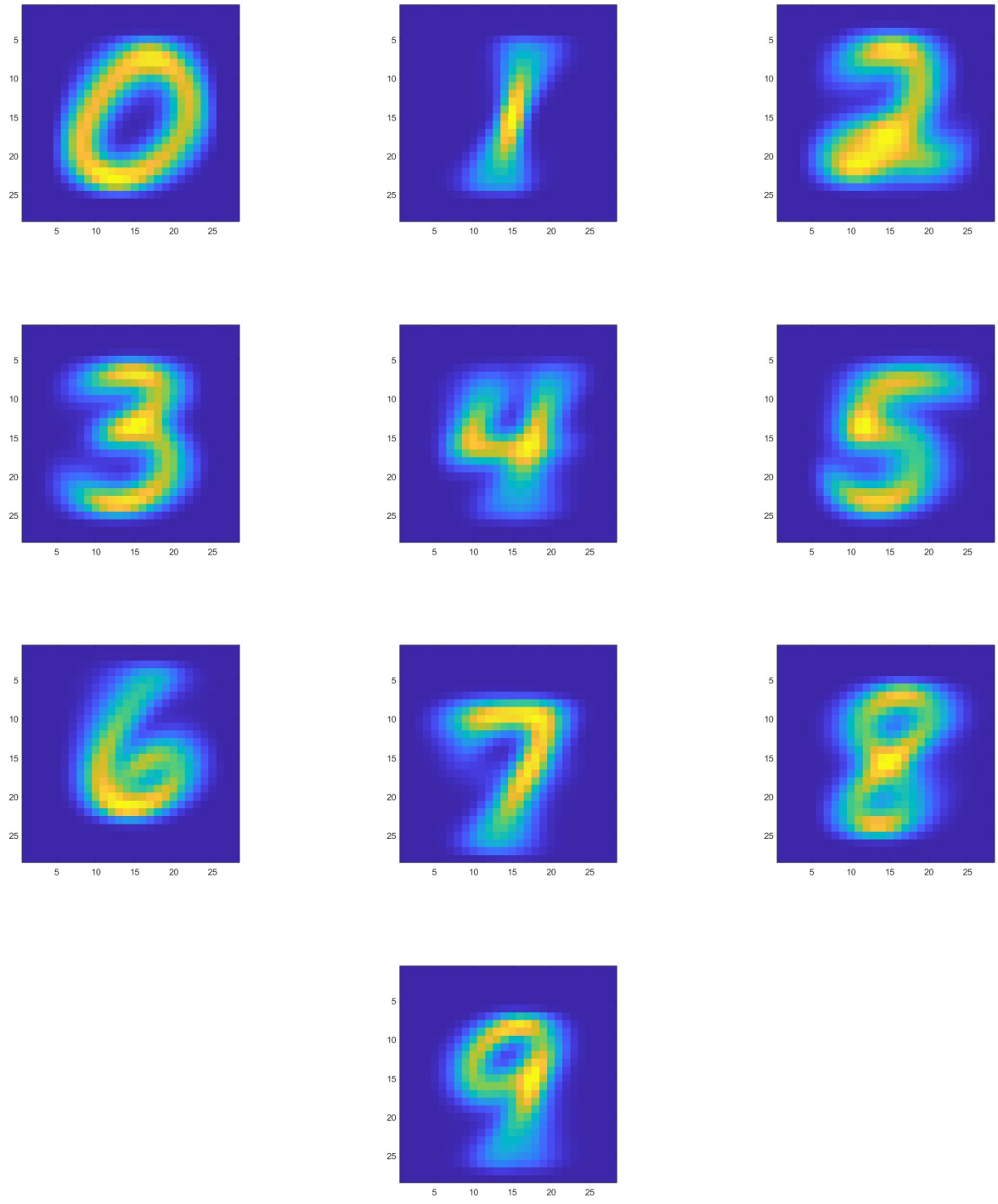
CS215 Assignment-2

210050115 Patil Vipul Sudhir

210050119 Hari Prakash Reddy

1 Principal Component Analysis (PCA)

1.1 Mean of digits



Instructions to run the code:
Run meanOfDigits.m from code folder of Q4

1.2 Covariance of Digits

Wrote code for finding covariance of all digits.

Instructions to run the code:
Run covarianceOfDigits.m from the code folder of Q4, for printing covariance of digits.

