Exercises
Computational Intelligence Lab
SS 2012

Machine Learning Laboratory
Dept. of Computer Science, ETH Zürich
Prof. Dr. Joachim M. Buhmann
Web http://cil.inf.ethz.ch/

Series 4, March 14-15, 2013 (The K-means Algorithm)

Problem 1 (K-means Theory):

In this exercises, you will elaborate on some of the formal results connecting K-means Theory and Matrix Factorization.

1. Show that the K-means algorithm always converges. Consider the following cost function:

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} z_{k,n} \|\mathbf{x}_n - \mathbf{u}_k\|_2^2.$$

   Show that step 2 and 3 of the K-means algorithm from the lecture minimize this cost function for $\mathbf{z}_n$ and $\mathbf{u}_k$, respectively.

2. Here, you will formally see how the K-means Algorithm can be recast as a Matrix Factorization problem.

   - Show that at **Step 2**, for a given $\mathbf{u}$, the K-means algorithm solves:

$$\min_{\mathbf{Z}} \sum_{n=1}^{N} \sum_{k=1}^{K} \|\mathbf{x}_n - z_{k,n}\mathbf{u}_k\|_2^2$$

   - Show that at **Step 3**, for a given $\mathbf{Z}$, the K-means algorithm solves:

$$\min_{\mathbf{u}} \sum_{n=1}^{N} \sum_{k=1}^{K} \|\mathbf{x}_n - z_{k,n}\mathbf{u}_k\|_2^2$$