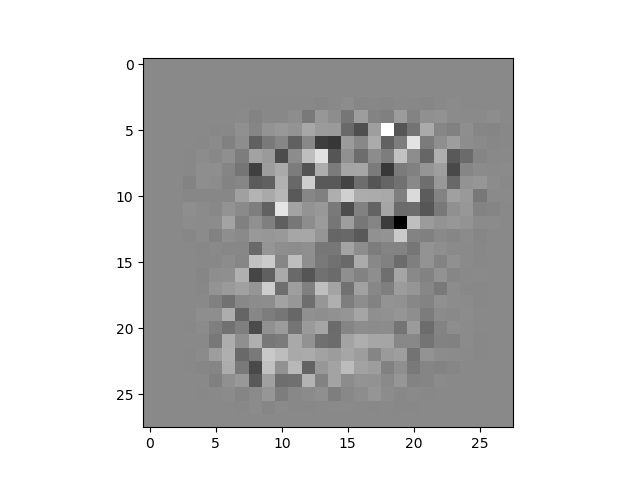
Assignment #6 – Programming Part

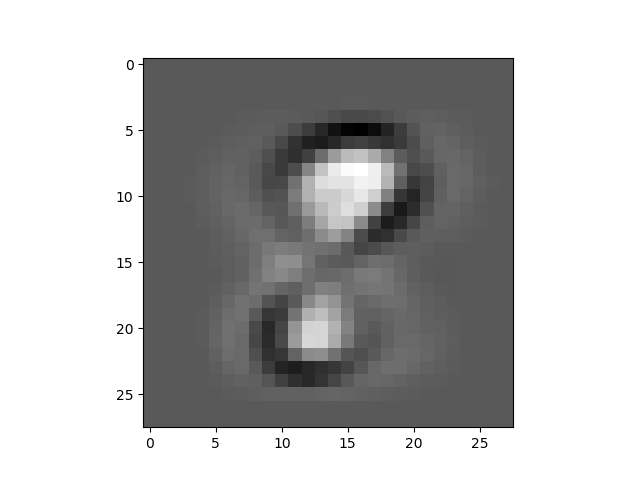
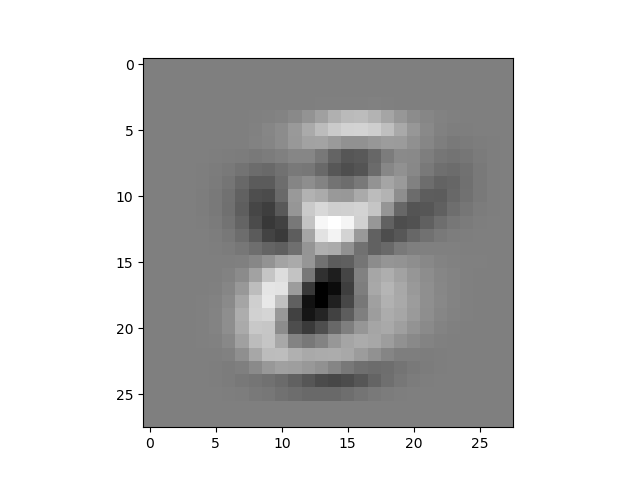
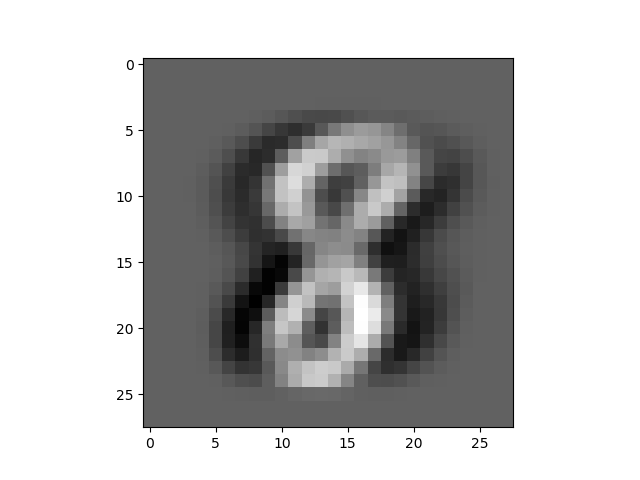
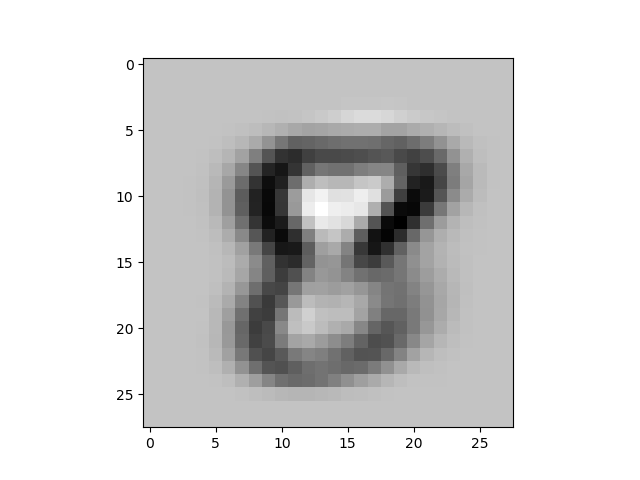
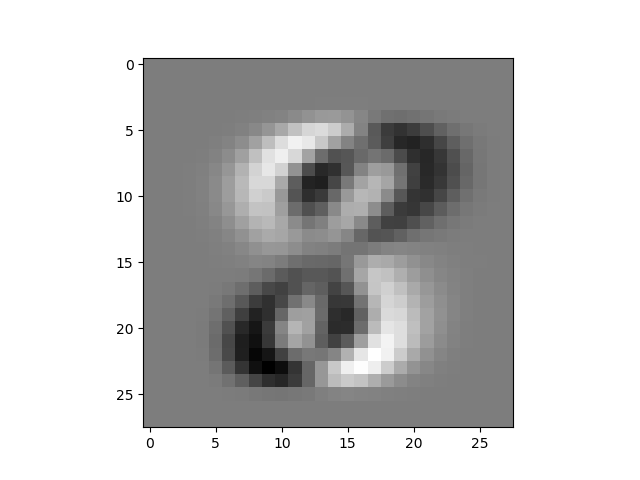
**Code Location:** /specific/a/home/cc/students/cs/itaysofer/IntroToML/Ex6

