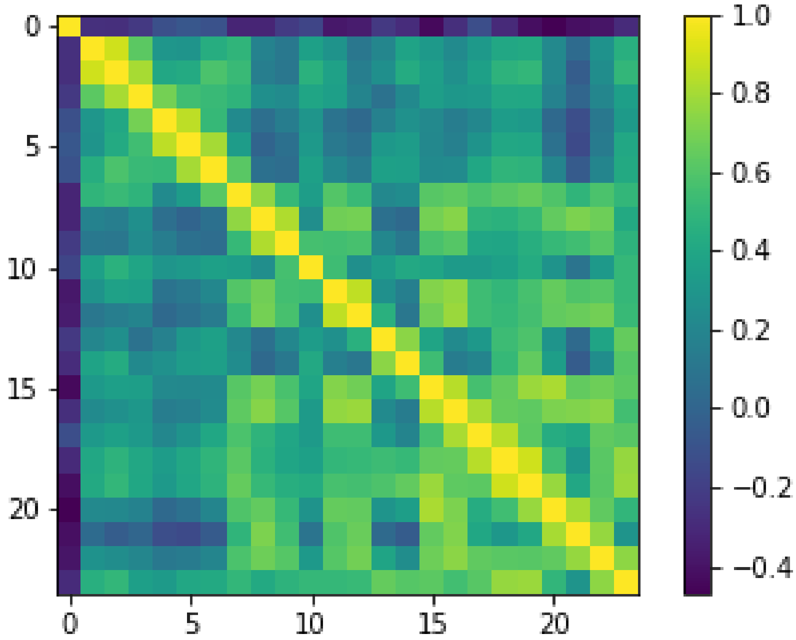
INF2B Coursework

April 14, 2020

1 Task 1

1.2

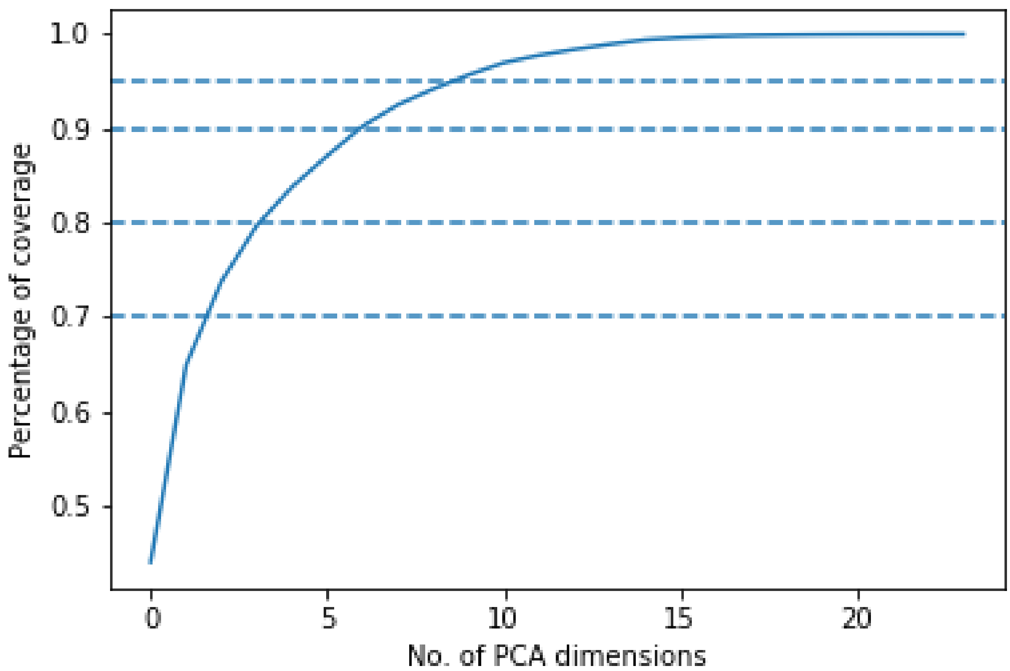


Looking at the correlation matrix the first thing that comes up to our eye is how all the diagonal elements 1. This however can easily be explained as entries of 1 denote a perfect correlation.

