COL726: LSI, PCA

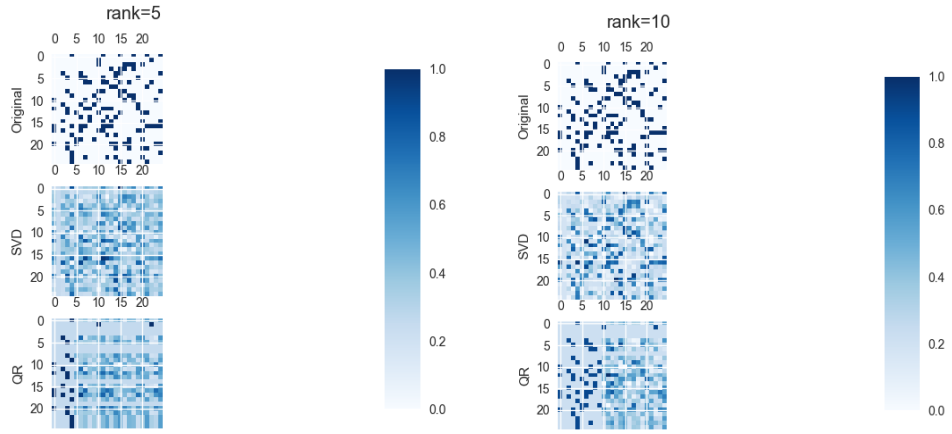
Haroun Habeeb 2013CS10225

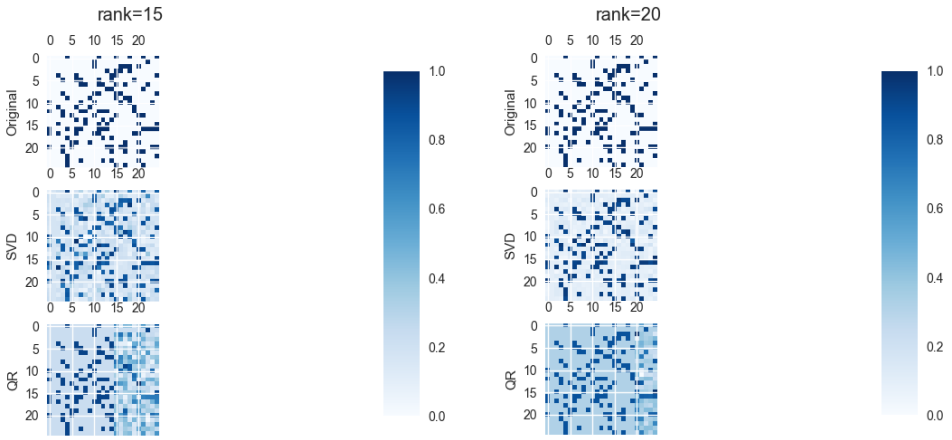
Part 1: LSI

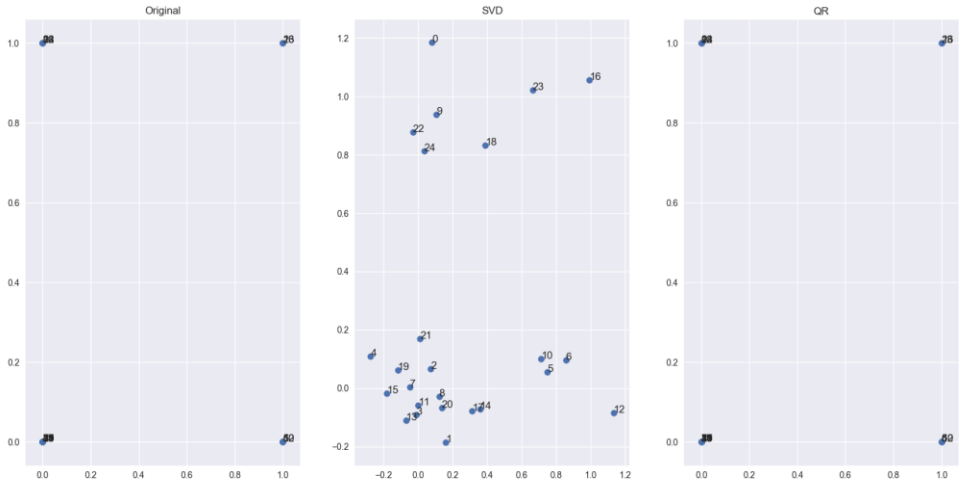
I used python and numpy instead of matlab.

Also, I'm not a fan of reading CS books, so I used a randomly generated binary matrix. The following images show that QR looks a lot more like the original matrix than SVD. Simply look at the rank 5 and rank 10 results.

On the other hand, the noise in SVD is visually less than the noise in QR. Look at the right half of rank 10 results.