**Question 1**

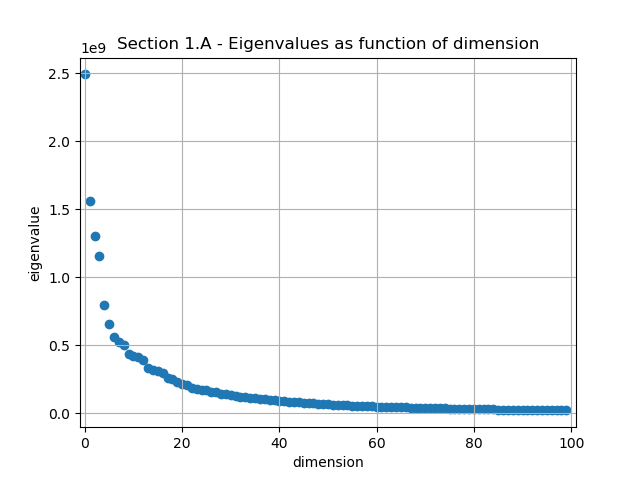
1. Mean Image:



First 5 eigenvectors:



Eigenvalues (in decreasing order) as a function of dimension:



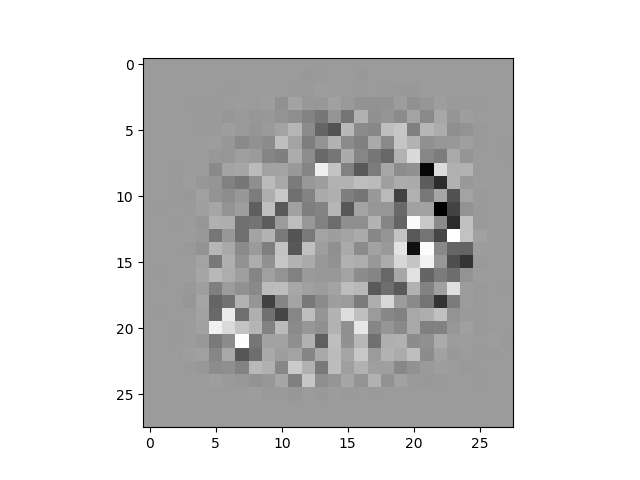
Explanations:

The first eigenvector looks like it determines if the digit is leaning to the left or right.