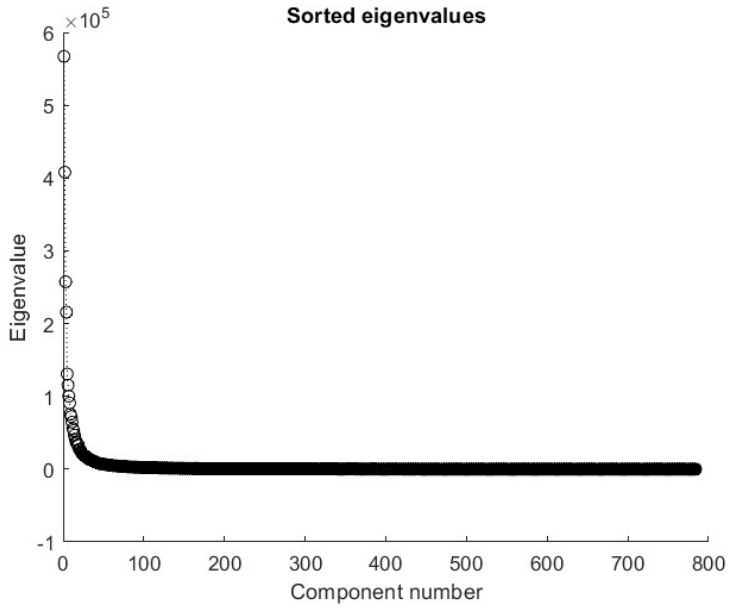
1.3 Principal mode of variation in digits

Digit	Eigenvalues
0	5.6716 x10^5
1	5.1199 x10^5
2	3.9687 x10^5
3	3.6442 x10^5
4	3.1714 x10^5
5	5.1741 x10^5
6	4.8539 x10^5
7	3.9185 x10^5
8	3.6633 x10^5
9	4.0329 x10^5

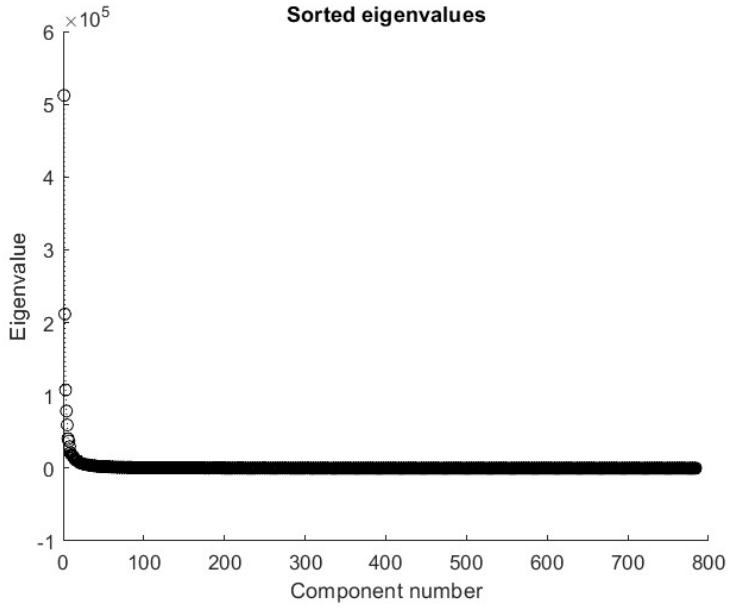
Instructions to run the code:
Run eigenvectorOfDigits.m from code folder of Q4

