Homework 4 - High Dimensional Probability

Bernardi Alberto, Cracco Gianmarco,
Manente Alessandro, Mazzieri Riccardo

January 2021

Exercise 1

Let A be a subset of the sphere $\sqrt{n}S^{n-1}$ such that

$$\sigma(A) > 2\exp\left(-cs^2\right) \qquad \text{for some } s > 0$$

(i) Prove that $\sigma(A_s) > 1/2$

(ii) Deduce from this that for any $t \geq s$

$$\sigma(A_{2t}) \geq 1 - 2\exp\left(-ct^2\right).$$

Solution:

(i) $\sigma(A_s) > \frac{1}{2}$: suppose by contradiction that $\sigma(A_s) \leq \frac{1}{2}$. This would imply that $\sigma(B) \geq \frac{1}{2}$, where $B = (A_s)^c$. We could therefore apply the Blow Up Lemma[1] on B:

$$\sigma(B_s) \geq 1 - 2\exp\left(-cs^2\right) > 1 - \sigma(A) \Longrightarrow \sigma(B_s) + \sigma(A) > 1.$$

This is a contradiction since $B = (A_s)^c \Longleftrightarrow B_s = A^c \Longrightarrow \sigma(B_s) + \sigma(A) = 1$. Therefore $\sigma(A_s) > \frac{1}{2}$, and the first point is proven.

(ii) Consider $t \geq s$. Since we proved that $\sigma(A_s) > \frac{1}{2}$, we can apply the Blow Up Lemma on A_s and obtain an inequality of this form:

$$\sigma((A_s)_\varepsilon) \geq 1 - 2\exp\left(-c\varepsilon^2\right) \tag{1}$$

Where we want to put $(A_s)_\varepsilon = A_{2t}$ (this makes sense since $2t > s$). Now we only need to find the value of ε and plug it into inequality (1).

$$\varepsilon = 2t - s = t + (t - s) = t + \tilde{c}.$$

[1] Let A be a subset of the sphere $\sqrt{n}S^{n-1}$ and let σ denote the normalized area on that sphere. If $\sigma(A) \geq 1/2$, then $\forall t \leq 0$, $\sigma(A_t) \geq 1 - 2\exp\left(-ct^2\right)$.

Where $\tilde{c} \geq 0$ since $t \geq s$. Therefore we have:

$$\sigma(A_{2t}) \geq 1-2\exp\left(-c(t+\tilde{c})^2\right) = 1-2\exp\left(-ct^2 - c\tilde{c}^2 - 2tc\tilde{c}\right) \geq 1-2\exp\left(-ct^2\right)$$

Since $c\tilde{c}^2, 2tc\tilde{c} \geq 0$.

■

Exercise 2

Consider a random vector X taking values in some metric space (T, d). Assume that there exists $K > 0$ such that

$$\|f(X) - \mathbb{E}f(X)\|_{\psi_2} \leq K\|f\|_{\text{Lip}} \tag{2}$$

for every Lipschitz function $f : T \to \mathbb{R}$. For a subset $A \subset T$, define $\sigma(A) := \mathbb{P}(X \in A)$. (Then σ is a probability measure on T).
Show that if $\sigma(A) \geq \frac{1}{2}$, then, for every $t \geq 0$,

$$\sigma\left(A_t\right) \geq 1 - 2\exp\left(-\frac{ct^2}{K^2}\right)$$

where $c > 0$ is an absolute constant.

Solution:

We apply the assumption to $f(x) := dist(x, A) = \inf\{d(x,y) : y \in A\}$ and we call M the median of $f(X)$. To find M we observE:

$$\sigma(A) \geq \frac{1}{2} \Longrightarrow \mathbb{P}\{X \in A\} \geq \frac{1}{2} \Longleftrightarrow \mathbb{P}\{f(X) = 0\} \geq \frac{1}{2} \Longrightarrow M = 0.$$

Recalling the result of Exercise 5.1.13 in the book and combining it with (2) and the fact that $M = 0$, we obtain

$$\|f(X)\|_{\psi_2} \leq C\|f(X) - \mathbb{E}f(X)\|_{\psi_2} \leq CK\|f\|_{\text{Lip}}, \quad \exists C, K > 0. \tag{3}$$

Now we can conclude observing that

$$\begin{aligned}\sigma(A_t) &= \mathbb{P}\{X \in A_t\} = \mathbb{P}\{f(X) \leq t\} \\ &= \mathbb{P}\{|f(X)| \leq t\} \geq 1 - 2\exp\left(-\frac{ct^2}{\|f(X)\|_{\psi_2}^2}\right) \\ &\geq 1 - 2\exp\left(-\frac{ct^2}{K^2}\right),\end{aligned}$$

where the first inequality follows from the properties of a sub-gaussian random variable.