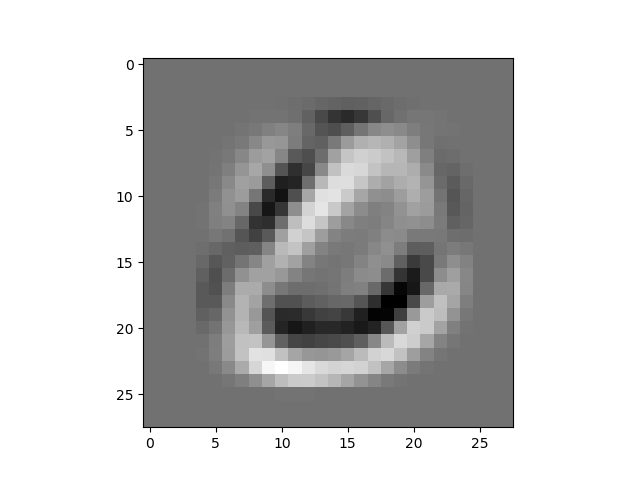
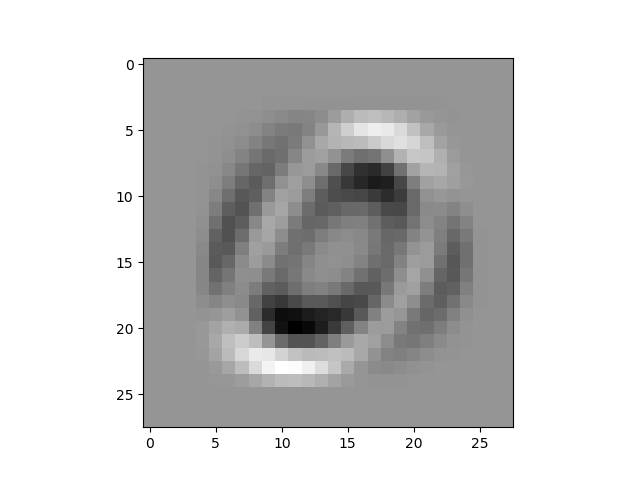
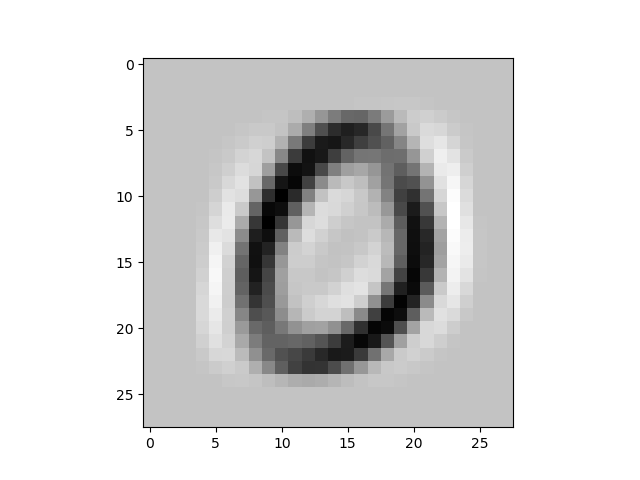
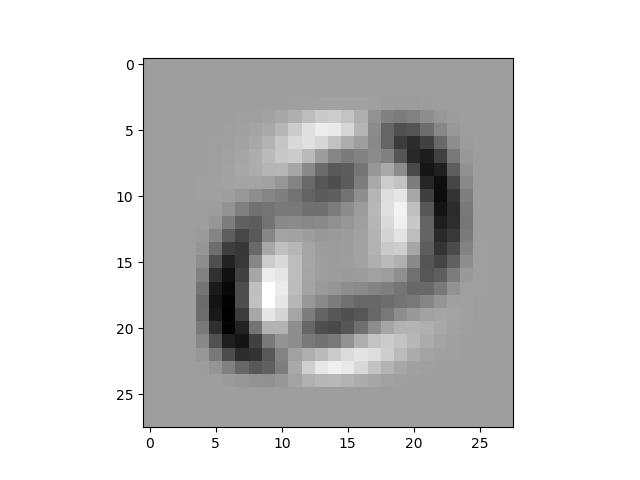
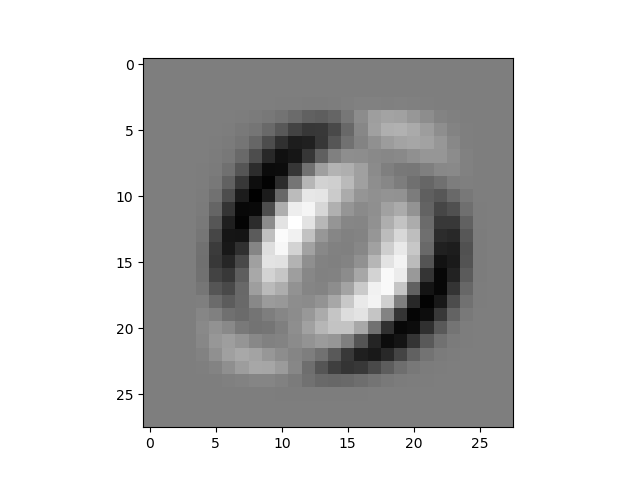
The second eigenvector looks like it determines if the digit has narrow or wide holes.

