

Phase 3 report

Armin Soltan 610396115

Preface

In this phase I should get eigenvectors of iris data set. So first of all i should tracking data from file and clean data. I use Min Max scalar algorithm for pre processing of data then i use pca.

Get eigen vectors

In this task first of all i implement covariance matrix using by numpy package then by builtin its method i get eigen values and eigen vectors. Also we can use eigenvectors by different normalization method like Standard Scalar and builtin sklearn method but i prefer using min max algorithm.

```
(venv) armin@armin-Lenovo-751-70:~/courses/data mining/project 3$ python3
enter task: 1
eigen values
[0.23231168 0.03239279 0.00963728 0.00175337]
eigen vectors
[[ 0.4252893  -0.42100611 -0.71434339  0.36276216]
 [-0.14612219 -0.90470509  0.33510219 -0.21877734]
 [ 0.61610511  0.06432645 -0.06825927 -0.78205964]
 [ 0.64667752  0.01116485  0.61054133  0.45708076]]
```

Figure 1: eigen vectors

Draw 3D space

In this task first we should reduce dimensionality from 4 dimension in to the 3 dimension then split component and then draw space using by scatter plot. I show different class in this space by different color. You can see below figure.

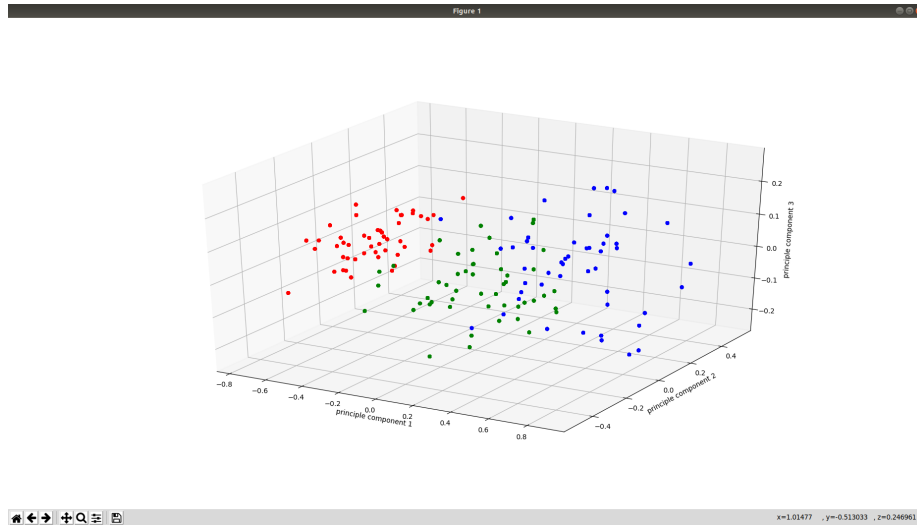


Figure 2: 3D space

Compare two way of PCA

In this task i should compare two different way of reduce dimensionality of 4 dimension space. first way is step by step in other word first we use pca to convert 4 dimension to 3 dimension then again use pca to get 2 dimensional space. In second way we use PCA to get 2 dimensional spcae directly. You can see there is no different between two ways and scatter plot of both ways are exactly the same.

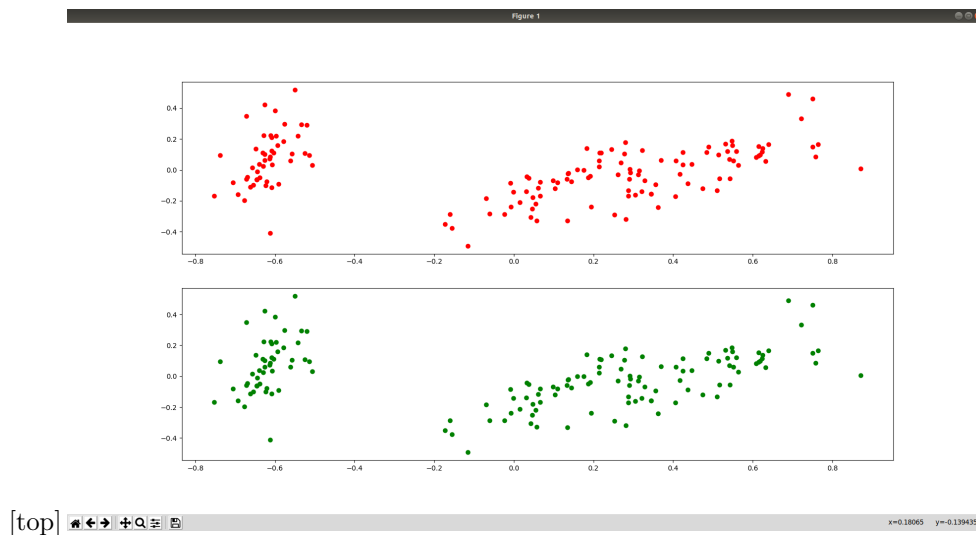


Figure 3: Comparison ways of pca