Course: Numerical Analysis for Machine Learning

Prof. E. Miglio - June 15th 2023

Duration of the exam: 2.5 hours.

Exercise 1

This dataset contains the chemical fingerprint of different samples of glass belonging to different classes. The study of classification of types of glass was motivated by criminological investigation. At the scene of the crime, the glass left can be used as evidence. If it is correctly identified!

Attribute Information:

- Id number: 1 to 214
- RI: refractive index
- Na: Sodium
- Mg: Magnesium
- Al: Aluminum
- Si: Silicon
- K: Potassium
- Ca: Calcium
- Ba: Barium
- Fe: Iron
- Type of glass: (class attribute)
    1. building_windows _float _processed
    2. building_windows_non_float_processed
    3. vehicle_windows_float_processed
    4. vehicle_windows_non_float_processed (none in this database)
    5. containers
    6. tableware
    7. headlamps

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
data = pd.read_csv('glass.data', names=['Id', 'RI', 'Na', 'Mg', 'Al', 'Si', 'K', 'Ca', \
  'Ba', 'Fe','Type'])
data.head()
```

1. Extract the samples of type class 1 (building_windows _float _processed) and 7 (headlamps). How many samples belong to the two groups?
2. Normalize the data so that each feature has zero mean and unitary standard deviation.
3. Perform PCA on the dataset (do not consider the "Id" as a feature!) by means of the SVD decomposition. Then, plot the trend of the singular values σ_k, the cumulate fraction of singular values $\frac{\sum_{i=1}^{k}\sigma_i}{\sum_{i=1}^{q}\sigma_i}$, the fraction of the "explained variance" $\frac{\sum_{i=1}^{k}\sigma_i^2}{\sum_{i=1}^{q}\sigma_i^2}$

4. Compute a matrix containing the principal components associated with the dataset.

5. Generate a scatterplot of the first two principal components of the dataset, grouped by label.

6. Consider now the first principle component. By looking at the components of the first principal direction, do you expect building windows float processed glasses to have a large or small concentration of magnesium and aluminium, respectively? And what about headlamps glasses?

7. Are your predictions confirmed by the original data?

Exercise 2
Consider the following function (where $\mathbf{x} = (x, y)$)

$$f(\mathbf{x}) = (x + 2y - 7)^2 + (2x + y - 5)^2. \tag{1}$$

1. Plot the function.

2. Compute analitically the value $\mathbf{x}_{min}$ where the function attains its minimum value and $f(\mathbf{x}_{min})$.

3. Implement the gradient descent (GD) method

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha \nabla f(\mathbf{x}^k), \tag{2}$$

where α is the learning rate. Use the stopping criterium $E < \epsilon$ where E is the absolute value of the difference of the last two functional values *i.e.*

$$E = |f(\mathbf{x}^{k+1}) - f(\mathbf{x}^k)|, \tag{3}$$

and ϵ is the required tolerance. Moreover set the maximum number of iterations to 200.

4. Apply the GD method with $\epsilon = 10^{-3}$, $\alpha = 0.01$ and 0.1. Plot the convergence history and comment the results.

5. Implement the GD with momentum (GDM) *i.e.*

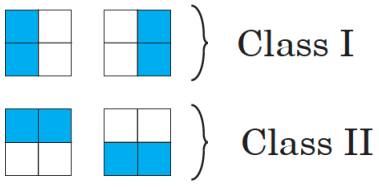
$$\mathbf{z}^{k+1} = \beta \mathbf{z}^k + \nabla f(\mathbf{x}^k), \tag{4}$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha \mathbf{z}^{k+1}, \tag{5}$$

where α is the learning rate and β is the momentum coefficient. Use the same stopping criterium used in the GD method.

6. Apply the GDM method with $\epsilon = 10^{-3}$, $\alpha = 0.01, 0.1$ and $\beta = 0.5, 0.9$. Plot the convergence history and comment the results.

Exercise 3
Consider the two classes of patterns that are shown in the following figure where Class I represents vertical lines and Class II represents horizontal lines.



1. Are these categories linearly separable ?

2. Design a multilayer network to distinguish these categories.