The third eigenvector looks like it determines the if the digit is thin or wide.

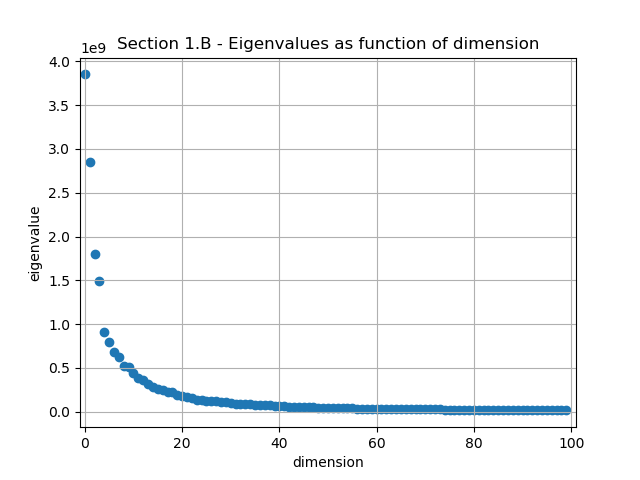
1. Mean Image:



First 5 eigenvectors:



Eigenvalues (in decreasing order) as a function of dimension:



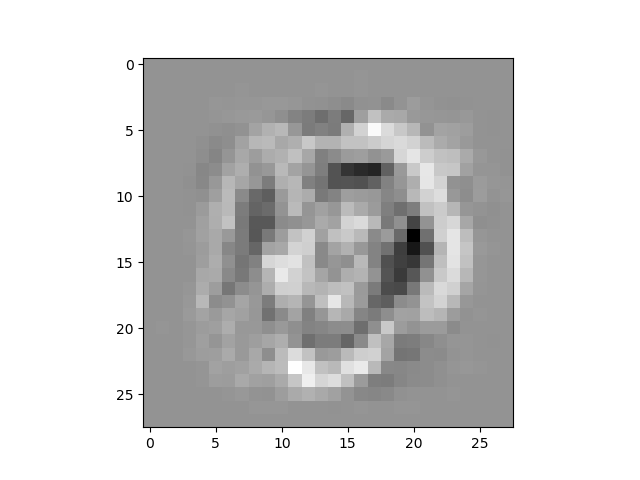
Explanations:

The first and second eigenvector looks like it determines if the digit is leaning to the right.