■

Exercise 3

Let D be a $k \times m$ matrix and B be an $m \times n$ matrix. Prove that

$$\|DB\|_F \leq \|D\| \|B\|_F.$$

Solution:

$$\begin{aligned}
\|DB\|_F = \sqrt{\sum_{i=1}^{\min(k,n)} \sigma_i^2(DB)} &\overset{(2)}{\leq} \sqrt{\sum_{i=1}^{\min(k,n)} \sigma_1^2(D)\sigma_i^2(B)} \\
&= \sigma_1(D) \cdot \sqrt{\sum_{i=1}^{\min(k,n)} \sigma_i^2(B)} \\
&= \|D\| \cdot \sqrt{\sum_{i=1}^{\min(k,n)} \sigma_i^2(B)} \\
&= \|D\| \cdot \|B\|_F
\end{aligned}$$

where σ_i are the singular values, (2) holds since $\sigma_i^2(DB) \leq \sigma_1^2(D)\sigma_i^2(B)$ and the last equalities were obtained in Homework 3.

■

Exercise 4

Cousider *i.i.d.* random variables $\delta_{ij} \sim \text{Ber}(p)$ where $i, j = 1, \ldots, \mu$. Assuming that $pn \geqslant \log n$, show that

$$\mathbb{E} \max_i \sum_{j=1}^{n} \left(\delta_{ij} - p\right)^2 \leqslant Cpn.$$

Solution:

$$\begin{aligned}
\exp\left(\mathbb{E} \max_i \sum_j \left(\delta_{ij} - p\right)^2\right) \overset{\text{Jensen}}{\leq} \mathbb{E} \exp\left(\max_i \sum_j \left(\delta_{ij} - p\right)^2\right) &\overset{\star}{=} \mathbb{E} \max_i \exp\left(\sum_j \left(\delta_{ij} - p\right)^2\right) \\
&\leq \mathbb{E} \sum_{i=1}^{n} \exp\left(\sum_j \left(\delta_{ij} - p\right)^2\right) = \sum_{i=1}^{n} \mathbb{E} \exp\left(\sum_j \left(\delta_{ij} - p\right)^2\right) \\
&= n\mathbb{E}\left[\prod_{j=1}^{n} \exp\left((\delta_{ij} - p)^2\right)\right] = n \prod_{j=1}^{n} \mathbb{E}\left[\exp\left((\delta_{ij} - p)^2\right)\right] \\
&= n\left(\mathbb{E}\left[\exp\left((\delta_{i1} - p)^2\right)\right]\right)^n \\
&= n\left(p \exp\left((1-p)^2\right) + (1-p)\exp\left(p^2\right)\right)^n,
\end{aligned}$$

where in $\star$ we leveraged the monotonicity of the exponential and in the last equalities we the i.i.d. property of our variables.
We now show that $\exists k$ such that $p \exp\left((1-p)^2\right) + (1-p)\exp\left(r^2\right) \leqslant \exp(kp)$:

- If $p \geqslant 1-p$: $\quad \exp\left((1-p)^2\right) \leqslant \exp\left(p^2\right) \leqslant \exp(p)$

  $\Longrightarrow p \exp\left((1-p)^2\right) + (1-p)\exp\left(p^2\right) \leq \exp(p) \Longrightarrow k = 1;$

- If $1-p > p$: $\quad \exp\left((1-p)^2\right) + (1-p)\exp\left(p^2\right) \leq \exp\left((1-p)^2\right) \leqslant \exp(1-p)$, which is smaller than $\exp(kp)$, for $k > \frac{1}{p}$.

Hence we have:

$$\exists k \quad s.t. \quad \exp\left(\mathbb{E}\max_i \sum_j \left(\delta_{ij} - p\right)^2\right) \leq ne^{kpn}.$$

Finally, taking the $log(\cdot)$ of both sides:

$$\mathbb{E}\max_i \sum_j \left(\delta_{ij} - p\right)^2 \leq log(n) + kpn \leq pn + kpn = (1+k)pn = Cpn.$$

$\blacksquare$