Mathematical Underpinnings Lab 11 29.05.2024

Task 1 (sparse PCA)

a) Load dataset from `sklearn package`:
`faces, _ = datasets.fetch_olivetti_faces(return_X_y=True, shuffle=False)`. Draw all the faces. Standardize the data.

b) Fit Sparse PCA on the dataset with 10 principal components. You can control the sparsity of the solution through the parameter `alpha`. Fit the model for `alpha`=1 and `alpha`=3 (start with `alpha`=3 as the training may take some time, possibly around 5 minutes, and if it is that long you can omit `alpha`=1). Plot all the principal components and compare the number of non-zero values.

c) Based on the sparse PCA with `alpha`=3, how can you isolate the eyes in the picture (using which principal component)? How can you extract the whole lower or upper part of the face?

d) Which pair of principal components is the most predictive in determining whether the person wears glasses?

e) NMF (Non-negative Matrix Factorization) is another method for representing data. As the name suggests, the components cannot be negative. Fit NMF to the dataset (just scaled not centered), plot all the principal components, and then answer the questions from c) and d).

f) Fit PCA. Compare the results with SPCA and NMF.