

homework1

Name: Mukul Sati

GTID: 902917135

February 7, 2016

Contents

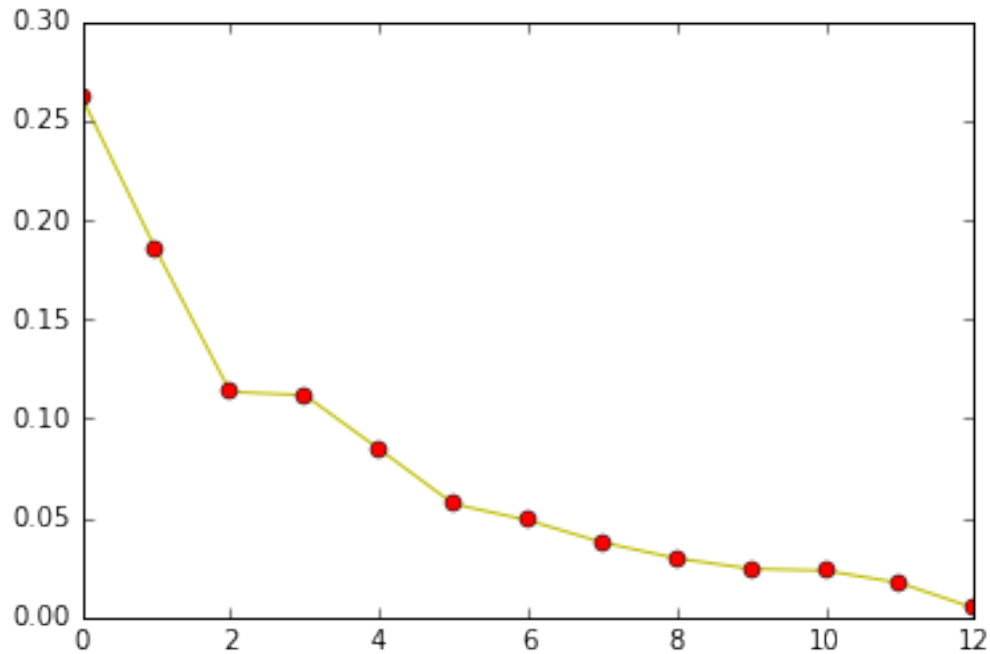
1	Homework 1
---	------------

2

Homework 1

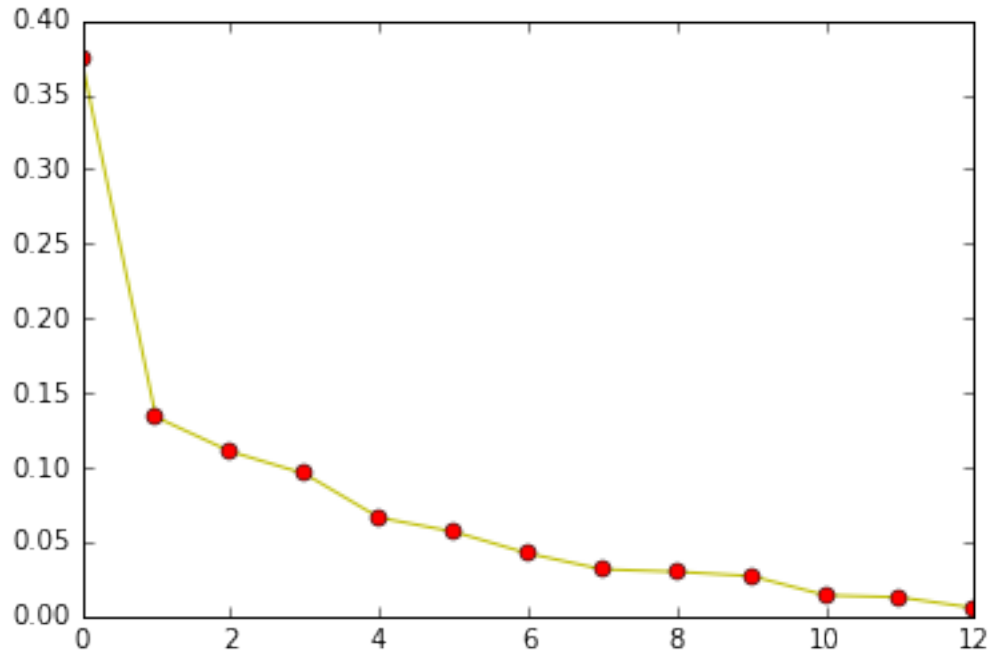
For my implementation of PCA, I show the 10th eigenvector instead of the 20th, as the 20th is zero as the data is 13-dimensional. The most important eigenvector, the 10th eigenvector and the plot for the ratio of the sorted eigenvalues to the sum of the eigenvalues for the wine dataset for two runs, one with 5 and the other time with 50 random samples of each class selected for training are:

