

Introduction to Machine Learning  
Instructor: Lara Dolecek  
TA: Zehui (Alex) Chen, Ruiyi (John) Wu

1. Here is an example where we would want to regularize clusters. Suppose  $n$  students are seated for taking an endterm exam in an  $\mathbf{R}^2$  Euclidean room. There are  $K$  teaching assistants who must collect the answer scripts once the time is up. The TAs need to figure out good locations to position themselves so that the students can walk to the nearest TA and submit their answers. Once the TAs have all the answer sheets, they must return to the front desk located at  $(0,0)$  while handling the returned answer sheets carefully. To reduce the possibility of mishaps related to handling of the papers, write down an objective which can be used to minimize the total distance that both students and TAs need to walk to bring the papers to the front desk. Assume that everyone can walk by taking the shortest path between two points.

2. We learned that the  $\mu$  that maximize

$$\sum_{i=1}^N (x_i - \mu)^2$$

for  $x_i \in \mathbf{R}$  is given by the mean of  $\{x_1, \dots, x_N\}$ , i.e.,  $\mu^* = \frac{1}{N} \sum_{i=1}^N x_i$ .

Show that the  $\mu$  that maximize

$$\sum_{i=1}^N |x_i - \mu|$$

is given by the mode of  $\{x_1, \dots, x_N\}$ . What if  $x_i \in \mathbf{R}^n$ ?

3. In class, you learned that the direction that maximize the variance of the projection onto a one-dimensional space is the eigenvector that corresponds to the largest eigenvalue of the data covariance matrix  $S = \frac{1}{N}X^T X$ . Formally, the solution to the following maximization problem

$$\max_{u_1} u_1^T S u_1 \quad \text{subject to } \|u_1\|^2 = 1,$$

is the eigenvector that corresponds to the largest eigenvalue of  $S$ .

Suppose  $u_2$  is orthogonal to  $u_1$  and have unit norm. We want to maximizes the variance of the data projected on  $u_2$ . Show that the optimal  $u_2$  is defined by the second eigenvectors of the data covariance matrix  $S$  that corresponds to the second largest eigenvalues.