CS 215 ASSIGNMENT

Tathagat Verma - 180050111
Neel Aryan Gupta - 180050067

Date:

Question - 4

Eigenvalues plot:
For each digit, we see that the graph of eigenvalues decreases rapidly.
From this, we infer that only some of the eigenvalues (consequently, eigenvectors) have major contribution in forming the dataset ie. we can effectively reproduce the dataset using lower dimension variables.
We see that eigenvalues with index > 100 are almost zero relative to first 100 eigenvalues.
Out of a total of 784 (28*28) only 100 of them are significant.
This happens because in real-life datasets (such as this one), the data points/vectors are not uncorrelated. There is a certain 'pattern' observed in the data & PCA helps us quantify it.

The three figures obtained for each digit are quite similar except that a certain 'pattern' is changed between the three. This 'pattern' is characterised by the principal mode of variation. For example, in case of the digit 1, the principal mode of variation characterised the rotation of a '1' about the center, which would imply that most people write 1 as a single line segment / rod in a slanted / rotated way. And the data mostly contains 1's with different rotations.

Scanned with CamScanner