1.4 Significant eigenvalues

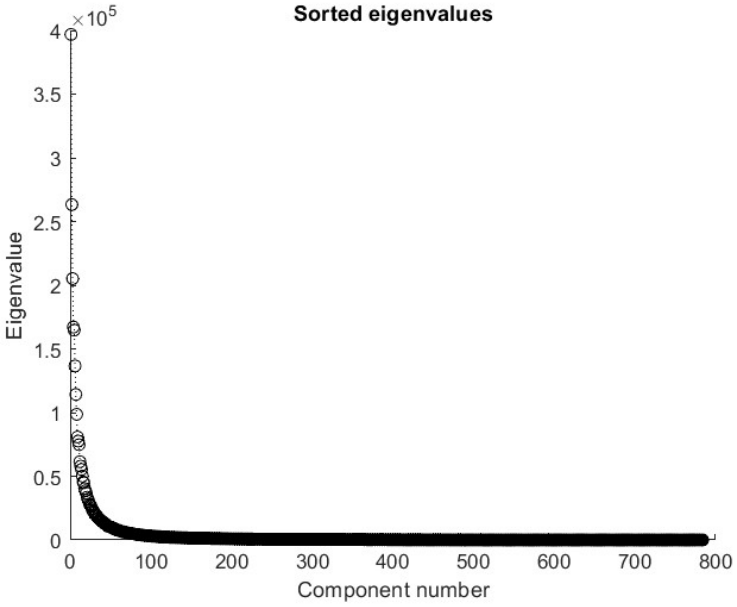
N=0



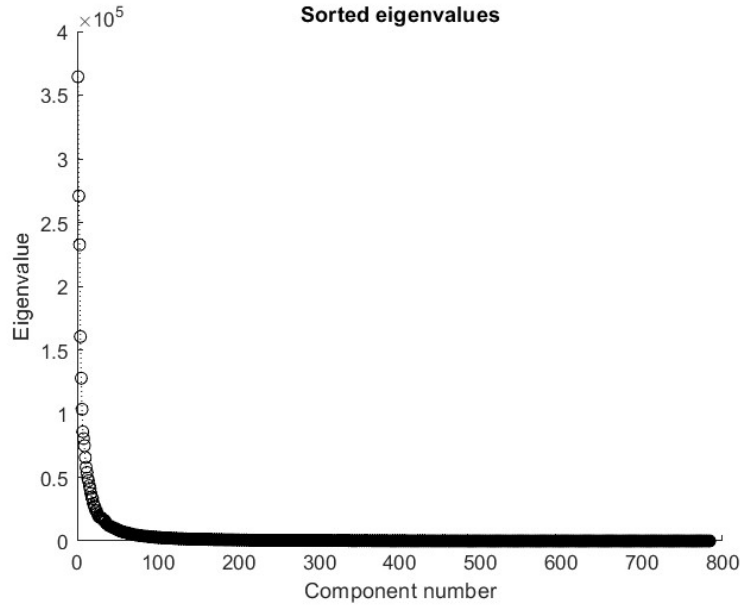
N=1


Sorted eigenvalues
×10^5
6
5
4
3
2
1
0
-1
Eigenvalue
0
100
200
300
400
500
600
700
800
Component number

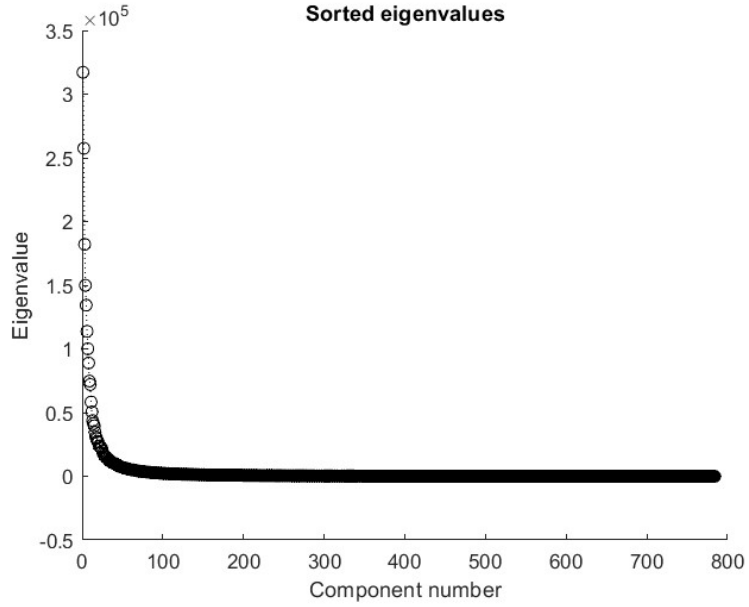
N=2


Sorted eigenvalues
×10^5
4
3.5
3
2.5
2
1.5
1
0.5
0
Eigenvalue
0
100
200
300
400
500
600
700
800
Component number

