

WRITE FIRST NAME, LAST NAME, AND ID NUMBER (“MATRICOLA”) ON YOUR ASSIGNMENT. TIME: 1.5 hours.

FIRST NAME:

LAST NAME:

ID NUMBER:

Exercise 1 [6 points]

1. Discuss the general classification problem of machine learning.
2. Introduce the logistic regression model for binary classification discuss how we can train that model based on training data \mathcal{S} .

Exercise 2 [6 points]

1. Explain the main difference between supervised learning and unsupervised learning and introduce the clustering problem
2. Define the Gaussian Mixture Model for clustering and, in particular:
 - Explain the main idea behind the mixture model
 - Assuming the GMM has been learned from data and is GIVEN, once a new (unlabeled) data point becomes available, based on which criterion would you assign this point to one cluster? Would you call this a “hard” assignment or a “soft” assignment (explain).

Exercise 3 [6 points]

Consider a linear regression problem with predictor function

$$h_w(x) = x^\top w \quad x \in \mathbb{R}^2 \quad w \in \mathbb{R}^2$$

Let us denote with (x_i, y_i) , $i = 1, \dots, m$ the training data and denote with $[x_i]_1$ and $[x_i]_2$ respectively the first and second component of each data point x_i . The input training data x_i are plotted in the figure below.

We, as usual, shall define the regression matrix

$$X = \begin{bmatrix} x_1^\top \\ x_2^\top \\ \vdots \\ x_m^\top \end{bmatrix}$$

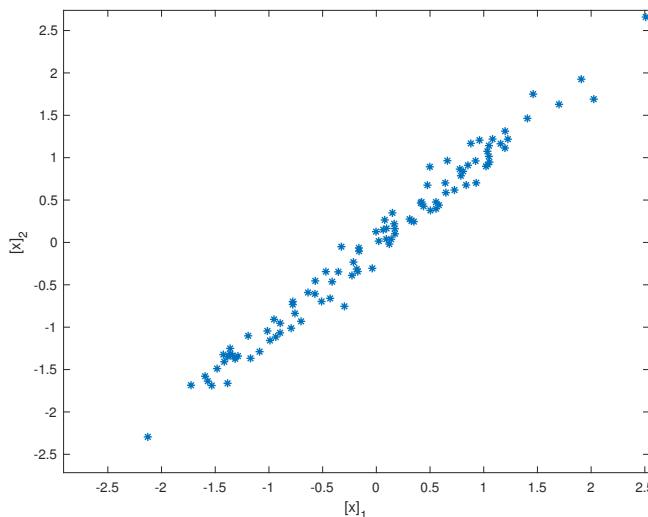
and the output matrix $Y :=$

$$\begin{bmatrix} y_1^\top \\ y_2^\top \\ \vdots \\ y_m^\top \end{bmatrix}$$

1. Do you expect that the least square problem you have to solve to find \hat{w}_s is ill-conditioned? Please motivate your answer.
2. Is it possible to find a reasonable predictor

$$\hat{Y} = X_1 \hat{\alpha} \quad \hat{\alpha} \in \mathbb{R}$$

for a suitably defined column vector X_1 ? If so, how would you find X_1 and $\hat{\alpha}$? Motivate your answer.



Exercise 4 [6 points]

1. Introduce the hard and soft SVM problems for binary classification.
2. With reference to the figure below, the red circles and blue crosses are the data points, the pink line is the optimal classification boundary and the green dashed lines identify the margin.
 - Indicate which are the misclassified points.
 - Indicate which are the support vectors.
 - Indicate for which points the slack variables satisfy the condition $\xi_i > 1$

Motivate your answers.

