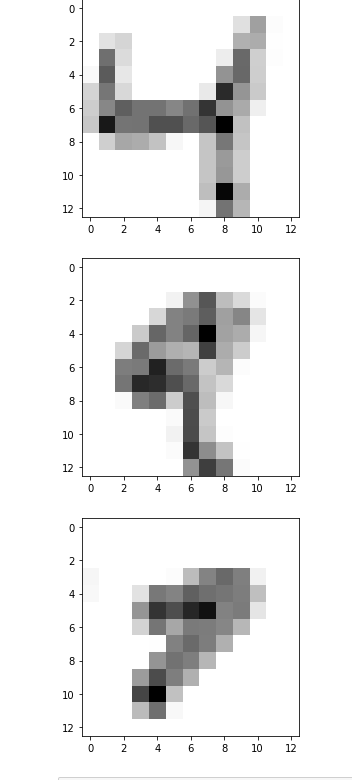
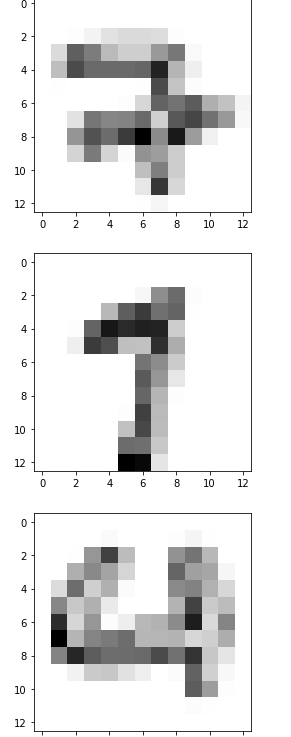
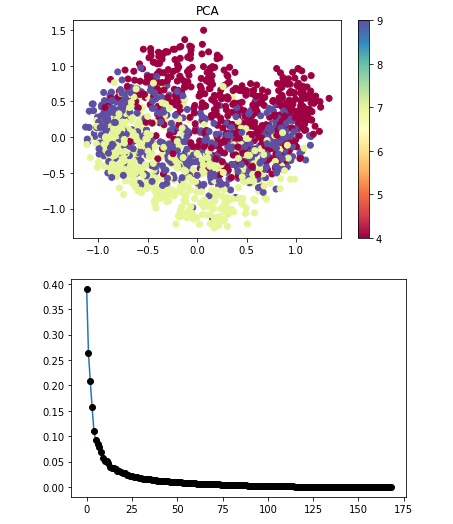
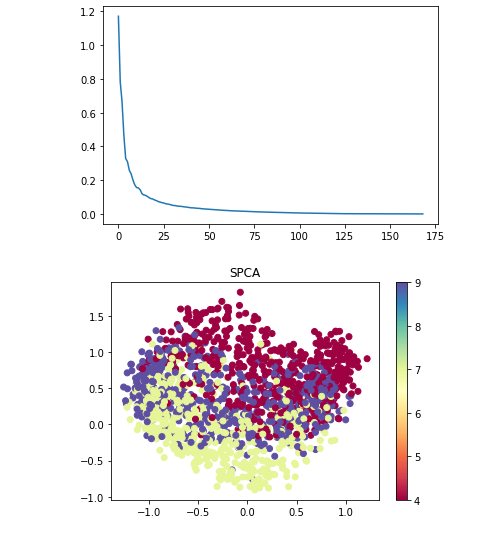
We have a dataset (it is a specially prepared part of MNIST), with 1508 vectors corresponding to 13x13 images of digits 4,7 and 9, so the size of the data matrix is 1508x169. The digits look like this:

The dataset is located at <https://github.com/Mirkes/DAPCA/raw/main/datasets/MNIST_M/MNIST_ex479.mat>

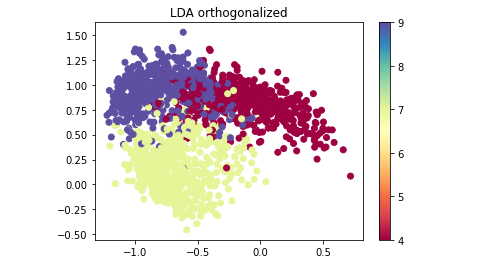
Here is the PCA of this dataset:

We apply Supervised PCA (e.g., with alpha=1 meaning points of the same class will be contracted with the same force and the point of different classes will repulse).

