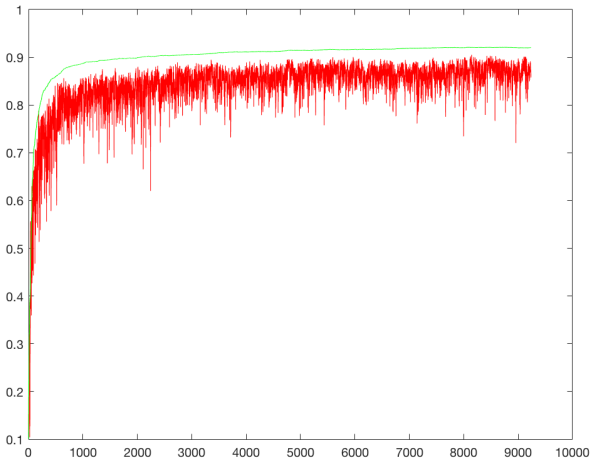
Machine Learning (CS771A) - Homework 2

Sai Kishan Pampana(13458)

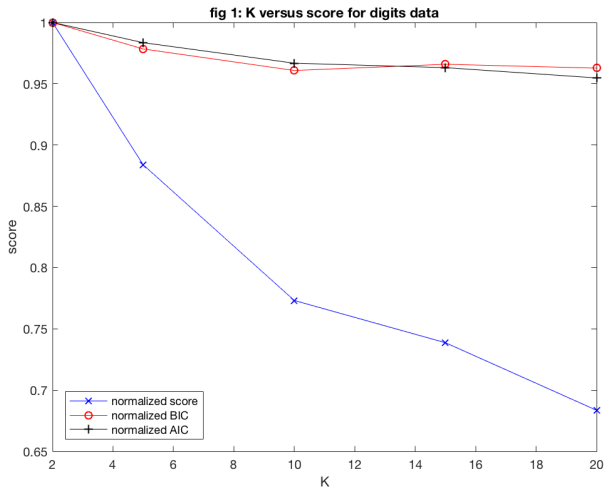
---

1. **Problem 1**



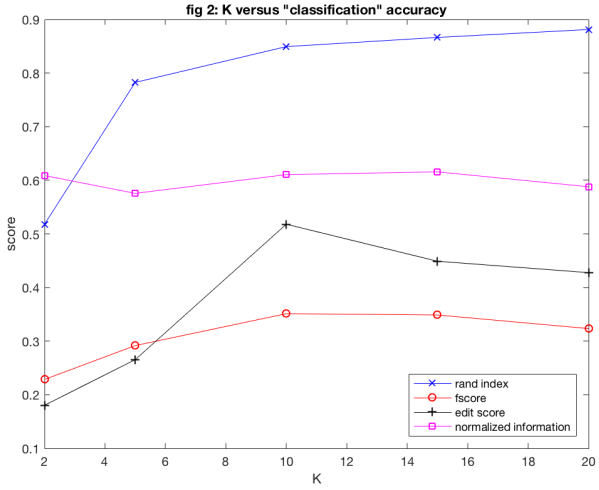
Green line in the graph represents accuracy for $\mathbf{W}_{avg}$. We observe that this is always higher than$\mathbf{W}$.
In this we observe such ocillatory graph because, $\mathbf{W}$ changes suddenly but the $\mathbf{W}_{avg}$ is the average so it doesn't change suddenly. Initial we have a sudden rise in the accuracy and later we have an almost constant rate this is because initially we did not have any correction but later on we will have lesser rate of corrections.

2. **Problem 2**

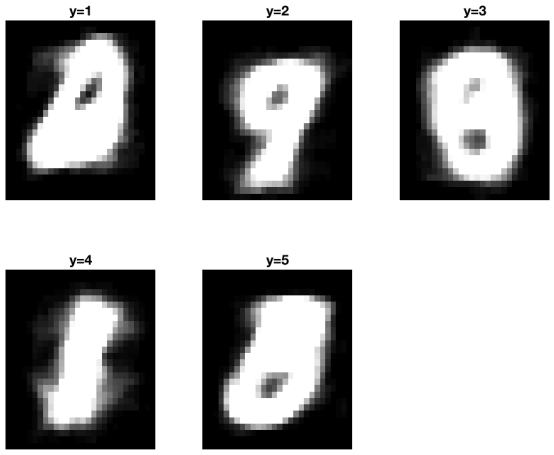


(a) Here we see a regular trend with the Normalized scores, they keep on decreasing as we increase K, so we are getting better results.
But for the case of BIC we have a slightly different behavior, meaning that we have a minimum

