Problem sheet # 3

The number of ☕ symbols is proportional to the amount of coffee needed to solve each question. Do not hesitate to contact `aubrun@math.univ-lyon1.fr` for hints.

Exercise 1.1 Subgaussian random variables

Let X be a random variable. Define the following quantities

$$\|X\|_{\psi_2} = \sup_{p \geqslant 1} \frac{\|X\|_{L^p}}{\sqrt{p}},$$

$$\|X\|'_{\psi_2} = \inf\{t \geqslant 0 \,:\, \mathbf{E}\exp(|X/t|^2) \leqslant 2\},$$

$$Tail(X) = \inf\{K > 0 \,:\, \mathbf{P}(|X| \geqslant t) \leqslant 2\exp(-t^2/K^2) \text{ for every } t \geqslant 0\}$$

1. ☕☕☕ Show that these 3 quantities are equivalent, in the sense that

   $$\|\cdot\|_{\psi_2} \leqslant C_1 \|\cdot\|'_{\psi_2}, \quad \|\cdot\|'_{\psi_2} \leqslant C_2 Tail(\cdot), \quad Tail(\cdot) \leqslant C_3 \|\cdot\|_{\psi_2}$$

   for some absolute constants C_1, C_2, C_3.

2. ☕☕☕ Show that if (X_i) are independent subgaussian random variables with mean zero, then

   $$\left\| \sum X_i \right\|_{\psi_2} \leqslant C \left(\sum \|X_i\|_{\psi_2}^2 \right)^{1/2}$$

   for some absolute constant C.

3. ☕☕ Show that if $X_1, \dots, X_N$ are subgaussian random variables, then

   $$\mathbf{E} \sup_{1 \leqslant i \leqslant N} X_i \leqslant C\sqrt{\log N} \sup_{1 \leqslant i \leqslant N} \|X_i\|_{\psi_2}$$

   for some absolute constant C.

Exercise 1.2 Nets and packing in the sphere.

☕☕ Show that if (T,d) is a metric space, then for every $\varepsilon > 0$,

$$P(T,d,2\varepsilon) \leqslant N(T,d,\varepsilon) \leqslant P(T,d,\varepsilon).$$

☕☕ Let d be the distance on S^{n-1} induced from the Euclidean distance. Show that if $0 < \varepsilon < 1$, then

$$N(S^{n-1}, d, \varepsilon) \leqslant \left(1 + \frac{2}{\varepsilon}\right)^n \leqslant (3/\varepsilon)^n$$

(to bound the cardinality of an ε-separated set $P \subset S^{n-1}$, consider the volume of the union of n-dimensional balls of radius ε centered at P)

Exercise 1.3 An example where Dudley inequality is not sharp.

☕☕☕ Let (Z_k) be a sequence of i.i.d. $N(0,1)$ random variables and set $X_k = Z_k/\sqrt{1+\log k}$. Show that $\mathbf{E}\sup X_k$ is finite, but that the upper bound given by Dudley's integral is of order $\log\log n$ for the Gaussian process $(X_k)_{1 \leqslant k \leqslant n}$.

Exercise 1.4 Concentration on a subspace via the Dudley integral

Let $f : S^{n-1} \to \mathbf{R}$ be a 1-Lipschitz function with mean zero. Let $O \in \mathsf{O}(n)$ be a random matrix distributed according to the Haar measure. Let E_0 be a fixed subspace of dimension k, and $E = O(E_0)$; then E is distributed according the to uniform measure $\mu_{n,k}$.

1. Let $\mathcal{N}$ be a ε-net in the unit sphere of E_0. ☕☕ Show that

   $$\mathbf{P}(\exists x \in S^{n-1} \cap E_0, |f(O(x))| > \varepsilon) \leqslant \mathbf{P}(\exists x \in \mathcal{N}, |f(O(x))| > 2\varepsilon) \leqslant (\operatorname{card}\mathcal{N}) 2\exp(-cn\varepsilon^2)$$

   (you need to realize that for every $x \in S^{n-1}$, the distribution of $O(x)$ is σ) and conclude that with high probability, the supremum of $|f|$ on a random k-dimensional subspace is $\leqslant \varepsilon$ whenever $k \leqslant c\varepsilon^2 \log(1/\varepsilon) n$.

2. In order to improve the above analysis, we now use the Dudley integral. Let (S, d) the unit sphere in E_0 equipped with the induced Euclidean distance. ☕☕☕ Show that the process $(X_s)_{s \in S}$ defined by $X_s = f(O(s))$ is subgaussian with constant $C/\sqrt{n}$. ☕☕ Deduce that

$$\mathbf{E} \sup_{x \in S} X_s \leqslant \frac{C}{\sqrt{n}} \int_0^\infty \sqrt{\log N(S, d, \varepsilon)} \, \mathrm{d}\varepsilon \leqslant C' \sqrt{k/n}$$

and ☕ conclude that the factor $\log(1/\varepsilon)$ can be removed from the conclusion in question 1.