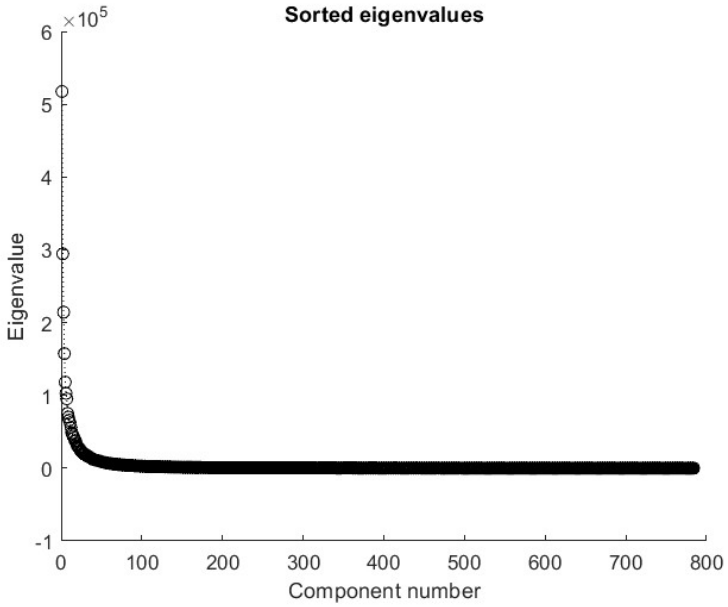
N=3


Sorted eigenvalues
×10^5
4
3.5
3
2.5
2
1.5
1
0.5
0
Eigenvalue
0
100
200
300
400
500
600
700
800
Component number

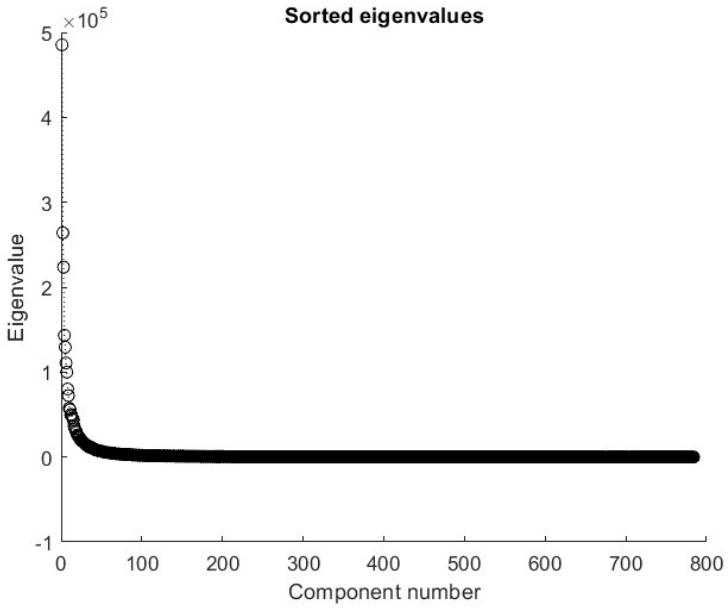
N=4


Sorted eigenvalues
×10^5
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
Eigenvalue
0
100
200
300
400
500
600
700
800
Component number

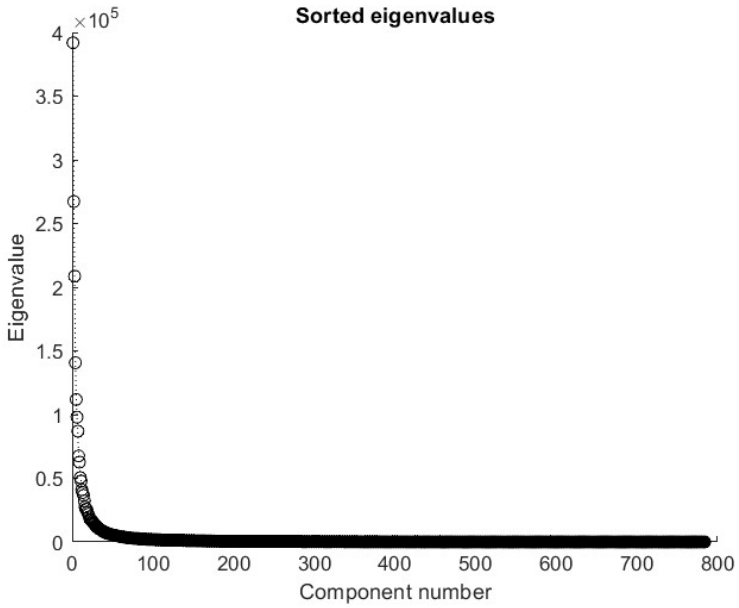
N=5


Sorted eigenvalues
×10^5
6
5
4
3
2
1
0
-1
Eigenvalue
0
100
200
300
400
500
600
700
800
Component number

