Exercises for the course
Machine Learning for Data Science
Winter Semester 2024/25

G. Montavon
Institute of Computer Science
Department of Mathematics and Computer Science
Freie Universität Berlin

Exercise Sheet 5 (theory part)

Exercise 1: K-Means Clustering (15 + 10 P)

The K-means optimization problem is given by $\arg\min_{\boldsymbol{\mu},\boldsymbol{c}} \sum_{i=1}^N \|\boldsymbol{x}_i - \boldsymbol{\mu}_{c_i}\|^2$ where $\boldsymbol{c} \in \{1,\dots,K\}^N$ is the cluster assignment function. When considering the latter to be fixed, and only letting the centroids $\boldsymbol{\mu}$ vary, the optimization problem can be restated as:

$$\arg\min_{\boldsymbol{\mu}} \sum_{k=1}^K \sum_{i\in\mathcal{C}_k} \|\boldsymbol{x}_i - \boldsymbol{\mu}_k\|^2$$

where $\mathcal{C}_k$ is the set of instances that are assigned to cluster k.

(a) Show that the solution of this optimization problem is given by:

$$\boldsymbol{\mu} = (\boldsymbol{\mu}_k)_{k=1}^K \quad \text{where} \quad \boldsymbol{\mu}_k = \frac{\sum_{i\in\mathcal{C}_k} \boldsymbol{x}_i}{\sum_{i\in\mathcal{C}_k} 1}$$

(b) A data point $\boldsymbol{x}$ is assigned onto cluster c if

$$\forall_{k\,:\,k\neq c}:\ \|\boldsymbol{x} - \boldsymbol{\mu}_c\| < \|\boldsymbol{x} - \boldsymbol{\mu}_k\|.$$

Show that this condition for assignment onto cluster c can be equivalently formulated as a min-pooling over affine functions, specifically, we assign to cluster c if

$$\min_{k\,:\,k\neq c} \big\{\boldsymbol{w}_k^\top \boldsymbol{x} + b_k\big\} > 0$$

where $\boldsymbol{w}_k = (\boldsymbol{\mu}_c - \boldsymbol{\mu}_k)$ and $b_k = \frac{1}{2}(\|\boldsymbol{\mu}_k\|^2 - \|\boldsymbol{\mu}_c\|^2)$.

Exercise 2: Spectral Clustering (15 + 10 P)

In the lecture, it was mentioned that the eigenvalues λ of the Laplacian matrix $L = D - A$ (where D and A are the degree and adjacency matrices respectively) can be related to the corresponding eigenvector u as:

$$\lambda = \frac{1}{2}\sum_{i=1}^N \sum_{j=1}^N A_{ij}\big(u_i - u_j\big)^2.$$

(a) *Prove* the equation above.

(b) From the equation above, we can see that the eigenvalue λ influences the extent by which the associated eigenvector u can vary between connected nodes.

Show that eigenvectors associated to the eigenvalue $\lambda = 0$, cannot vary within a connected component, that is, denoting by $\boldsymbol{u}$ the eigenvector, show that if i and j are part of the same connected component (i.e. if there is a sequence of edges connecting i and j), then $u_i = u_j$.