Machine Learning

Exercise Sheet 12

Winter Term 2023 — Available: 09.02.2024
Prof. Dr. Niels Landwehr — Hand in until: 16.02.2024 11:59am
Dr. Ujjwal — No exercise session, model solution will be made available

Task 1 – K-Means [5 points]

As stated in the lecture (Slide 7), the optimization criterion of K-Means is given by

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} ||\mathbf{x}_n - \boldsymbol{\mu}_k||^2 \tag{1}$$

Show that, as claimed in the lecture, the minimum in $\boldsymbol{\mu}_k$ of the expression J is given by

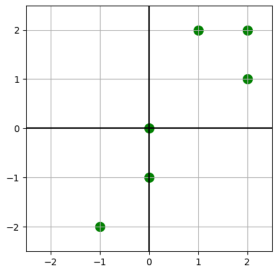
$$\boldsymbol{\mu}_k = \frac{\sum_{n=1}^{N} r_{nk} \mathbf{x}_n}{\sum_{n=1}^{N} r_{nk}} \tag{2}$$

by setting the derivative $\frac{\partial J}{\partial \boldsymbol{\mu}_k}$ to zero and resolving for $\boldsymbol{\mu}_k$.

Task 2 – Gaussian Mixture Model [15 points]

In this task, we study Gaussian mixture models for a small toy data set. Consider the data set $\{\mathbf{x}_1, ..., \mathbf{x}_6\}$ with $\mathbf{x}_n \in \mathbb{R}^2$ given by

	x_1	x_2
$\mathbf{x}_1$	1	2
$\mathbf{x}_2$	2	2
$\mathbf{x}_3$	2	1
$\mathbf{x}_4$	0	0
$\mathbf{x}_5$	0	-1
$\mathbf{x}_6$	-1	-2



We want to cluster this data using a Gaussian mixture model with $K = 2$ cluster components. Execute one iteration of the EM-algorithm for the Gaussian mixture model for this data set, using initial cluster probabilities $\pi_1 = \pi_2 = 0.5$, initial cluster means $\boldsymbol{\mu}_1 = (1, 1)^\mathsf{T}$ and $\boldsymbol{\mu}_2 = (-1, -1)^\mathsf{T}$, and initial covariance matrices $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \mathbf{I}$, where $\mathbf{I} \in \mathbb{R}^{2\times 2}$ denotes the 2-by-2 identity matrix.
Specifically, report the responsibilities $\gamma(z_{nk})$, new cluster probabilities $\pi_k^{(*)}$, new cluster means $\boldsymbol{\mu}_k^{(*)}$, and covariance matrices $\boldsymbol{\Sigma}_k^{(*)}$ for the first iteration of the EM-algorithm. Hint: implement the necessary calculations in a matrix programming language such as Numpy.