Mathematics of Machine Learning
Homework 1

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

February 26, 2021

Try to solve the questions before looking to the answers. Every item must be proved rigorously.

Problem 1

Given a set of points $x_1 \dots, x_n \in \mathbb{R}^d$ and a partition of them into k clusters $S_1, \dots, S_k$, recall that the k-means objective is

$$\min_{S_1,\dots,S_k} \min_{\mu_1,\dots,\mu_k} \sum_{l=1}^{k} \sum_{i \in S_l} \|x_i - \mu_l\|_2^2.$$

If $\hat{S}_1, \dots, \hat{S}_k$ are the minimizers for the k-means objective, show that they are also minimizers for the optimization problem below

$$\min_{S_1,\dots,S_k} \sum_{l=1}^{k} \frac{1}{|S_l|} \sum_{i,j \in S_l} \|x_i - x_j\|_2^2.$$

Problem 2

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ its eigenvalues. Describe the solutions of the optimization problems below

(a)

$$\max_{x\in\mathbb{S}^{n-1}} \langle Ax, x\rangle.$$

(b)

$$\max_{x\in\mathbb{S}^{n-1}} \|Ax\|_2.$$