Homework 2

Submission Instructions

- Homework is due on: **Thursday 15/08/19 23:55**.
- Homework should be done **only in pairs**. Each pair is to do their own work, separate from the other pairs.
- We prefer you type your submission, however, you may submit scanned handwritten material as long as it is **clear and readable**.
- Submit **only one** PDF file. Please **write your ID** on the top of the file.
- Submission is done via **Moodle** website.
- Homework can be done using either MATLAB or Python.

Question 1

Vectors $x \in \mathbb{R}^6$ are generated in the following manner:

First, we sample $z \sim \mathcal{N}(\mu, \Sigma)$ where

$$\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 21 & 0 & 0 \\ 0 & 0 & 13 & 8 \\ 0 & 0 & 8 & 13 \end{bmatrix}.$$

Then, apply a linear transform $A \in \mathbb{R}^{6\times 4}$ on z to get $x = Az$ where $A^T A = I$ (See Tutorial 6).

(a) Calculate the eigenvalues of the covariance matrix Σ_x of the random vector x. Compute the sum of eigenvalues denoted by S_λ.

(b) How many PCA directions should be taken so that the mean reconstruction error will be smaller than $S_\lambda/4$?

(c) We define new vector $y \triangleq x + v \in \mathbb{R}^6$ where $v \in \mathbb{R}^6$. The vector $v \sim \mathcal{N}(0, \Sigma_z)$ is normally distributed where

$$\Sigma_z = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}.$$

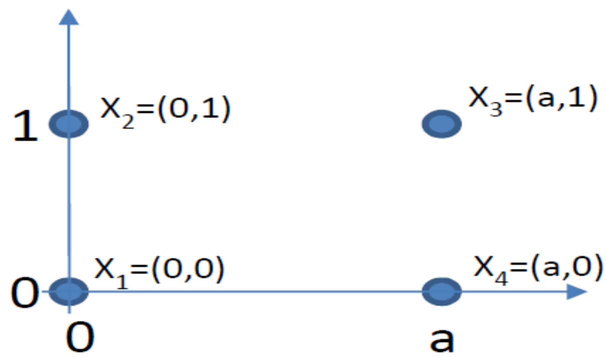
Repeat sections (a) and (b) for the vector y (note that the sum of the eigenvalues has changed).

Question 2

Reminder: The K-Means algorithm divides samples into clusters, and changes a sample's cluster association from cluster A to cluster B only if its distance to the centroid of group B is strictly smaller than its distance to the centroid of group A.

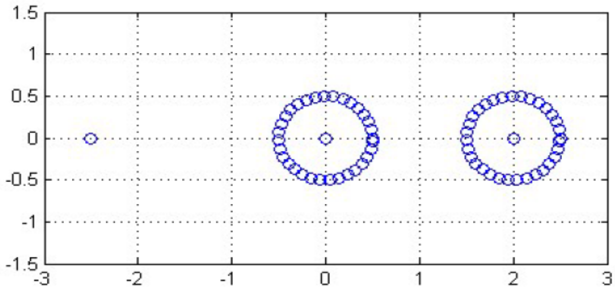
Fix-Point: A specific clustering of the samples into certain clusters is called a **fix point** of K-Means, if the algorithm does not change the sample clustering when it used as initial conditions.

Consider a training set of four samples in $\mathbb{R}^2$ located on the vertices of a rectangle parallel to the axes (see figure below). Find all possible fixed point divisions/clusterings of the four samples into two groups. Your answer may depend on the rectangle's width parameter $a > 0$.



Question 3

For the set of samples shown in the figure below and for K-Means with k=2, is there one option for convergence or more? If only one option exists, write it down and explain why it is the only option. If there is more than one option, mention at least two other options, and explain which one of them attains a lower value of the objective function.



Question 4

Reminder: The perceptron learning algorithm

Input: set of labeled examples $\{x_i, d_i\}_{i=1}^{n}$ where $d_i \in -1, 1$ and $x_i \in \mathbb{R}^m$.
Initialization: the weights vector w_0 is initialized with zeros.
For each step $t = 1, 2, ..$:

- Choose one example x_t from the dataset.
- Calculate the perceptron output for that sample using the current weight vector w_t:

$$y_t = sign(w_y^T x_t).$$

- Update the weights vector - $w_{t+1} = w_t + \eta(d_t - y_t)x_t$.

Assume

- $sign(0) = 0$.
- $\eta = \frac{1}{2}$.
- The example x_t that the algorithm receives at step t is one of the columns of an orthogonal matrix with size $D \times n$ (i.e. the norm of each example is one $||x_t||_2 = 1$).
- The dataset is linearly separable.

How many updates will the algorithm preform, at the least?