Mathematics of Machine Learning
Homework 8

Instructors: Afonso S. Bandeira & Nikita Zhivotovskiy
Course Coordinator: Pedro Abdalla Teixeira

April 23, 2021

Try to solve the questions before looking to the answers. Every item must be proved rigorously. Starred problems are harder.

Problem 1

Let $g_1, \dots, g_n$ denote a collection of standard Gaussian random variables, not necessarily independent. The goal of this exercise is to provide a simple bound for the expectation of the maxima of Gaussian random variables

(a) Prove that, for every $\lambda > 0$, we have

$$\mathbb{E}e^{\lambda \max_{i=1,\dots,n} g_i} \leq ne^{\lambda^2/2}$$

(b) Prove that, for every $\lambda > 0$,

$$e^{\lambda \mathbb{E} \max_{i=1,\dots,n} g_i} \leq \mathbb{E}e^{\lambda \max_{i=1,\dots,n} g_i}$$

(Hint: Apply Jensen's inequality)

(c) Prove that

$$\mathbb{E} \max_{i=1,\ldots,n} g_i \leq C\sqrt{\log n},$$

where $C > 0$ is an absolute constant. (Hint: Follows the steps above and optimize in λ)

Problem 2

Let X be a bounded random variable taking values in the interval $[a.b]$. The goal of this exercise is to prove a sharper version of Hoeffiding's lemma.

(a) Prove that

$$\mathbb{E}e^{\lambda X} \leq \frac{b - \mathbb{E}X}{b - a}e^{\lambda a} + \frac{\mathbb{E}X - a}{b - a}e^{\lambda b}.$$

(Hint: Use the convexity of the function $e^{\lambda x}$ and take expectation of both sides.)

(b) Let $\theta := \frac{-a}{b-a}$, $u := \lambda(b-a)$ and $f(u) := -\theta u + \log(1 - \theta + \theta e^u)$. Check that

$$\mathbb{E}e^{\lambda(X - \mathbb{E}X)} \leq e^{f(u)}.$$

(c) Use Taylor expansion to bound $e^{f(u)}$ by $e^{\frac{\lambda^2(b-a)^2}{8}}$.

(d) Derive a sharper Hoeffiding lemma.

Problem 3

A random variable X is said to be symmetric if X and $-X$ have the same distribution.

(a) Check that the standard Gaussian random variable is symmetric. Give an example of a random variable that is not symmetric.

(b) Let ε denote a Rademacher random variable, i.e, takes values ± 1 with probability $\frac{1}{2}$ each. Prove that if X is a symmetric random variable independent on ε, then εX and X have the same distribution.

(c) Prove via conditional expectation that if X is a symmetric random variable then all odd moments vanish, i.e, $\mathbb{E}X^k = 0$ for all odd k.