Using 5 samples for training:



```
1st eigenvector: [-0.20301402 -0.31257139  0.01919546  0.45675768 -0.04065617  0.07177575
-0.59649598  0.0355507   0.17157012 -0.45488468  0.02213474  0.2260994
 0.015657  ]
10th eigenvector: [-0.30530884 -0.22437366  0.28270088  0.39554347  0.1035333  -0.08530881
 0.31634353 -0.11676717  0.4330378   0.26617685 -0.12875669 -0.31277113
 0.33998248]
```

Using 50 samples for training:

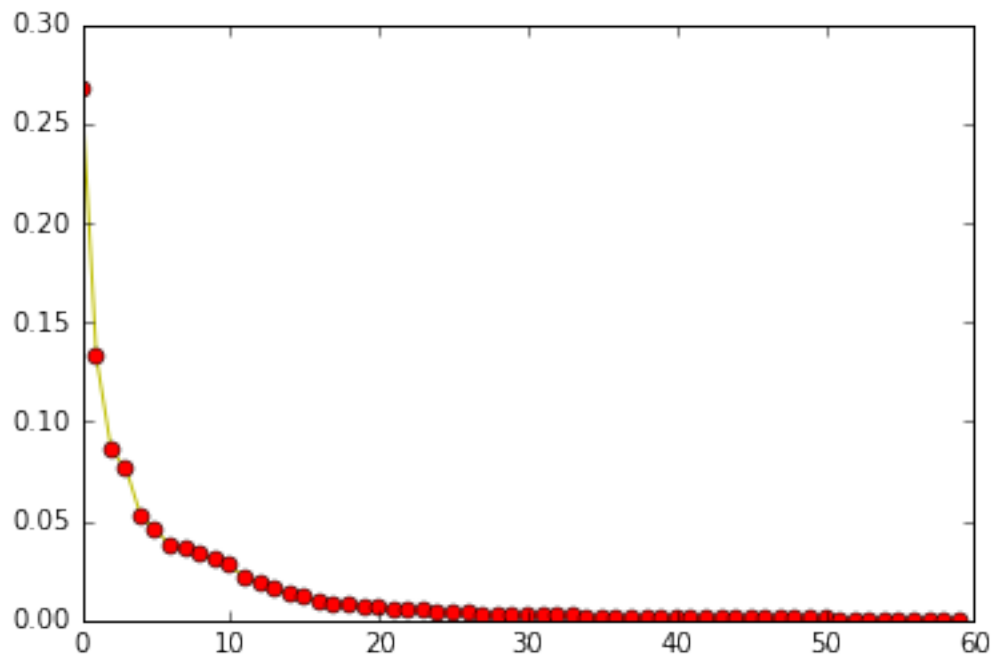


```
1st eigenvector: [-0.36238596 -0.15505812  0.05559563 -0.28095178  0.19543318 -0.12934828
 0.20626521 -0.18849597  0.0075404  -0.34747772 -0.68146153 -0.18718634
 0.09685696]
10th eigenvector: [-0.37476352 -0.22292995  0.10433449 -0.21738561  0.12079655  0.18908253
 -0.14714305 -0.00722471  0.09693343 -0.07920049  0.1982744  0.71330035
 0.34042052]
```

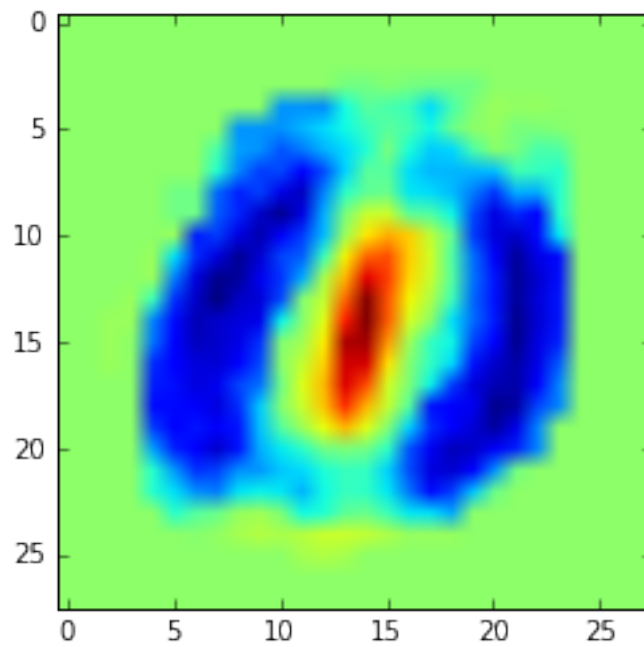
I also use a variable number of eigenvectors for the reconstruction, discarding those for which the ratio: (eigenvalue)/(sum of eigenvalues) falls below a threshold (I use a threshold of 0.03). Using this to reconstruct a test-example, I get the following reconstruction errors:

```
(784, 784) (60, 784)
(60, 784)
```

Using 5 samples fototalTrainingDatar training:



1st eigenvector:



10th eigenvector:

