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. Define the PAC learning problem
2. Define the realizability and finite model class assumptions and *derive* the expression for lower bound on number of samples needed to achieve a desired accuracy with prescribed probability.

[Solution: Exercise 1]

[Solution: Exercise 1]

Exercise 2 [6 points]

With reference to binary classification:

1. describe the logistic regression model and discuss the shape of the separation boundary;

2. how could logistic regression be extended to obtain more "general" separation boundaries (always for binary classification)?

[Solution: Exercise 2]

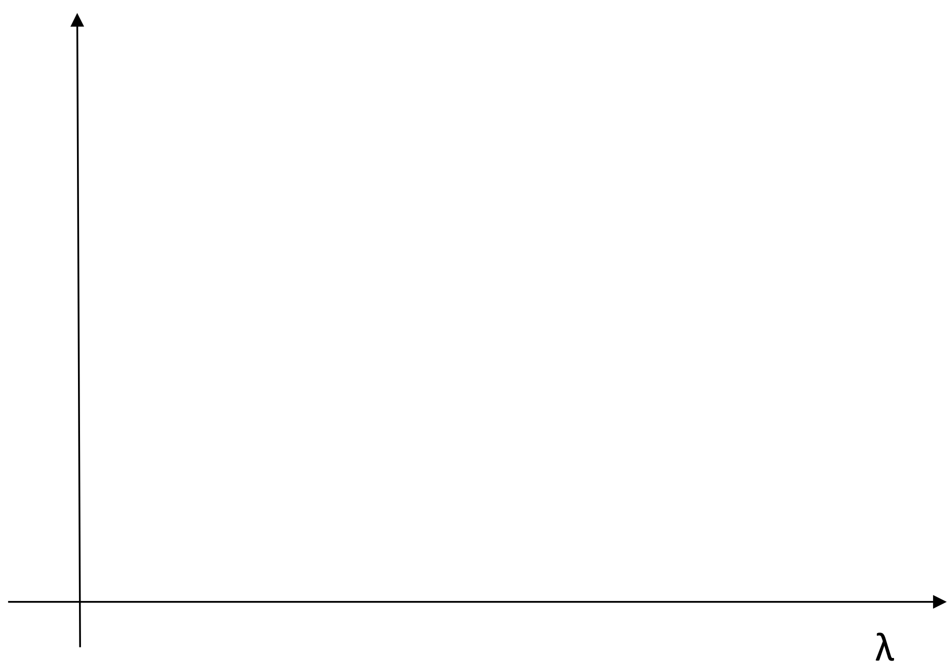
[Solution: Exercise 2]

Exercise 3 [6 points]

Introduce (with proper motivations) ridge regression in the context of linear regression. Assuming the parameter vector that describe the linear model is $w \in \mathbb{R}^d$ and the regularization parameter is $\lambda \geq 0$:

1. provide a formal definition of the true risk $L_{\mathcal{D}}(w)$, of the training error $L_S(w)$ and of the optimization problem to be solved to find $\hat{w}_R(\lambda)$ (the ridge regression estimator of w as a function of λ).

2. Draw in the graph below, the *qualitative* behaviour of $L_s(\hat{w}_R(\lambda))$, $L_{\mathcal{D}}(\hat{w}_R(\lambda))$ and discuss how an "optimal" value of λ can be chosen in practice.



[Solution: Exercise 3]

[Solution: Exercise 3]

Exercise 4 [6 points]

1. Introduce the dimensionality reduction problem in unsupervised learning

2. Assume you are given a dataset $\{x_i\}_{i=1,..,m}$, $x_i \in \mathbb{R}^d$, describe how you can use the singular value decomposition to represent the points x_i in a lower dimensional space

$$x_i \simeq V\alpha_i$$

In particular show how the matrix V and the vectors α_i can be found.

[Solution: Exercise 4]

[Solution: Exercise 4]