Barbara Hammer, Philip Kenneweg Universität Bielefeld

Modern Data Analysis (WiSe 2023/2024)
2. Sheet

Start: Wednesday, 25.10.2023.
End: The worksheets should be solved using Python, in groups of 2-3 people and will be presented in the Tutorials.
Discussion: Monday, 04.11.2023 in den Tutorien.

Information

The worksheets and necessary toolboxes will be made available in the Lernraum "392246 Modern Data Analysis (V)". Worksheets will usually be released every two weeks on Wednesday, and discussed during the exercises on Monday two weeks later. In order to successfully finish the course, 60% of the available points have to be obtained and each participant has to present his/her results at least once. The Monday in between the release and discussion of the sheet will be used to discuss the implementation of the various algorithms presented in the lecture.

Exercise 1:

(*10 Points*)

(a) *(6 Pts.)* The algorithm Alternating Least Squares was introduced in the lecture as one way to perform NMF. Visually inspect the Olivetti Faces Dataset included in the materials. Implement the ALS algorithm and apply it to the dataset using 10 components. Analyze the results, e.g. by visualizing the progress of the reconstruc- tion error, comparing different parameter choices, visually inspecting the primitives and comparing the reconstructed images to the originals.

(b) *(4 Pts.)* Compare your results to an alternative NMF implementation with sparsity constraints. For this purpose you can use the sklearn NMF with alpha_W = 1 and l1_ratio = 1. Discuss advantages and drawbacks of sparsity.