# <u>“Machine Learning and Computational Statistics”</u>

## <u>2nd Homework</u>

**<u>Exercise 1</u>** <u>(python code + text)</u>**<u>:</u>**

(a) **Use** the attached python code to read the data from the file “regression.mat”. These data correspond to a data set $X$ of 101 points of the form $X=\{(y_i,x_i), i=1,\ldots,101\}$, where the $x_i$ coordinates of all points (instances of the independent variable) are stored to the vector $X$ and the associated $y_i$ coordinates (instances of the dependent variable) are stored to the vector y.

(b) **Plot** the data and observe that these are modeled by a third degree curve.

(c) **Propose** a model that explains the relationship between $x_i$ ‘s and $y_i$ ‘s.

(d) Based on the above model, **define** a suitable transformation $\varphi(.)$ that maps all $x_i$’s from $R$ to a higher order space, where the problem becomes linear. Denoting by $\mathbf{x}_i'$ the image of $x_i$ in the transformed space, we form a new data set $X'=\{(y_i,\mathbf{x}_i'), i=1,\ldots,101\}$

(e) **Adopting** the linear model assumption in the transformed space and the SSE criterion, estimate the parameters of the model (do not use functions from python libraries to compute the LS solution).

**<u>Exercise 2:</u>**

Consider the following nonlinear model:

$$y = 3x_1^2 + 4x_2^2 + 2x_1x_2 + 7x_1 + 5x_2 + \eta$$

Define a suitable function $\varphi$ that transforms the problem to a space where the problem of estimating the model becomes linear. What is the dimension of the original and the transformed space?

**<u>Exercise 3:</u>**

Consider the following two-class nonlinear classification task:

$$\mathbf{x} = [x_1, x_2, x_3]^T :$$
$$x_1^2 + 3x_2^2 + 6x_3^2 + x_1x_2 + x_2x_3 > (<) 3 \rightarrow \mathbf{x} \in \omega_1(\omega_2)$$

Define a suitable function $\varphi$ that transforms the problem to a space where the problem of estimating the border of the two classes becomes linear. What is the dimension of the original and the transformed space?

**Exercise 4:**

Verify the sum, the product and the Bayes rule for the discrete-valued case, using the relative frequency definition of the probability.

**Exercise 5:**

Consider the matrix $A=\begin{bmatrix} 100 & 10 & 40 & 30 \\ 200 & 20 & 30 & 10 \\ 300 & 30 & 20 & 40 \\ 400 & 40 & 10 & 20 \end{bmatrix}$, where the rows correspond to feature vectors (representing entities) and the columns correspond to the features $x_1$, $x_2$, $x_3$, $x_4$. Defining $\mathbf{x}=[x_1, x_2, x_3, x_4]^T$, compute the **covariance matrix** ($\text{cov}(\mathbf{x})=E[(\mathbf{x}-E[\mathbf{x}])(\mathbf{x}-E[\mathbf{x}])^T]$), the **correlation matrix** ($R_x=E[\mathbf{x}\mathbf{x}^T]$) and the **correlation coefficient matrix** (whose (i,j) element equals to $\frac{\text{cov}(x_i,x_j)}{\sqrt{\text{var}(x_i)}\sqrt{\text{var}(x_j)}}$). Draw your conclusions on the way the features are related to each other.