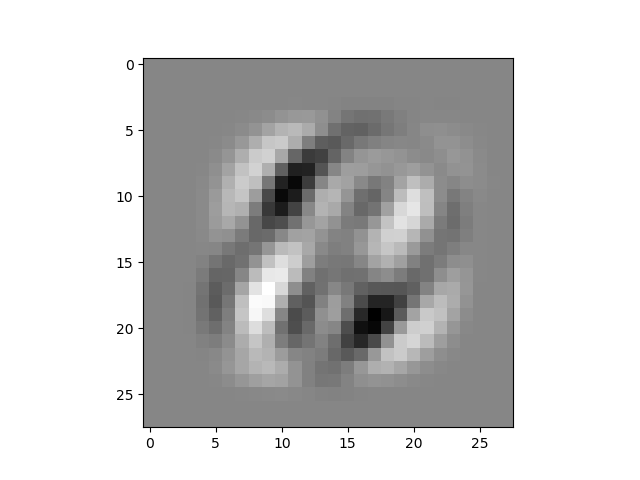
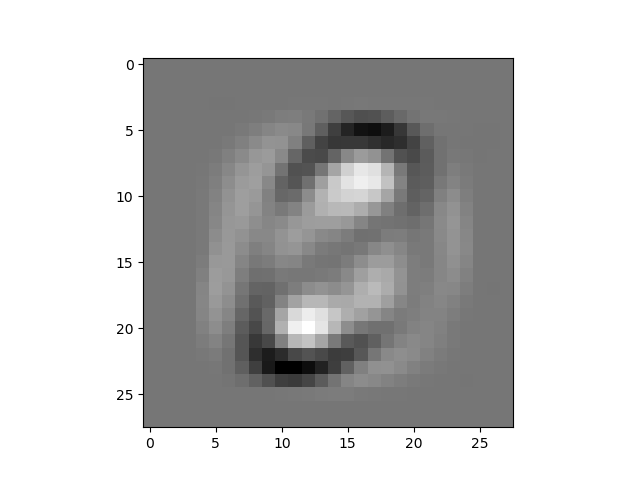
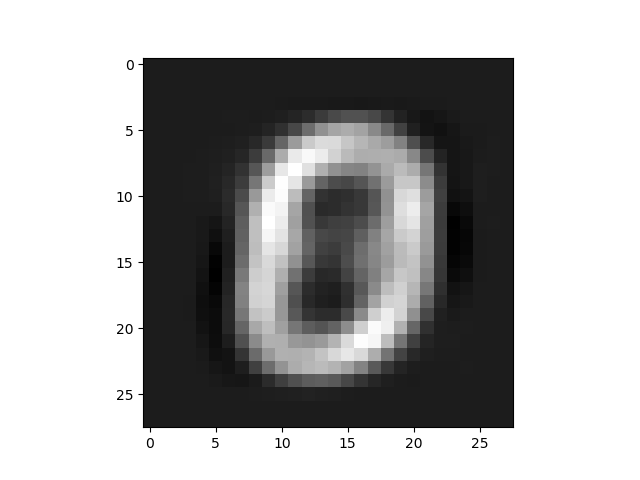
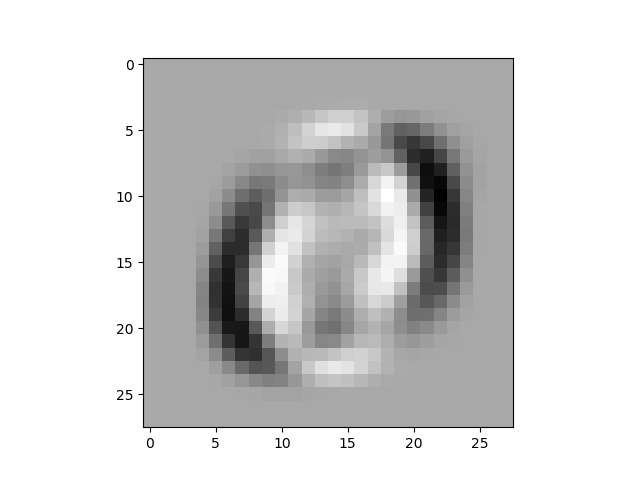
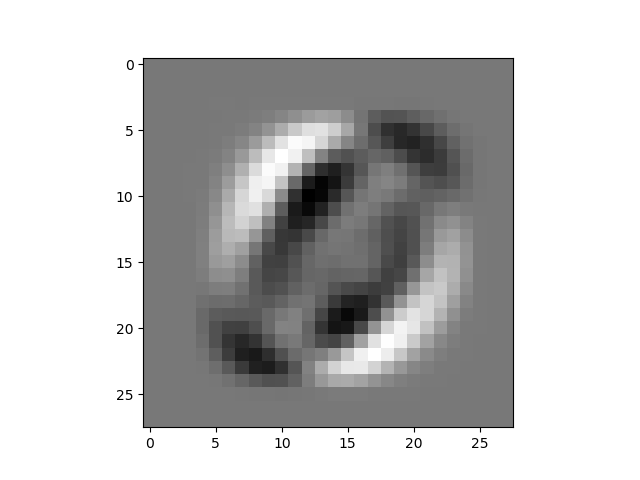
The third eigenvector looks like it determines the size of the digit.