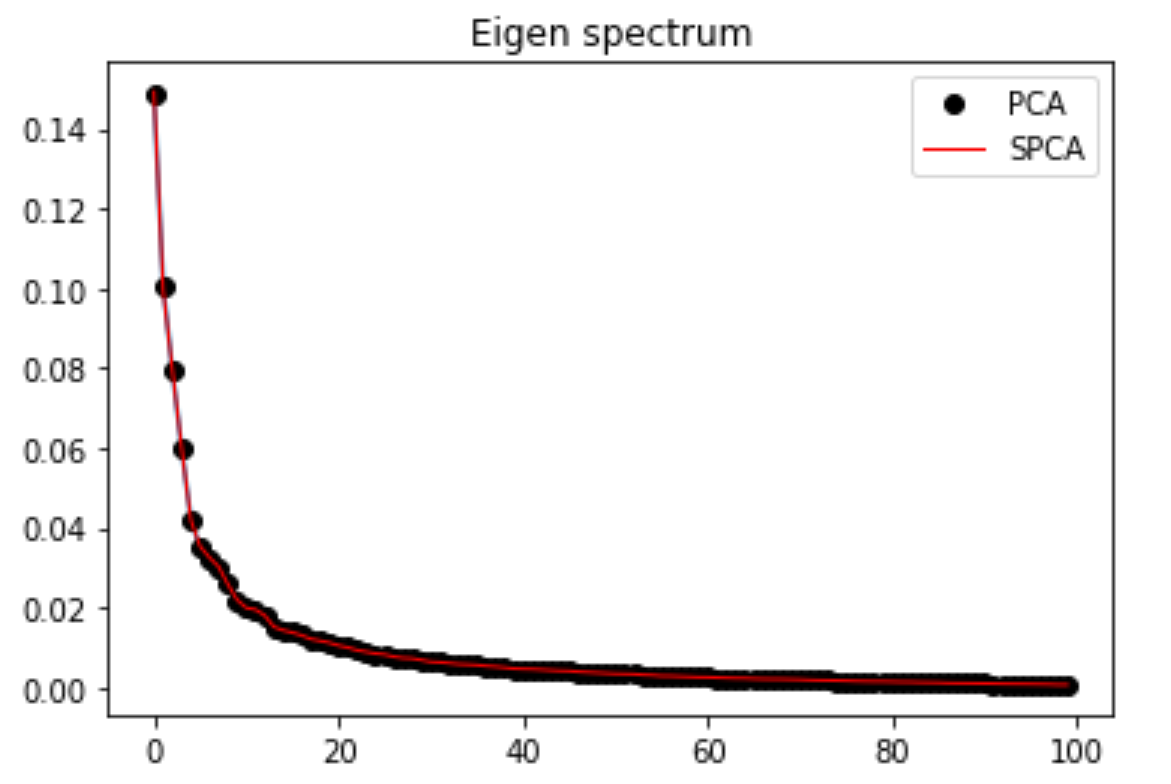


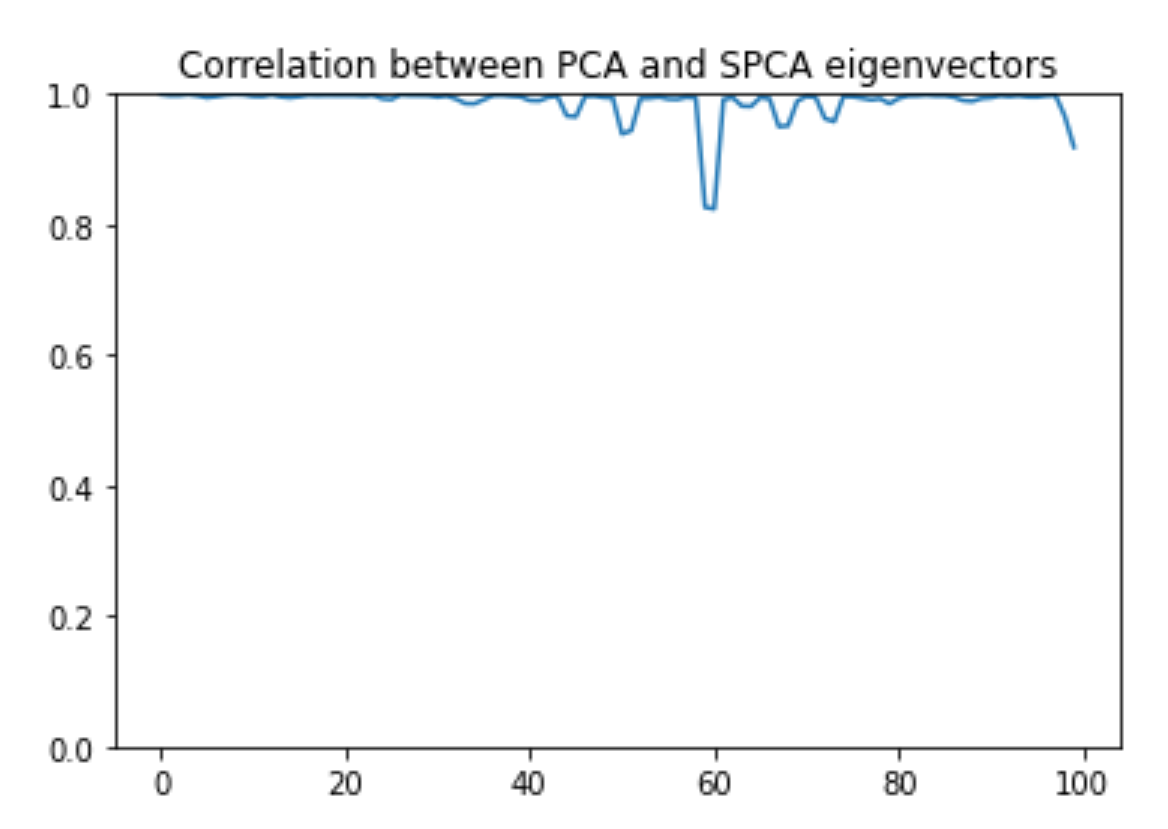
The result looks identical to PCA. No other choice of alpha can affect this result: it always looks like PCA.

