$, $x \in R^m$, $y \in R^n$, and $\|x\|_{\ell_2} = \|y\|_{\ell_2} = 1$

Then $\max\limits_{\substack{\|x\|_{\ell_2}=1 \\ \|y\|_{\ell_2}=1}} \langle x, Ay\rangle \le \max\limits_{\substack{\|x\|_{\ell_2}=1 \\ \|y\|_{\ell_2}=1}} \|x\|_{\ell_2} \cdot \|Ay\|_{\ell_2} \cdots (1)$

due to $\|x\|_{\ell_2} = 1$,

The right side of inequality (1) is $\max\limits_{\|y\|_{\ell_2}=1} \|Ay\|_{\ell_2}$

Also, $\|A\| \triangleq \max\{\|Ax\|_{\ell_2} : \|x\|_{\ell_2} = 1\}$

$\therefore$ The right side of (1) is $\|A\|$

Hence $\|A\| = \max\{\langle x, Ay\rangle : x \in R^m, y \in R^n$, and $\|x\|_{\ell_2} = \|y\|_{\ell_2} = 1\}$

Since $\|A\| \triangleq \max\{\|Ax\|_{\ell_2} : \|x\|_{\ell_2} = 1\}$, I consider $\|Ax\|_{\ell_2}$

Let $A = U\Sigma V^T$ with the help of SVD decomposition

$\sup\limits_{\|x\|=1} \|Ax\| = \sup\limits_{\|x\|=1} \|U\Sigma V^T x\| \cdots (2)$

due to that for unitary U, $U^TU = I \Rightarrow \|Ux\|_{\ell_2}^2 = x^T U^T U x = \|x\|_{\ell_2}^2$

$\therefore$ (2) is equal to $\sup\limits_{\|x\|=1} \|\Sigma V^T x\|$

$\because V \in R^{m\times n}$ and unitary, $y \in R^n$ and $\|y\|_{\ell_2} = 1$

$\therefore V^T x = y$

Therefore $\sup\limits_{\|x\|=1} \|Ax\| = \sup\limits_{\|y\|=1} \|\Sigma y\|$

Since Σ is a diagnoal matrix with singular values, let $y = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

$\therefore \sup\limits_{\|x\|=1} \|Ax\| = \sigma_1(A)$, which is the maximum singular value

$\Rightarrow \|A\| = \sigma_1(A)$

(2) $\because trace(U^T A V) = \sum\limits_{i=1}^{r} u_i^T a_{ii} v_i$

$\max\left[trace(U^T A V)\right] = \max\left[\sum\limits_{i=1}^{r} u_i^T a_{ii} v_i\right]$

$= \sum\limits_{s=1}^{r} \sigma_s(A)$ since U and V are unitary matrices

3. Prove:

For the upper bound: $\|A\| = \sup_{\|x\|_2=1} \|Ax\|_2 = \sup_{\|x\|_2=1} \sqrt{\sum_{i=1}^{m} |\sum_{j=1}^{n} A_{ij} \cdot x_j|^2} \cdots (1)$

After using Cauchy Schwartz inequality: (1) is $\leq \sup_{\|x\|_2=1} \sqrt{\sum_{i=1}^{m} \left[(\sum_{j=1}^{n} A_{ij}^2) \cdot (\sum_{j=1}^{n} x_j^2) \right]}$

$$= \max_{i \in \{1,2,\dots,m\}} \sqrt{m \sum_{j=1}^{n} A_{ij}^2}$$

$$= \sqrt{m} \max_{i \in \{1,2,\dots,m\}} \sqrt{\sum_{j=1}^{n} A_{ij}^2}$$

$$\therefore \|A\| \leq \sqrt{m} \max_{i \in \{1,2,\dots,m\}} \sqrt{\sum_{j=1}^{n} A_{ij}^2}$$

For equality $|\sum_{j=1}^{n} A_{ij} \cdot x_j|^2 = \sum_{j=1}^{n} A_{ij}^2 \cdot \sum_{j=1}^{n} x_j^2$

$\therefore$ When $m=n=1$, A is a real number, the upper bound is satisfied.

For the lower bound:

consider $x = (\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}})$

$$\|A\| = \sup_{\|x\|_2=1} \|Ax\| \geq \sqrt{\sum_{i=1}^{m} (\sum_{j=1}^{n} A_{ij} \cdot \frac{1}{\sqrt{n}})^2}$$

$$= \frac{1}{\sqrt{n}} \sqrt{\sum_{i=1}^{m} (\sum_{j=1}^{n} A_{ij} \times 1)^2}$$

$$\geq \frac{1}{\sqrt{n}} \sqrt{\frac{1}{m} (\sum_{i=1}^{m} |\sum_{j=1}^{n} A_{ij}|)^2} = \frac{1}{\sqrt{mn}} \sum_{i=1}^{m} |\sum_{j=1}^{n} A_{ij}|$$

So when A is the identity matrix, the lower bound is satisfied

4. Prove: $\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}^2}$

Since $\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2 = \text{trace}(a^T a)$

$\therefore \|A\|_F = [Tr(A^T A)]^{\frac{1}{2}}$

$= [Tr(V\Sigma U^T U \Sigma V^T)]^{\frac{1}{2}}$

$= [Tr(V\Sigma\Sigma V^T)]^{\frac{1}{2}} \quad \cdots (1)$

$\because$ trace $(AB) =$ trace (BA)

$\therefore$ (1) is equal to $[Tr(V^T V\Sigma\Sigma)]^{\frac{1}{2}}$

$= [Tr(\Sigma\Sigma)]^{\frac{1}{2}}$

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

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

From problem 2, we know that $\|A\| = \sigma_1(A)$, now $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_{\min(m,n)}(A)$

$\because \sigma_1(A) \leq \sqrt{\sigma_1^2(A) + \sigma_2^2(A) + \cdots + \sigma_{\min(m,n)}^2(A)}$

$\therefore \|A\| \leq \|A\|_F$

Also, $\because \sigma_1^2(A) + \sigma_2^2(A) + \cdots + \sigma_{\min(m,n)}^2(A) \leq \underbrace{\sigma_1^2(A) + \sigma_1^2(A) + \cdots + \sigma_1^2(A)}_{\text{The \# is rank}(A)}$

$\therefore \|A\|_F \leq \sqrt{\text{rank}(A)} \cdot \|A\|$

Therefore, $\|A\| \leq \|A\|_F \leq \sqrt{\text{rank}(A)} \cdot \|A\|$.