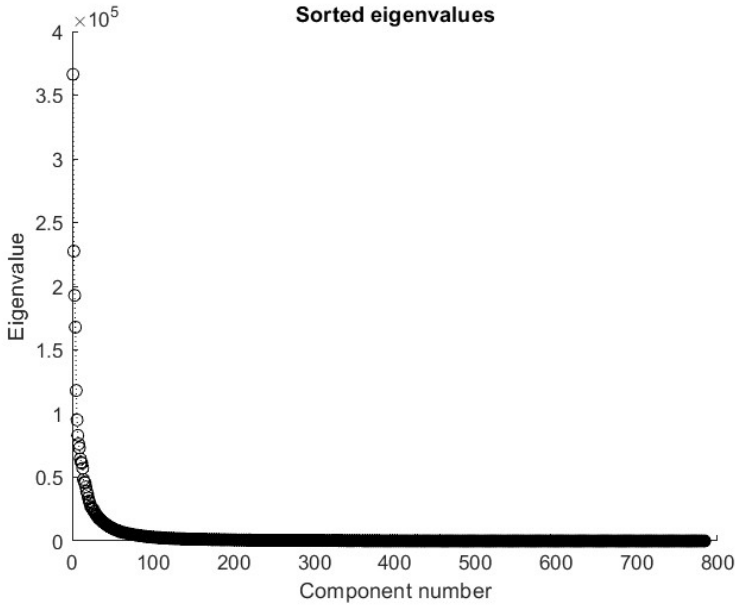
N=6


Sorted eigenvalues
×10^5
5
4
3
2
1
0
-1
Eigenvalue
0
100
200
300
400
500
600
700
800
Component number

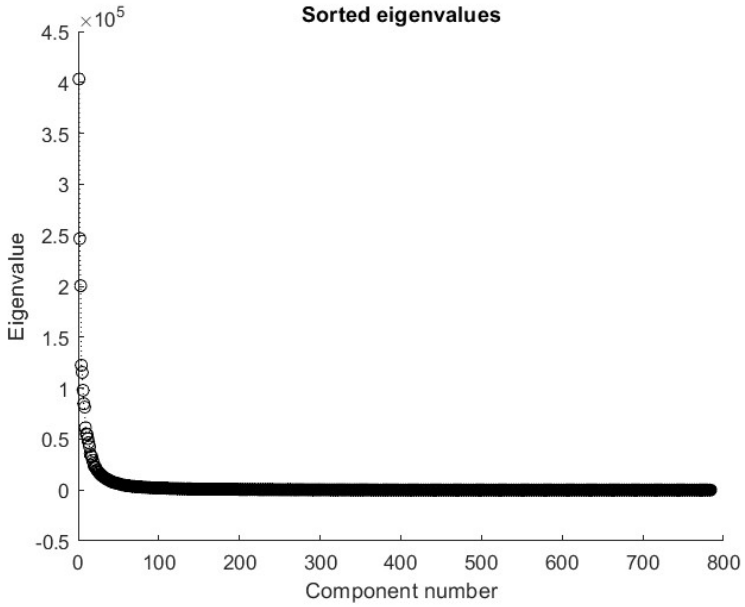
N=7


Sorted eigenvalues
×10^5
Eigenvalue
Component number
0
0.5
1
1.5
2
2.5
3
3.5
4
0
100
200
300
400
500
600
700
800

N=8


Sorted eigenvalues
×10^5
Eigenvalue
Component number
0
0.5
1
1.5
2
2.5
3
3.5
4
0
100
200
300
400
500
600
700
800