1. We used only ~7000 samples of digits 0 and 8 for efficiency.

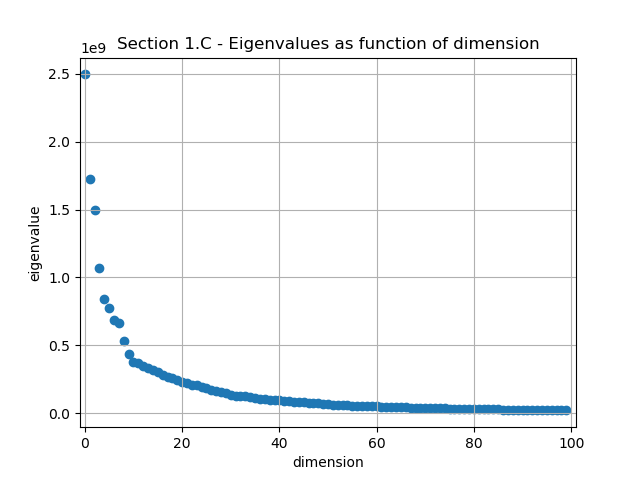
Mean image:



First 5 eigenvectors:



Eigenvalues (in decreasing order) as a function of dimension:



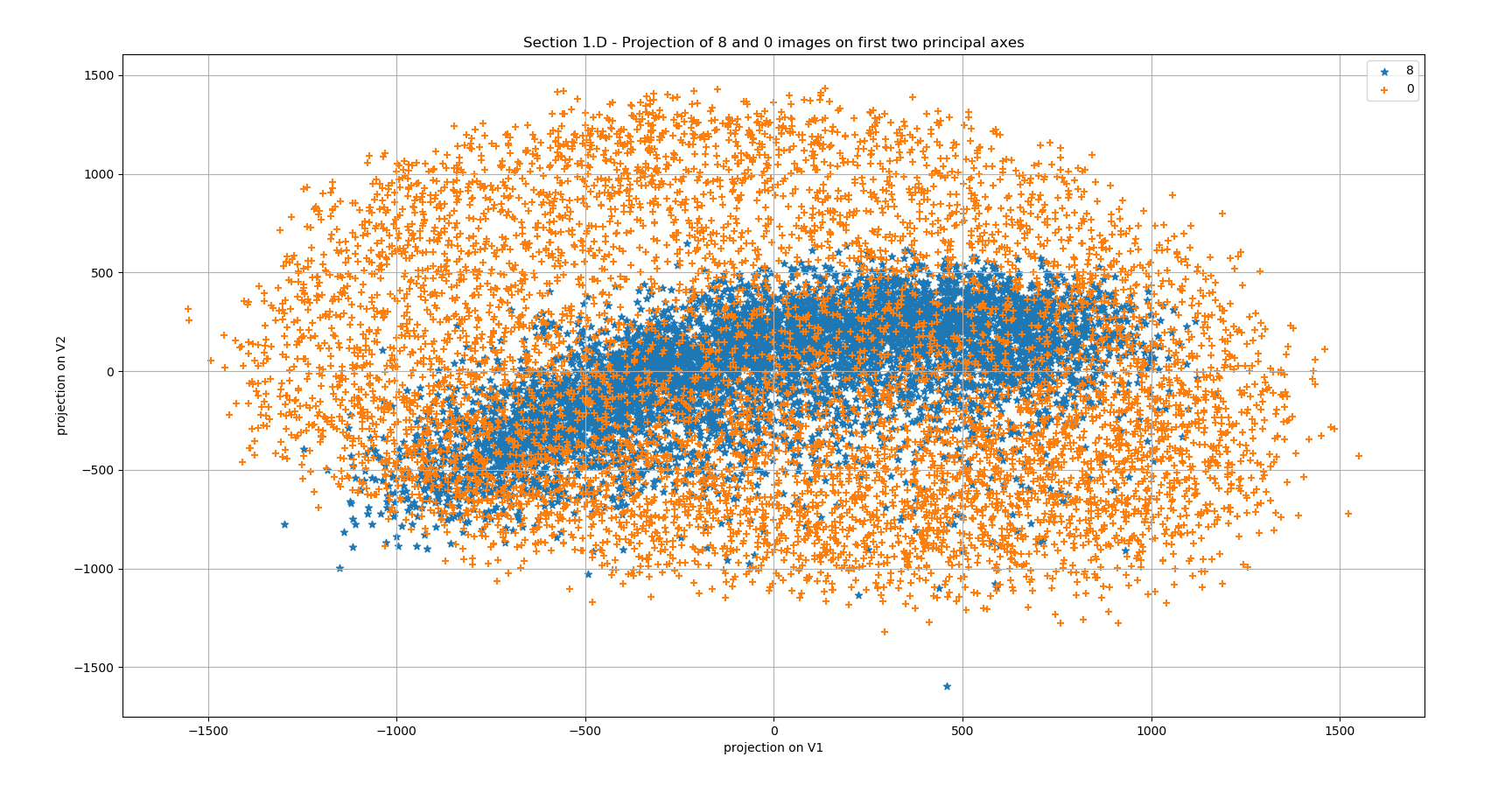
Explanations:

The mean image doesn’t have a general shape of an 8 or a zero, but a mixed shape of both.

The eigenvectors looks like a combination of the digits 0 and 8. This is probably because now the subspace which all the data points lay on is a combined space of 8 and 0.

There is a difference between the magnitude of the eigenvalues. In this case it is closer to the values in section a, and they are both lower than in section b. As we saw in class, as the sum of eigenvalues increases, the objective of PCA decreases. This means that the digits 0 alone are placed closer to a subspace of size 5 than the digits 8 or digits 8 and 0 combined (we looked at the sum of the first 5 eigenvalues).

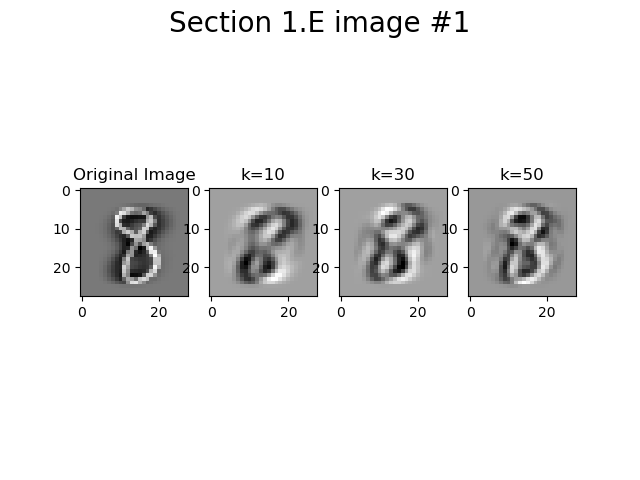
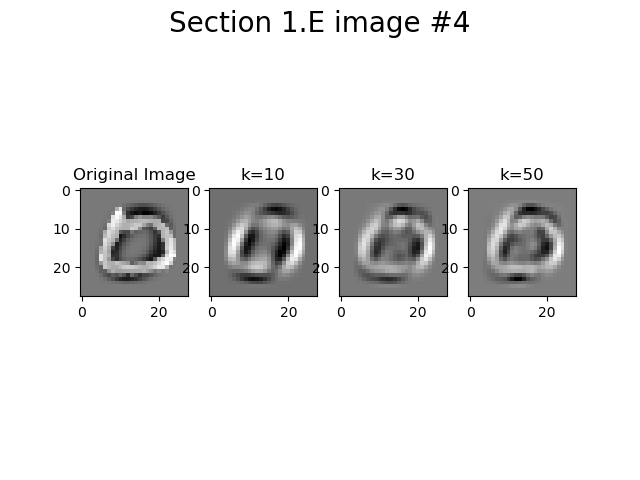
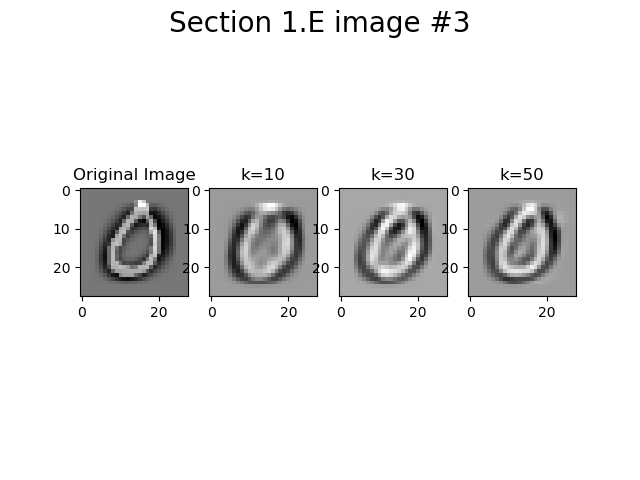
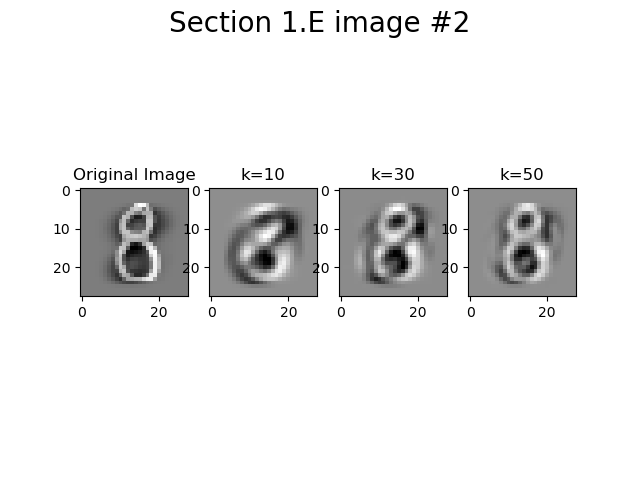
1. The plot:



