Experiment 8: Principle Component Analysis

Table of Contents

Name: Ventrapragada Sai Shravani

PRN:17070123120

Batch:Entc(2017-21) G-3

Aim:

Implementation of Principle Component Analysis (PCA)

Theory:

PCA is a method for dimensionality reduction. pcs is a n-by-n matrix of principal components. scrs is an m-by-n matrix containing the data transformed using the linear coordinate transformation matrix pcs (first output). pexp is a vector of length n containing the percentage of variance explained by each principal component. data is an m-by-n numeric matrix. The n columns correspond to n observed variables. Each of the m rows corresponds to an observation.

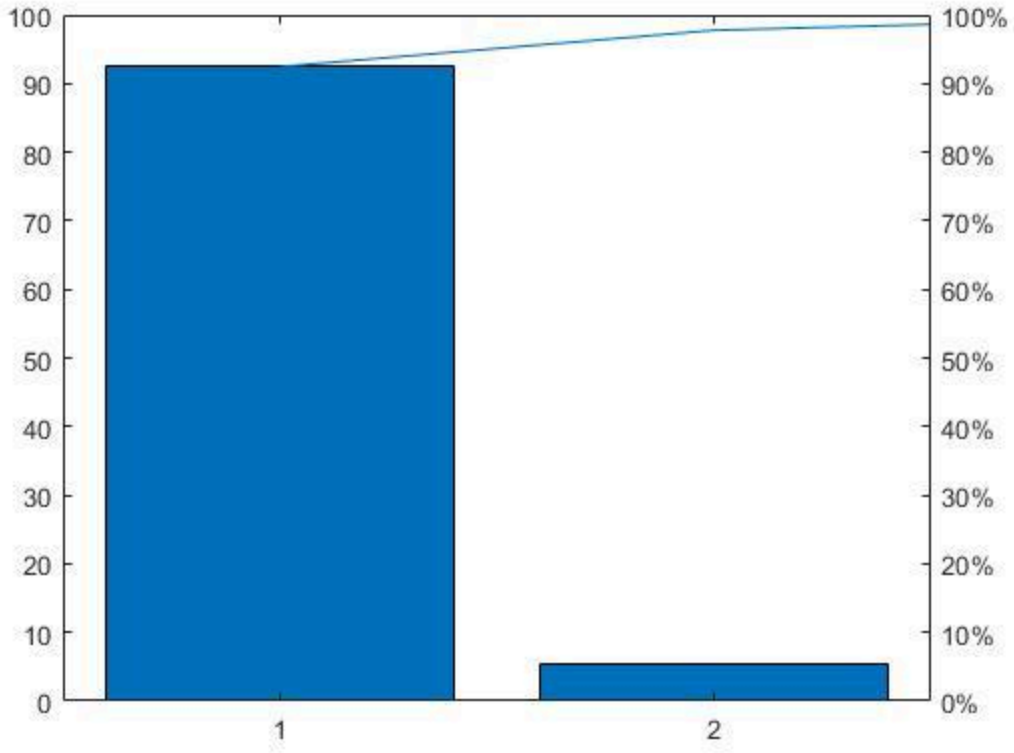
The output matrix P contains the principal components, represented in terms of the original variables x1 and x2. The first column of P contains the coefficients of the first principal component p1, and the second column contains the coefficients of the second principal component p2.

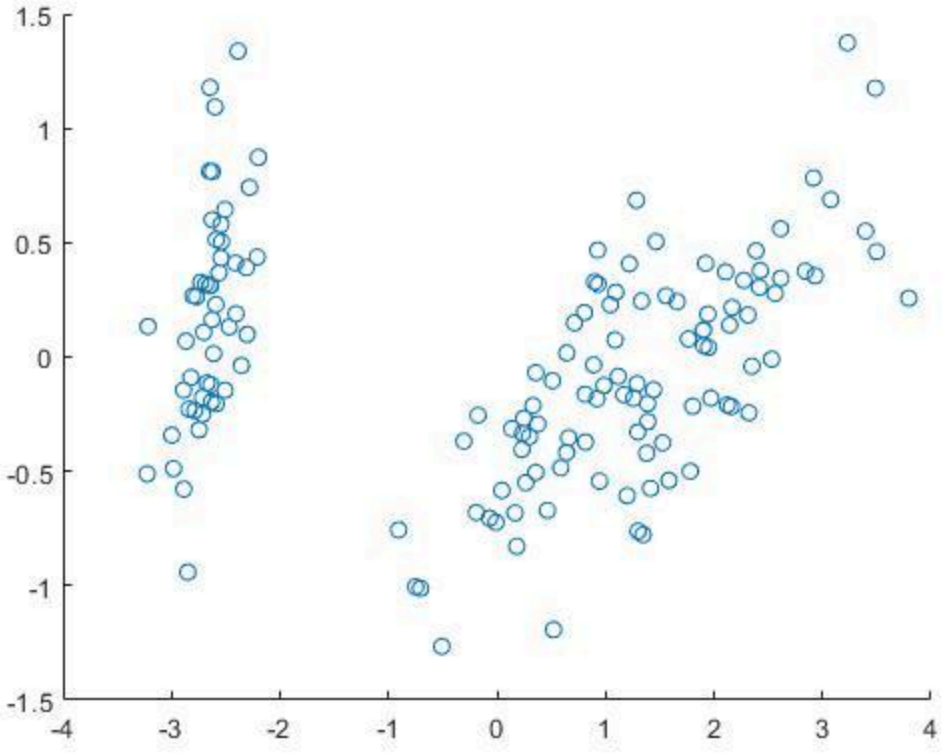
The second output scrs is a matrix containing the observations in data expressed in the coordinate space of the principal components p1 and p2. The last output pexp is a vector containing the percent variance explained by each principal component. Here, most of the variance in the observations is explained by the first principal component. We can use only the first column of the transformed data, reducing the dimension of the data from 2 to 1.

Principle Component Analysis Part 1

```
clc;
clear all;
close all;
load fisheriris
[pcs,scrs,~,~,pexp] = pca(meas);
figure(1);
```

```
pareto(pexp)
figure(2);
scatter(scrs(:,1),scrs(:,2))
```





Conclusion

In the above experiment I have used fisheriris dataset to scale the feature in data. I learnt how the accuracy could be compromised for better results. Also how pareto chart works and plotting scatter plot.

Published with MATLAB® R2020a