N=9



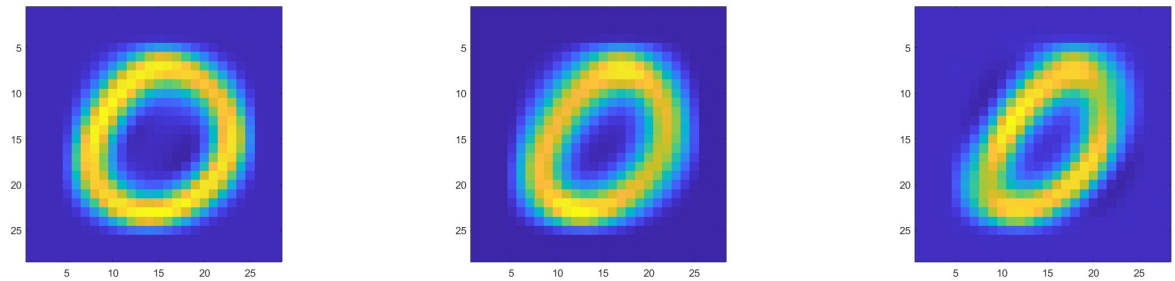
1.4.1 Comments and justifications

After plotting each digit's eigenvalues, only some are higher than others, and most of the eigenvalues are tending to be 0 according to the graph drawn.
The number of Significant eigenvalues for each digit is about 80-90(considering the threshold to be around 3000 as the order of eigenvalues is high), and the significant nodes of variation are far less than 784 as 784 dimensions are redundant to be used for containing a simple image such as digits, but can be used for getting more precision on images.
Instructions to run the code:

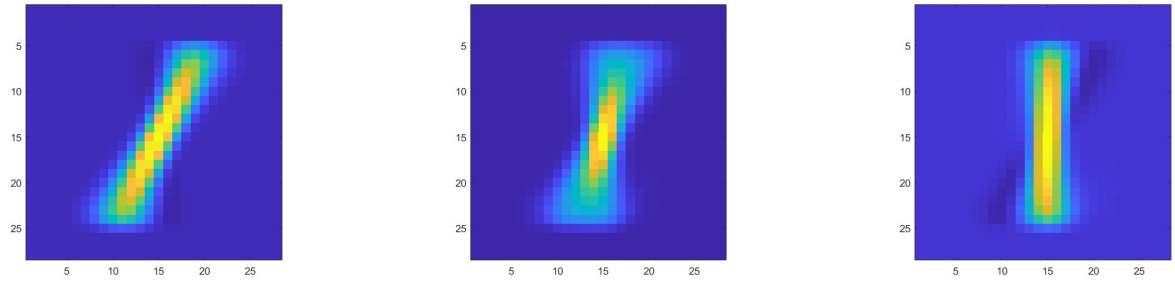
Run significantEigenvalues.m from the code folder of Q4 with an argument, for which number the graph of eigenvalues is to be plotted.

