Course: Numerical Analysis for Machine Learning
Prof. E. Miglio - January 16th 2024
Duration of the exam: 2.5 hours.

Exercise 1
Consider the following dataset

```
from sklearn.datasets import fetch_olivetti_faces

olivetti = fetch_olivetti_faces()
imgs = olivetti.images
labels = olivetti.target

X = imgs.reshape((400, 4096)).transpose()
```

1. Visualize 10 randomly selected pictures with the corresponding labels.
2. Compute and visualize the average of the images.
3. Perform SVD by first setting the attribute `full_matrices = True` and then `full_matrices = False`. Comment the results.
4. Plot the trend of the singular values and the fraction of "explained variance".
5. Implement a function computing the randomized SVD of rank k for a generic matrix.
6. Set $k = 1, 5, 10, 50, 100$ and plot the approximated singular values together with the exact ones.
7. Use PCA to perform dimensionality reduction on the dataset of images for rank $k = 1, 5, 10, 50, 100$ by means of exact SVD. Compute the reconstruction error and plot it as a function of k . Comment the results.
8. Visualize the first 30 principal axes.
9. Compute the first two principal components related to the subset of images corresponding to labels = 0, 39.
10. Create a scatterplot for the first 2 principal components of the subset of images grouped by label. Comment what you see.

Exercise 2
Consider the Ridge regression.

1. Write the loss function for the Ridge regression.
2. Derive the expression of the solution $\mathbf{w}^*$ (weight vector) for the Ridge regression.
3. Consider the dataset

   ```
   np.random.seed(55)
   x = np.arange(np.pi,3*np.pi,0.1)
   y = np.sin(x) + np.random.normal(0,0.1,len(x))
   ```

   and the following values of λ (regularization parameter) $\lambda = 0, 10^{-32}, 10^{-16}, 10^{-8}, 10^{-2}, 1, 16, 32, 1024$. Compute the values of $\mathbf{w}^*$ and plot the solution of the Ridge regression for the above mentioned values of λ. Comment the obtained results.

Exercise 3

Consider the quadratic function

$$J(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{b}^T\mathbf{x}, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ is SPD and $\mathbf{b} \in \mathbb{R}^n$.

1. Compute the gradient and the Hessian of J.
2. Verify that J is strictly convex and find the unique global minimum of J.
3. Let $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ a direction s.t. $\nabla J(\mathbf{x})^T\mathbf{q} < 0$. Compute analytically the step length α that solve the following exact line-search problem

$$\min_{\alpha>0} J(\mathbf{x} + \alpha\mathbf{q}). \tag{2}$$