Mathematics of Machine Learning
Homework 11

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

May 14, 2021

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

Problem 1

We say that two unit vectors $u, v \in \mathbb{R}^d$ are ε-almost orthogonal if $|\langle u, v\rangle| \leq \varepsilon$. A set of vectors in which every vector is ε-almost orthogonal to each other is called ε-almost orthogonal set. For example, for $\varepsilon = 0$, we recover the usual notion of orthogonal vectors and the canonical basis is an 0-almost orthogonal set. Clearly, the maximal size of an 0-orthogonal set in $\mathbb{R}^d$ is d. The goal of this exercise is to show that $\varepsilon > 0$, the size of an ε-almost orthogonal set is exponentially large in d and discuss its relevance for frame theory. The proof strategy relies on the probabilistic method, a nonconstructive strategy to show the existence of certain object by using probability.

The following inequality will be useful (you do not need to prove it), if $X_1, \dots, X_n$ are i.i.d Rademacher random variables, i.e, taking values 1 and -1 with probability $\frac{1}{2}$ each, then

$$\mathbb{P}(\sum_{i=1}^{n} X_i \geq t) \leq e^{\frac{-t^2}{2n}}.$$

(a) Let $\varphi \in \mathbb{R}^d$ be a random vector drawn as follows: Each entry is generated independently at random following a Rademacher distribution. Consider N i.i.d copies of the random vector φ, namely $\varphi_1, \dots, \varphi_N$. Prove the following statement: If $N = \lfloor e^{d\varepsilon^2/4} \rfloor$, then the probability that set of normalized vectors $\{\frac{1}{\sqrt{d}}\varphi_1, \dots, \frac{1}{\sqrt{d}}\varphi_N\}$ is an ε-almost orthogonal set is positive (strictly larger than 0).

(b) Conclude that letter "a" implies in the existence of ε-almost orthogonal set with size exponentially large in d.

(c) Give an interpretation of this fact in terms of mutual coherence of frames in the sense of Challenge 11.4 in the lecture notes.

(d) * Improve the estimate of letter "a" from $N = \lfloor e^{d\varepsilon^2/4} \rfloor$ to $N = \lfloor ce^{d\varepsilon^2/2} \rfloor$, where c is an absolute constant less than one. (Hint: Compute the expected number of unordered pair that violates the condition of ε-almost orthogonality, then delete one vector from each pair).

Problem 2

The goal of this exercise is to solve the Challenge 12.4 in the lecture notes. Recall that a random graph $G \sim \mathcal{G}(n,p)$ is a graph of n vertices generated by placing each possible edge independently at random with probability p. We use the standard asymptotic notation, $f(n) \ll g(n)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$, also $g(n) \gg f(n)$ if $\lim_{n\to\infty} \frac{g(n)}{f(n)} = \infty$. Moreover $f(n) = o(1)$ if $\lim_{n\to\infty} f(n) = 0$. Finally we define $p := \frac{\lambda \log n}{n}$ for some constant $\lambda > 0$.

(a) Prove that if $\lambda \leq 1 - c$, where $c > 0$ is an absolute constant, then the graph G has an isolated vertex with probability $1 - o(1)$. (Hint: Consider a random variable X_n to be the number of isolated vertices in the graph G, you need to prove that the probability of X_n be larger than zero is $1 - o(1)$. Use the inequality $\mathbb{P}(X_n = 0) \leq \frac{\operatorname{Var} X_n}{\mathbb{E} X_n^2}$).

(b) Now observe the following: A graph is disconnected if and only if thre exits a set of k nodes such that $k \leq \left\lfloor \frac{n}{2} \right\rfloor$ and there is no edge connecting the set of k nodes with the complement set of $n - k$ nodes. Use this fact to prove that if $\lambda \geq 1 + c$ for an absolute constant $c > 0$, then the graph is connected with probability $1 - o(1)$. (Hint: Use union bound twice).

Problem 3

Let $\mathcal{A}$ and $\mathcal{B}$ be two family of events with finite VC-dimension. Prove the following facts:

(a) If $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$, then $\mathcal{S}_{\mathcal{C}}(n) \leq \mathcal{S}_{\mathcal{A}}(n) + \mathcal{S}_{\mathcal{B}}(n)$.

(b) If $\mathcal{C} = \mathcal{A} \cap \mathcal{B}$, then $\mathcal{S}_{\mathcal{C}}(n) \leq \mathcal{S}_{\mathcal{A}}(n)\mathcal{S}_{\mathcal{B}}(n)$.