1.5 Images reconstructed using mean and Eigen vectors

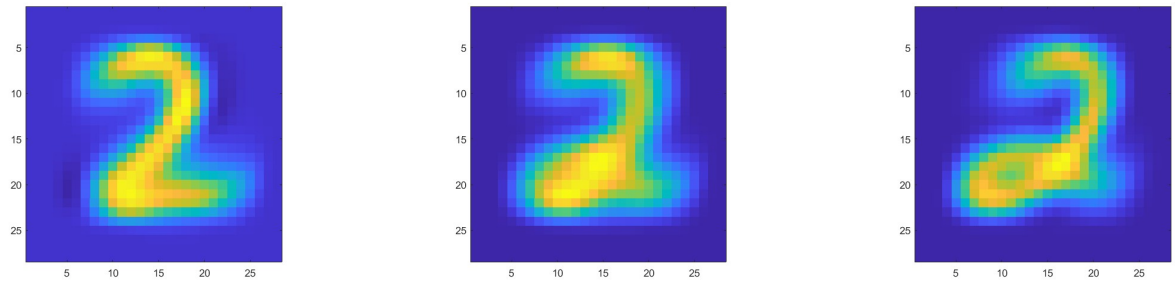
N=0



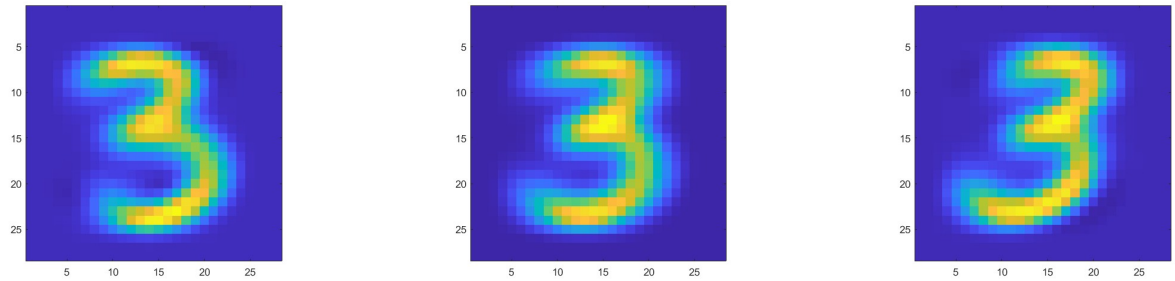
N=1



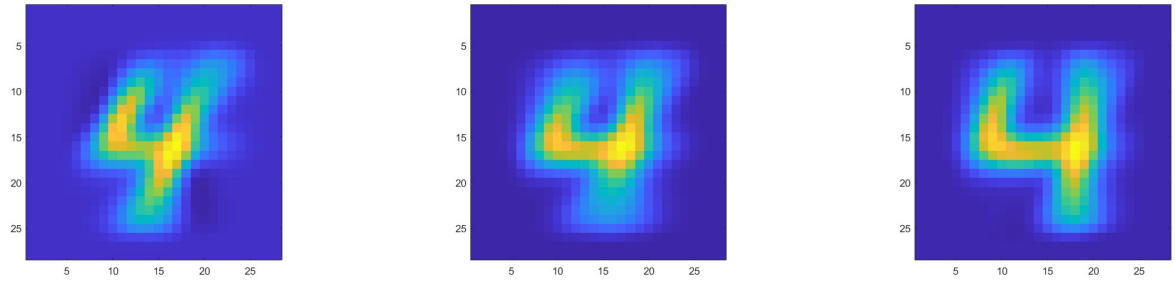
N=2



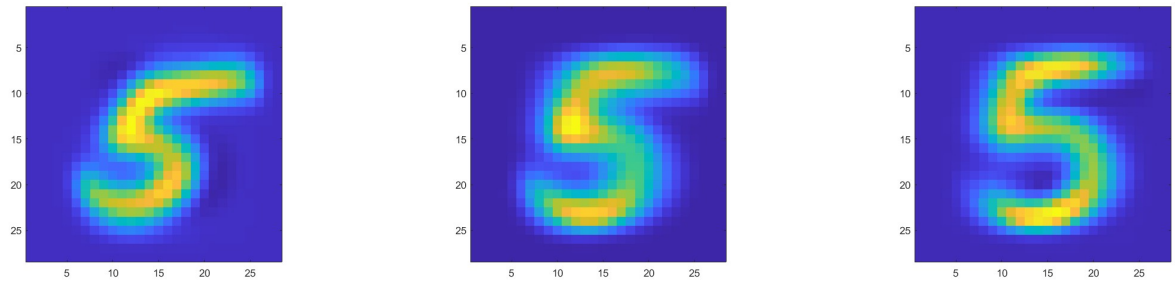
N=3



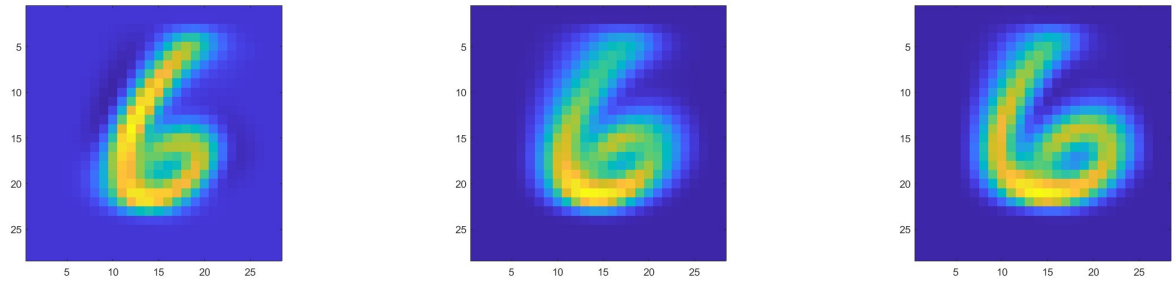
N=4


5
10
15
20
25
5 10 15 20 25

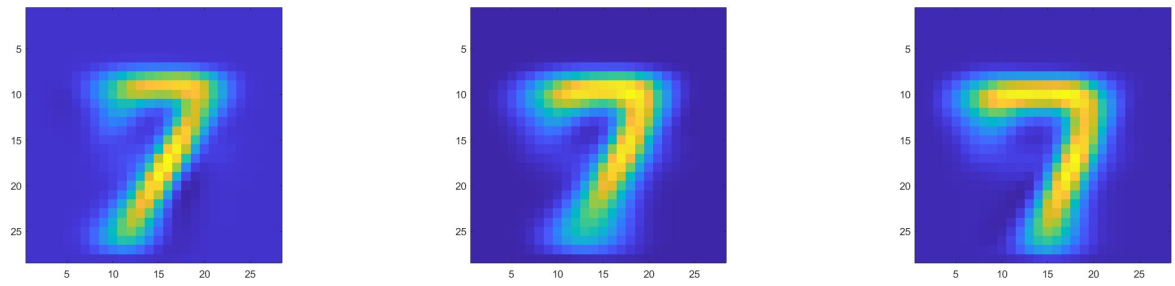
N=5


5
10
15
20
25
5 10 15 20 25

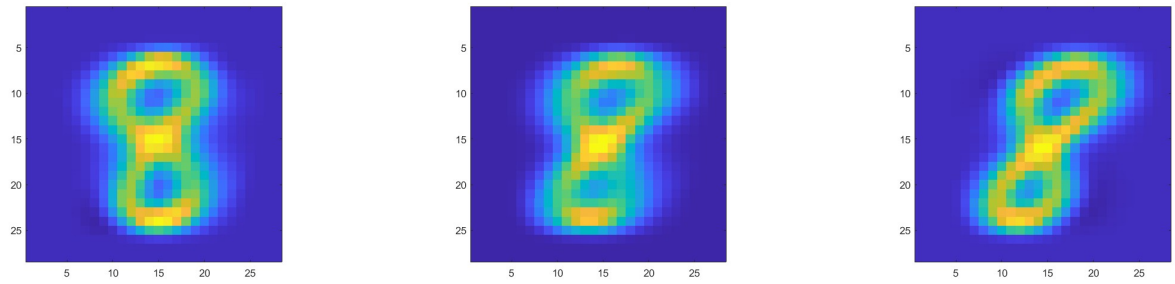
N=6


5
10
15
20
25
5 10 15 20 25

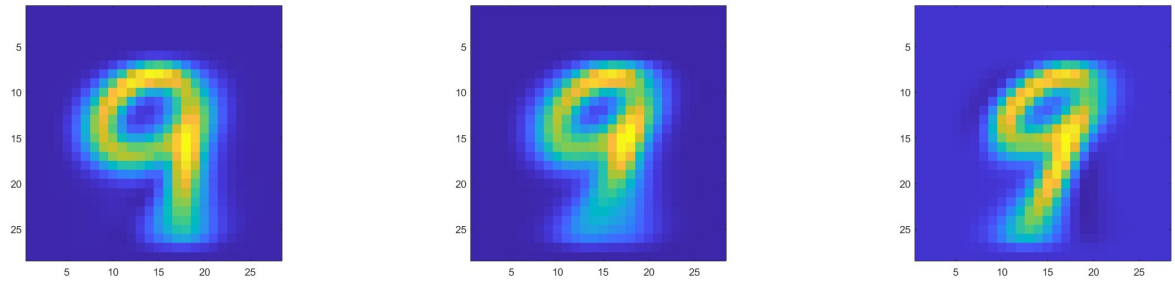
N=7



N=8



N=9



1.5.1 Comments and Justification

The results contain 3 images which are being produced using mean and principle modes of variation(eigenvector with the max of eigenvalue), it looks like different ways to write to all digits from 0 to 9, with mostly just a variation of angles they are written, 2 has a small circle which a significant number people tend to write.

For 1 it's a bit slant(drawn with mean), but after using the principal mode of variation, one got straighter on adding and slant on removing the principal mode of variation, so the principal mode of variation just changes the angle of 1 with the y-axis, so it tells that the main variation in people writing one is the angle they give with the

y-axis.

Instructions to run the code:

Run digitImages.m from code folder of Q4