Math 281C Homework 4 Solutions

1. A random variable X with mean μ is called *sub-Gaussian* if there is a positive parameter σ such that

$$\mathbb{E}e^{\lambda(X-\mu)} \le e^{\sigma^2\lambda^2/2}$$

for any $\lambda \in \mathbb{R}$. σ is referred to as the sub-Gaussian parameter. Suppose $\{X_i\}_{i=1}^n$ are independent with means $\{\mu_i\}_{i=1}^n$ and sub-Gaussian parameters $\{\sigma_i\}_{i=1}^n$. Show that for any $t \ge 0$,

$$\mathbb{P}\bigg(\sum_{i=1}^n (X_i - \mu_i) \ge t\bigg) \le \exp\bigg(-\frac{t^2}{2\sum_{i=1}^n \sigma_i^2}\bigg).$$

Solution: Denote $S_n = \sum_{i=1}^n (X_i - \mu_i)$, and for any $\lambda \ge 0$, by Markov inequality and the independence, we have

$$\mathbb{P}(S_n \ge t) = \mathbb{P}(e^{\lambda S_n} \ge e^{\lambda t}) \le e^{-\lambda t}\mathbb{E}[e^{\lambda S_n}] \le \exp\bigg\{-\lambda t + \lambda^2 \sum_{i=1}^n \sigma_i^2/2\bigg\}.$$

Taking $\lambda = t/\sum_{i=1}^n \sigma_i^2$ completes the proof.

2. Define the class of matrices

$$\mathbb{M}_{(1)}^{n,d} := \{\Theta \in \mathbb{R}^{n\times d} : \mathrm{rank}(\Theta) = 1, \|\Theta\|_{\mathrm{F}} = 1\},$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm. Prove that for any $0 < \epsilon < 1$,

$$\log N(\epsilon, \mathbb{M}_{(1)}^{n,d}, \|\cdot\|_{\mathrm{F}}) \le (n+d)\log(1+4/\epsilon).$$

Solution: The definition of $\mathbb{M}_{(1)}^{n,d}$ is equivalent as

$$\mathbb{M}_{(1)}^{n,d} := \{u \cdot v^\top : u \in \mathbb{S}^{n-1}, v \in \mathbb{S}^{d-1}\},$$

where $\mathbb{S}^{n-1}$ and $\mathbb{S}^{d-1}$ are unit spheres in $\mathbb{R}^n$ and $\mathbb{R}^d$ respectively. Suppose $\{u^1, \dots, u^{N_1}\}$ is a $(\epsilon/2)$-net of $\mathbb{S}^{n-1}$ and $\{v^1, \dots, v^{N_2}\}$ is a $(\epsilon/2)$-net of $\mathbb{S}^{d-1}$, both with respect to $\|\cdot\|_2$. For any $\Theta \in \mathbb{M}_{(1)}^{n,d}$, there are $u \in \mathbb{S}^{n-1}$ and $v \in \mathbb{S}^{d-1}$ such that $\Theta = u \cdot v^\top$, and for u and v, there are u^i and v^j in the aforementioned $(\epsilon/2)$-nets such that

$$\|u - u^i\|_2 \le \epsilon/2 \quad \text{and} \quad \|v - v^j\|_2 \le \epsilon/2.$$

This implies that for any $\Theta \in \mathbb{M}_{(1)}^{n,d}$, there exist u^i and v^j such that

$$\begin{aligned}
\|\Theta - u^i \cdot (v^j)^\top\|_{\mathrm{F}} &= \|u \cdot v^\top - u^i \cdot v^\top + u^i \cdot v^\top - u^i \cdot (v^j)^\top\|_{\mathrm{F}} \\
&= \|(u - u^i) \cdot v^\top + u^i \cdot (v - v^j)^\top\|_{\mathrm{F}} \\
&\le \|u - u^i\|_2 + \|v - v^j\|_2 \le \epsilon,
\end{aligned}$$

which means $\{u^i \cdot (v^j)^\top\}_{i=1,\dots,N_1, j=1,\dots,N_2}$ is a ϵ-net of $\mathbb{M}_{(1)}^{n,d}$. By Proposition 2.1 of Lecture 6, we have

$$N_1 \le (1+4/\epsilon)^n \quad \text{and} \quad N_2 \le (1+4/\epsilon)^d.$$

Combining all the results gives us

$$\log N(\epsilon, \mathbb{M}_{(1)}^{n,d}, \|\cdot\|_{\mathrm{F}}) \le \log N_1 + \log N_2 \le (n+d)\log(1+4/\epsilon).$$