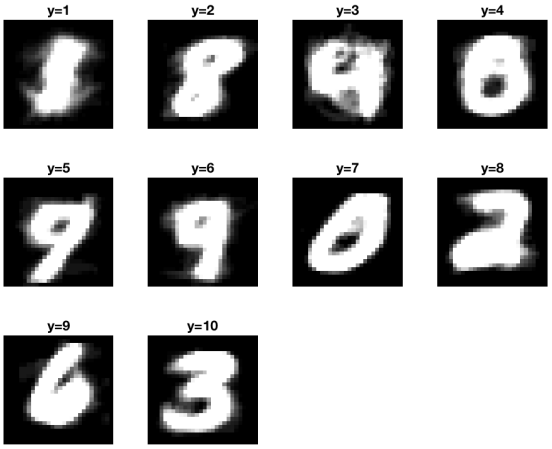
and not a decreasing trend. BIC reaches a minimum at K = 10.
In AIC we again see a regular trend of decreasing score with increasing K. We have a minimum at the maximum K and in this case 20.
So here for AIC I would choose K = 20 and for the BIC I would take K = 10.



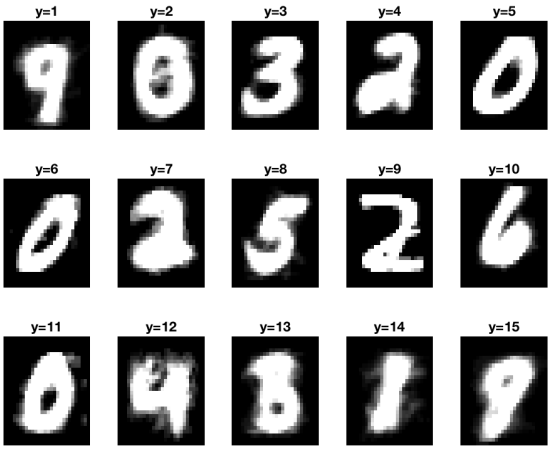
(b) Based on the scoring that we have that are represented in the graph I think it is better if we choose fscore, because this curve is more smooth than the rest of the scores and it is covex.
As for the value of K that we should be choosing, I think we should take K = 10 because the accuracy is highest for this value of K



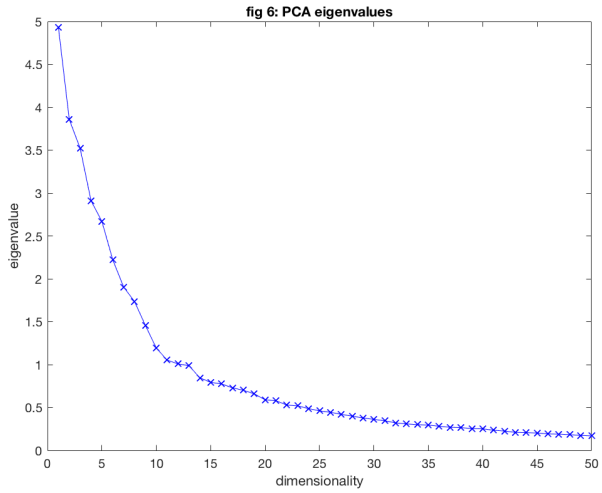
(c) In the given centers for K = 5 the figures that are recognizable are 1 and 9, the remaining images are not clear and seem to be a fusion of some digits.
In one of the images it seems that 3,8,5,0 have been merged together, in another it seems that 0,1,2 seem to be merged together, and in the last image it seems like 6,5,3 seem to be fused together.