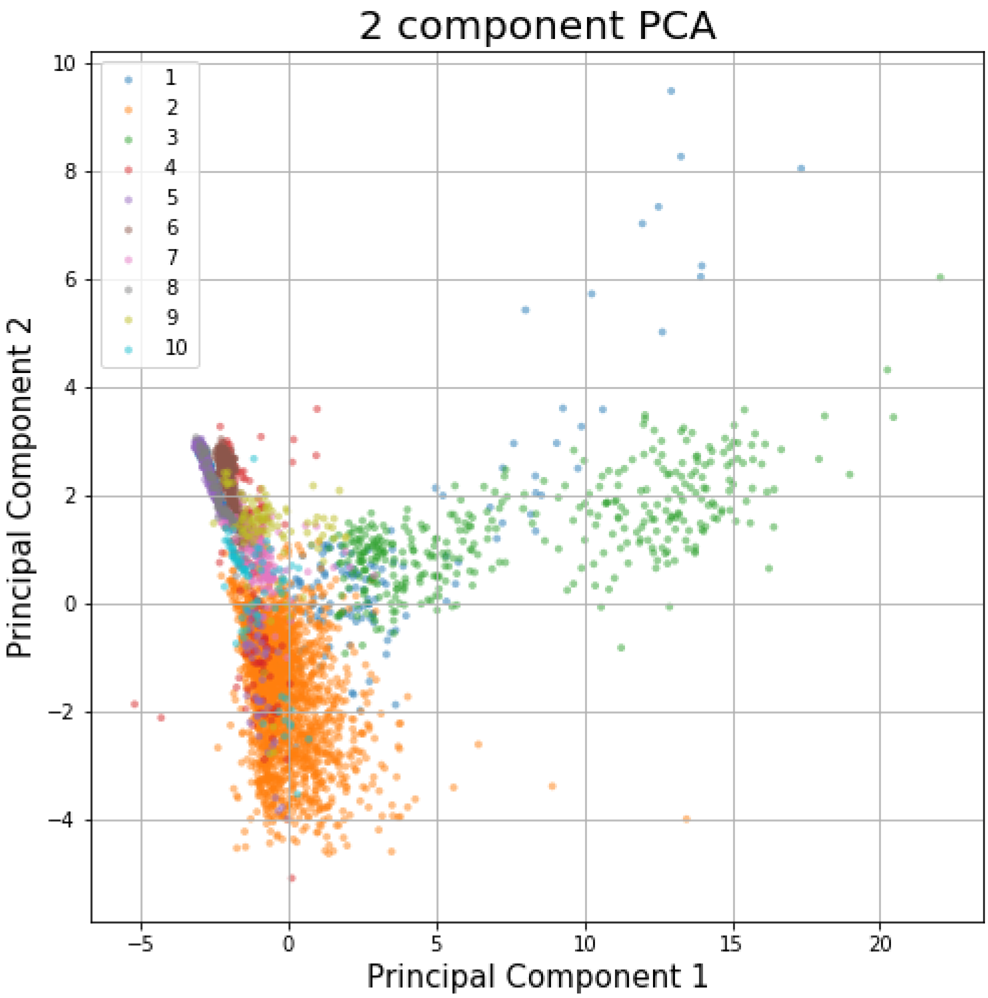
In a correlation matrix, positive values denotes positive correlation which mean as the value of feature i increases so does feature j, and negative values denote negative correlation where if feature i increases then feature j decreases. We try and look for such patterns in our correlation matrix R.

Another striking feature of our correlation matrix R is that if we start looking for feature pairs we can see almost all them seem to positively correlate pairwise. Except for the first feature vector which is negatively correlated to all others. This can be confirmed both visually and by using a scatter plot and then trying to find a line to fit them.

1.3.b



1.3.c


2 component PCA
Principal Component 1
Principal Component 2
1
2
3
4
5
6
7
8
9
10
10
8
6
4
2
0
−2
−4
−5
0
5
10
15
20

1.4.b

	0	1	2	3	4	5	6	7	8	9
0	0.032667	0.000000	0.000000	0.000000	0.041819	0.000000	0.000000	0.008933	0.000000	0.000000
1	0.000000	0.505667	0.000472	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.001417	0.000703	0.085994	0.000000	0.000000	0.000000	0.000000	0.000000	0.000236	0.000000
3	0.000236	0.013396	0.000000	0.016668	0.000000	0.007044	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000	0.000000	0.069781	0.000000	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000	0.000000	0.001175	0.117009	0.000000	0.000000	0.000000	0.000000
6	0.000000	0.000236	0.000000	0.000000	0.000000	0.000000	0.031256	0.003041	0.000708	0.000236
7	0.000000	0.000000	0.000000	0.000000	0.004222	0.000000	0.000000	0.013863	0.000000	0.000000
8	0.000000	0.000000	0.000000	0.000000	0.000461	0.000000	0.000000	0.000000	0.024904	0.000000
9	0.000000	0.008927	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.008927

Figure 1: Full co-variance matrix has an accuracy of 90.67372510187762%.

	0	1	2	3	4	5	6	7	8	9
0	0.021385	0.003053	0.007758	0.000000	0.000000	0.000000	0.000472	0.050752	0.000000	0.000000
1	0.000000	0.504959	0.000708	0.000000	0.000000	0.000472	0.000000	0.000000	0.000000	0.000000
2	0.002119	0.022278	0.063480	0.000000	0.000000	0.000000	0.000000	0.000000	0.000472	0.000000
3	0.000000	0.019041	0.000000	0.009388	0.000000	0.008680	0.000236	0.000000	0.000000	0.000000
4	0.000000	0.010591	0.000000	0.000000	0.040408	0.000000	0.000000	0.018551	0.000000	0.000230
5	0.000000	0.000000	0.000000	0.000000	0.000236	0.117481	0.000000	0.000467	0.000000	0.000000
6	0.000000	0.001169	0.000000	0.000000	0.000000	0.000236	0.030323	0.003041	0.000708	0.000000
7	0.000000	0.000000	0.000000	0.000000	0.000708	0.000000	0.000000	0.017376	0.000000	0.000000
8	0.000000	0.000703	0.000000	0.000000	0.000000	0.001647	0.000708	0.000000	0.021604	0.000703
9	0.000000	0.006347	0.000000	0.000000	0.000000	0.004227	0.000000	0.000000	0.000000	0.007280

Figure 2: Diagonal co-variance matrix has an accuracy of 83.36843508397761%.

	0	1	2	3	4	5	6	7	8	9
0	0.062951	0.005869	0.005644	0.000000	0.008720	0.000236	0.000000	0.000000	0.000000	0.000000
1	0.000000	0.506139	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.000000	0.027231	0.059708	0.000000	0.000000	0.000000	0.000000	0.000000	0.001411	0.000000
3	0.000230	0.026776	0.000000	0.005863	0.000000	0.003058	0.000000	0.000000	0.000000	0.001417
4	0.000000	0.010822	0.000000	0.000000	0.058960	0.000000	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.006105	0.000000	0.000000	0.000000	0.112079	0.000000	0.000000	0.000000	0.000000
6	0.000703	0.009152	0.000000	0.000000	0.000000	0.000000	0.025393	0.000230	0.000000	0.000000
7	0.003761	0.003064	0.000000	0.000000	0.009618	0.000000	0.000000	0.001641	0.000000	0.000000
8	0.000236	0.015044	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.010085	0.000000
9	0.000000	0.011519	0.000000	0.000236	0.000000	0.000000	0.000000	0.000000	0.000000	0.006099

Figure 3: Shared co-variance matrix has an accuracy of 84.89183836691168%.

1.5

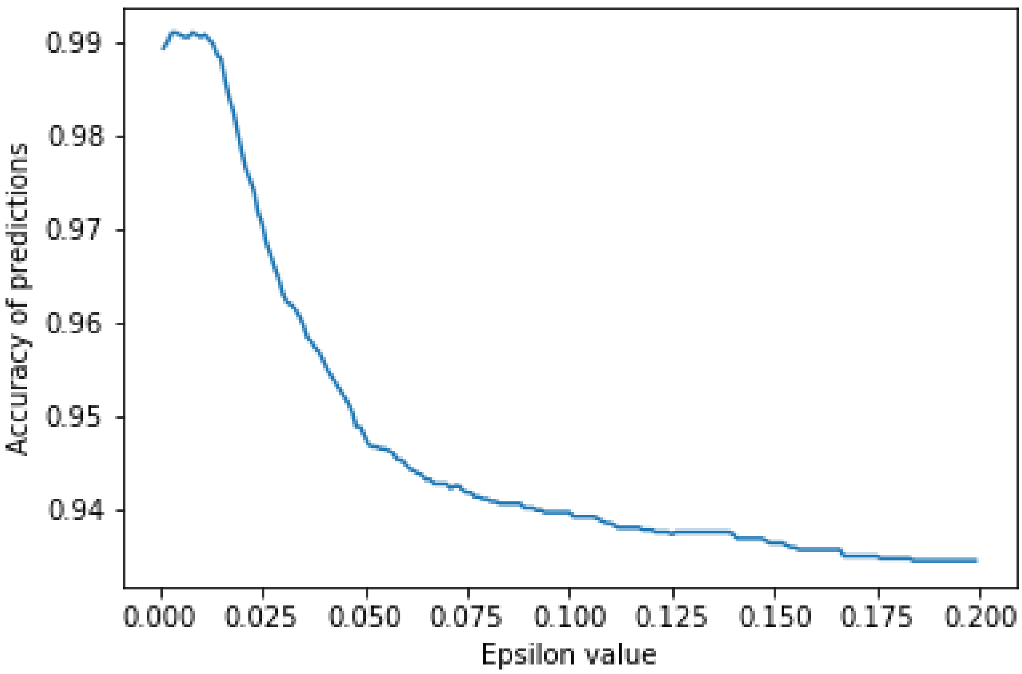


Figure 4: Graph of accuracy against epsilon values (excluding 0).

During the experiment we can see how we get the best results for the accuracy for extremely small values around 0. We get errors when calculating the accuracy for epsilon value 0, as we need some small parameter like epsilon to avoid calculation of inverse of a singular matrix. Following this we can clearly see how the accuracy value sharply drops for larger values of epsilon.