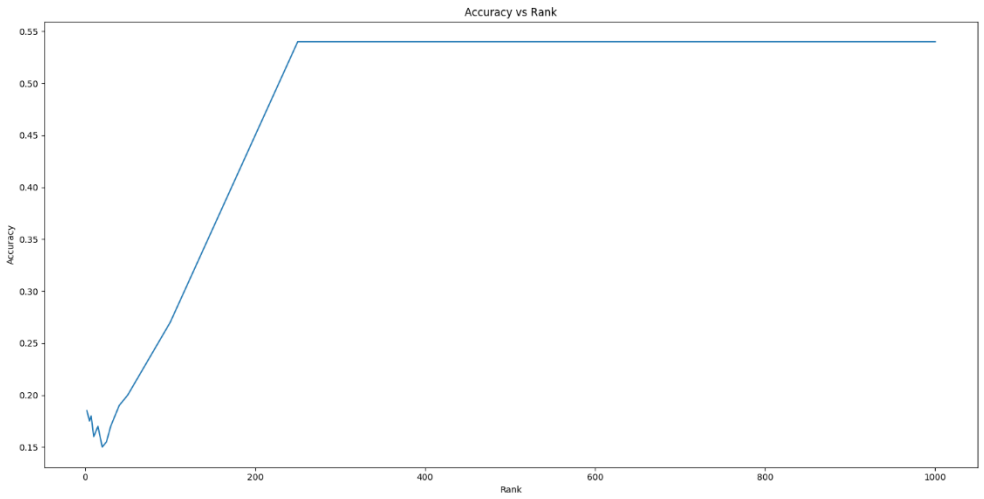


A quick look at the clusters tells us that both methods form good clusters.

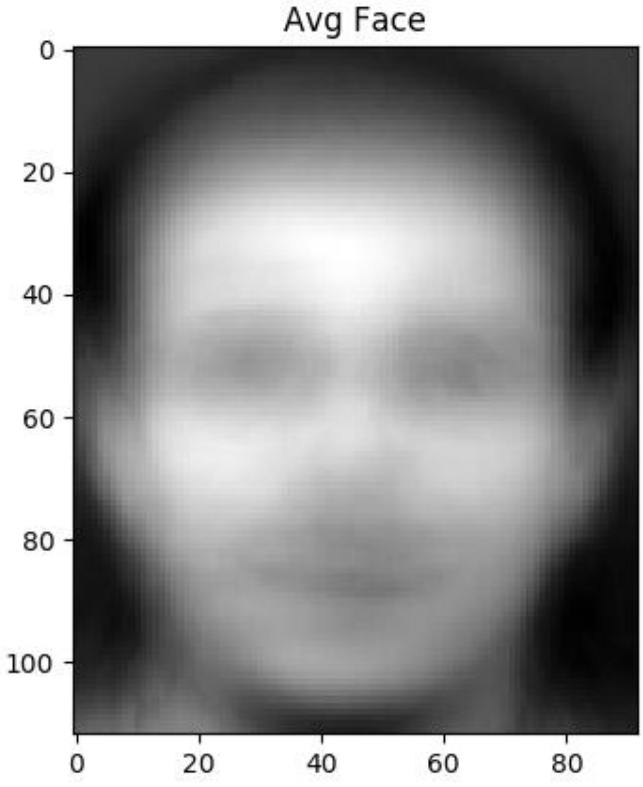
Part 2: PCA

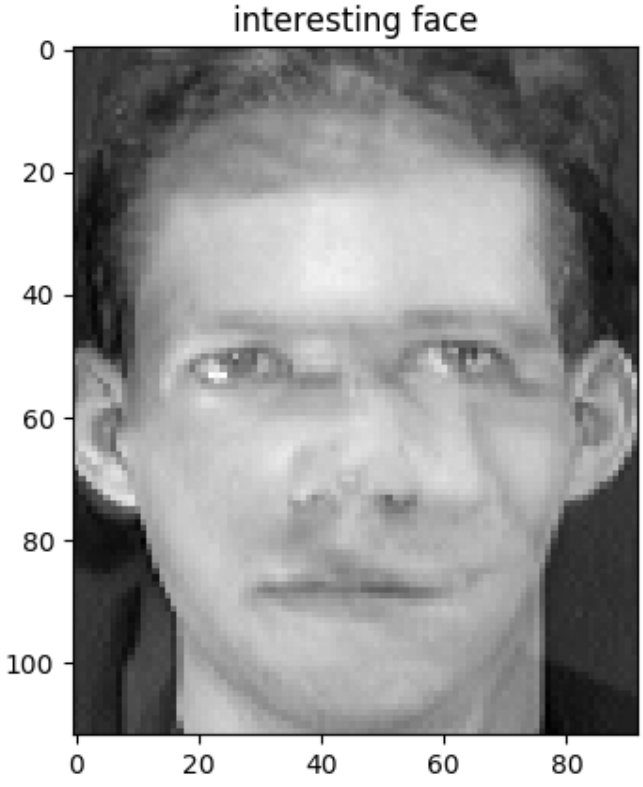
I don't have friends ☹, so I used a dataset that I found online. It was used in the Machine Learning course in IITD too. It has 40 subjects and 10 photos of each subject. Each face is in a slightly different pose from the others.

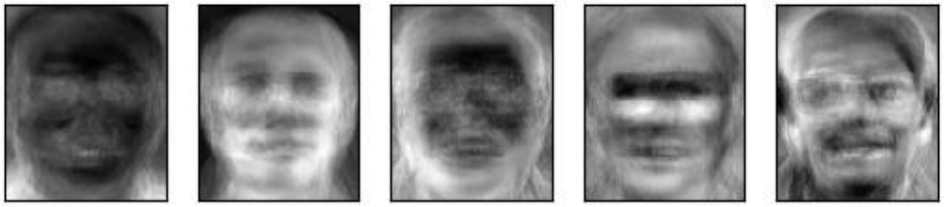
For face recognition, I trained an SVM. Python + sklearn library was used.



We see that the accuracy oscillates at very low rank, and then increases until it reaches a plateau.


Avg Face
0
20
40
60
80
100
0
20
40
60
80


interesting face
0
20
40
60
80
100
0
20
40
60
80



Eigen Faces

It appears as if the eigen values capture slightly different parts of the face. For example, the first one in the bottom row captures a forehead and so on.