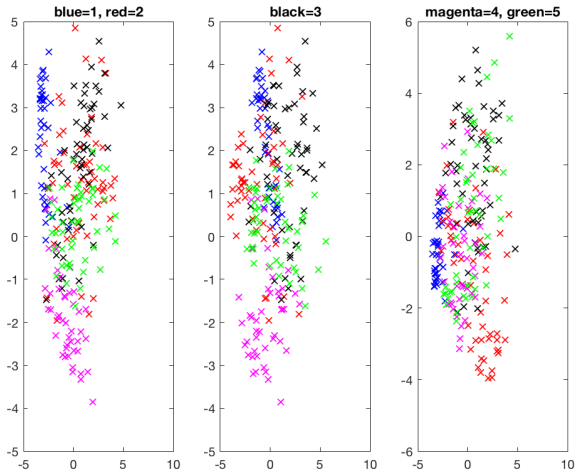
(d) In the given centers for K = 10 the figures that are recognizable are 2,0,9,3,8,6 and the remaining figures are not recognizable.
In one image 7,9 seems to be merged in another 8,0 seem to be merged in another 1 and 3 seem to be fused and in the last one 4,7,9,8 seems to be merged.



(e) In this image of centers for K = 15 the images that are recognizable are 0,2,3,5,6,9,0,0,0,1 and remaining are not.
0,3; 4,9,7; 3,8; 3,8; 2,7 seems to be have merged.

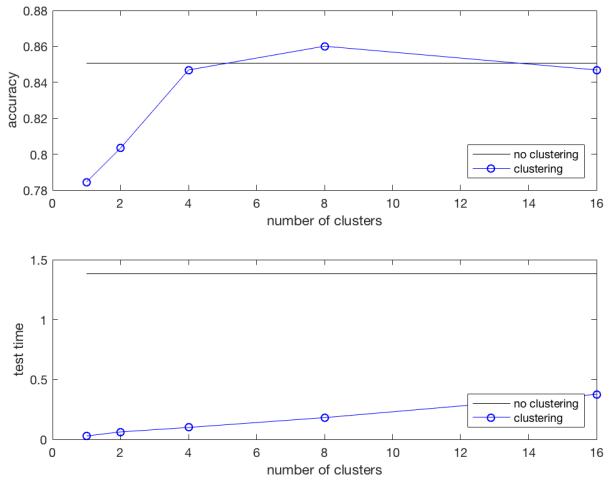


(f) For the case of high eigenvalues the eigenvector in that direction has high variance. bang for the buck start to significantly decrease as we include eigenvectors with low eigen values. Take those eigenvector for which eigenvalues till dimension reaches 25.

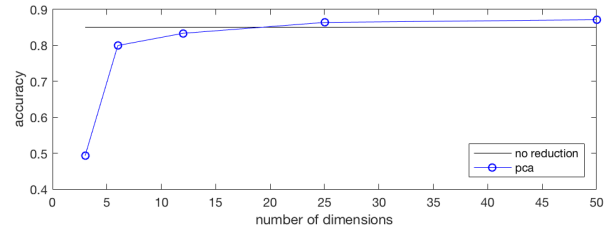


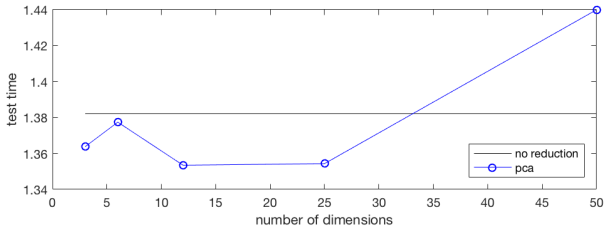
(g) In these images the classes represented by blue(class 1) and magenta(class 4) are clear the remaining are not clear. The images have merged because we have some of the parts of each center that are similar to the other center.

(h) They actually look as required in fig 11. From figure 6 if we use less than 25 dimensions means we have less eigen values, less standard deviation so numbers are not clear.
Edges are blurred in the artifacts at 50 dimensions. This blurness is due to the reconstruction of different distorted/oriented numbers in the training data.



(i) Initially accuracy increases with number of clusters, reaches maximum at 8 and then starts decreasing.
In order to achieve roughly the same accuracy as baseline 5 clusters per class are needed.
Test time increases with increase in number of clusters.
This is good in practice because we will be achieving almost the same accuracy with lesser time.





(j)

Accuracy is increasing with increase in number of clusters and it also exceeds the baseline accuracy. Here the trend that we see is that first the time increases and then decrease and after that it keeps on increasing. Also initially all the the time are less than the base time.
This is good in practice as it provides good accuracy in less test time.