However, we know that exists a LINEAR orthogonal projection where these three classes are well-separated. It can be obtained by simplest orthogonalization of the 2 vectors connecting two pairs of class centroids:

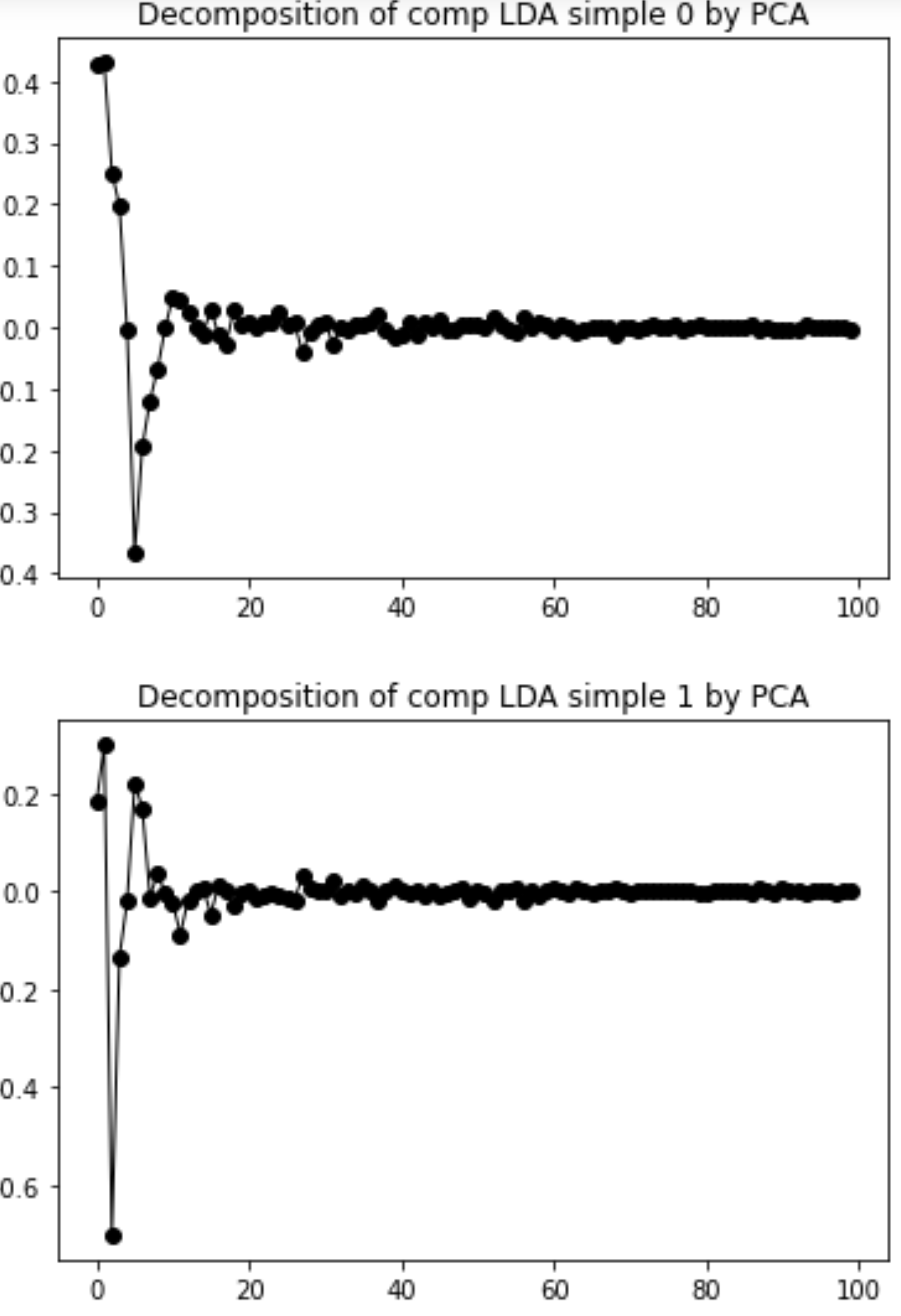


Here is comparison between eigenvalues and eigenvectors extracted by PCA and SPCA, for the first 100 components:





Here is how the 2 vectors connecting 3 class centroids (I call this approach ‘LDA simple’) are decomposed in the PCA basis:



**Question: why we can not improve the PCA projection in terms of class separation using Supervised PCA?**