Explanation:

The 8 digits are more concentrated than the 0 digits when projecting to the 2D subspace. This implies that reconstructing the 8 digits with 2 largest eigenvectors will result in a better result than constructing the 0 digits.

1. The images:

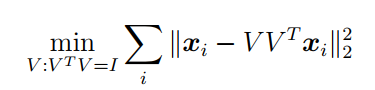
 

As we can see, reconstructing with more principal components gives a better result.

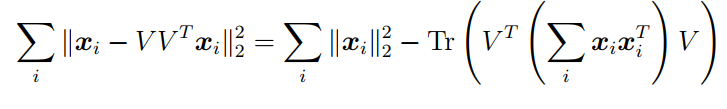
This implies that the images doesn’t lay on a subspace of dimention lower than 50.

Moreover, the reconstruction is still not very good with k=50 which means that the optimal subspace is of higher dimention.

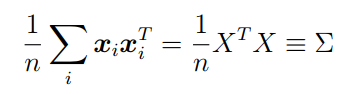
1. As we saw in class the PCA objective is:



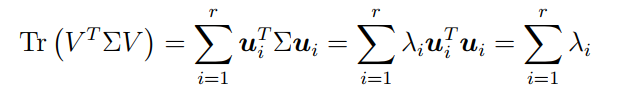
It is equivalent to:



Where:



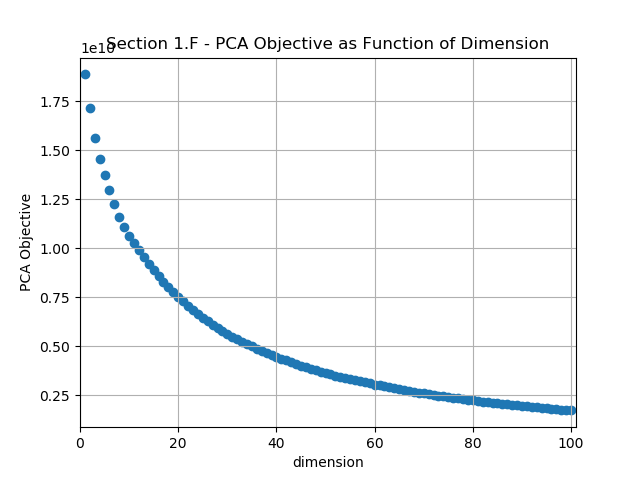
And:



Therefore we can compute the objective by:

(We used the non-scaled covariance matrix – X^T\*X)

The plot:



As the graph implies, using more principal components results in a lower PCA observation.