Assignment #2

Sara Bayat

17 November 2023

Contents

```r
knitr::opts_chunk$set(echo = TRUE)
knitr::opts_chunk$set(warning = FALSE, message = FALSE)
```

Interducing the Data (optional)

In order to showcase the proficiency of unsupervised techniques, the wine dataset, which has not any missing value, has been utilized, which comprises three diverse cultivars, each with **thirteen distinct features** that represent varying quantities of components. Below, you can find a detailed *statistical summary* for each variable.

```r
# set work directory
setwd("E:/Carleton - Master/Fall 2023 - semester 1/Data Mining/Data Mining - Dr. Mills/Assignment #2")
# read data file
winedata <- read.csv("wine.data.txt",sep=",", header=F)
names(winedata) <- c("Cvs", "Alcohol", "Malic acid","Ash", "Alcalinity of ash",
                     "Magnesium", "Total phenols", "Flavanoids", "Nonflavanoid phenols",
                     "Proanthocyanins", "Color intensity", "Hue",
                     "OD280/OD315 of diluted wines", "Proline")
head(winedata)
```

```
##   Cvs Alcohol Malic acid  Ash Alcalinity of ash Magnesium Total phenols
## 1   1   14.23       1.71 2.43              15.6       127          2.80
## 2   1   13.20       1.78 2.14              11.2       100          2.65
## 3   1   13.16       2.36 2.67              18.6       101          2.80
## 4   1   14.37       1.95 2.50              16.8       113          3.85
## 5   1   13.24       2.59 2.87              21.0       118          2.80
## 6   1   14.20       1.76 2.45              15.2       112          3.27
##   Flavanoids Nonflavanoid phenols Proanthocyanins Color intensity  Hue
## 1       3.06                 0.28            2.29            5.64 1.04
## 2       2.76                 0.26            1.28            4.38 1.05
## 3       3.24                 0.30            2.81            5.68 1.03
## 4       3.49                 0.24            2.18            7.80 0.86
## 5       2.69                 0.39            1.82            4.32 1.04
## 6       3.39                 0.34            1.97            6.75 1.05
##   OD280/OD315 of diluted wines Proline
## 1                         3.92    1065
## 2                         3.40    1050
## 3                         3.17    1185
## 4                         3.45    1480
## 5                         2.93     735
## 6                         2.85    1450
```

```r
sum(colSums(is.na(winedata)))
```

```
## [1] 0
```

```r
summary(winedata)
```

```
##       Cvs           Alcohol        Malic acid         Ash
##  Min.   :1.000   Min.   :11.03   Min.   :0.740   Min.   :1.360
##  1st Qu.:1.000   1st Qu.:12.36   1st Qu.:1.603   1st Qu.:2.210
##  Median :2.000   Median :13.05   Median :1.865   Median :2.360
##  Mean   :1.938   Mean   :13.00   Mean   :2.336   Mean   :2.367
##  3rd Qu.:3.000   3rd Qu.:13.68   3rd Qu.:3.083   3rd Qu.:2.558
##  Max.   :3.000   Max.   :14.83   Max.   :5.800   Max.   :3.230
##  Alcalinity of ash   Magnesium      Total phenols     Flavanoids
```

```
##  Min.   :10.60     Min.   : 70.00   Min.   :0.980   Min.   :0.340
##  1st Qu.:17.20     1st Qu.: 88.00   1st Qu.:1.742   1st Qu.:1.205
##  Median :19.50     Median : 98.00   Median :2.355   Median :2.135
##  Mean   :19.49     Mean   : 99.74   Mean   :2.295   Mean   :2.029
##  3rd Qu.:21.50     3rd Qu.:107.00   3rd Qu.:2.800   3rd Qu.:2.875
##  Max.   :30.00     Max.   :162.00   Max.   :3.880   Max.   :5.080
##  Nonflavanoid phenols Proanthocyanins Color intensity       Hue
##  Min.   :0.1300       Min.   :0.410   Min.   : 1.280   Min.   :0.4800
##  1st Qu.:0.2700       1st Qu.:1.250   1st Qu.: 3.220   1st Qu.:0.7825
##  Median :0.3400       Median :1.555   Median : 4.690   Median :0.9650
##  Mean   :0.3619       Mean   :1.591   Mean   : 5.058   Mean   :0.9574
##  3rd Qu.:0.4375       3rd Qu.:1.950   3rd Qu.: 6.200   3rd Qu.:1.1200
##  Max.   :0.6600       Max.   :3.580   Max.   :13.000   Max.   :1.7100
##  OD280/OD315 of diluted wines    Proline
##  Min.   :1.270                Min.   : 278.0
##  1st Qu.:1.938                1st Qu.: 500.5
##  Median :2.780                Median : 673.5
##  Mean   :2.612                Mean   : 746.9
##  3rd Qu.:3.170                3rd Qu.: 985.0
##  Max.   :4.000                Max.   :1680.0
```

Question 1 - part 1 - PCA (Principal Component Analysis)

The data was analyzed using a method that generated 13 principal components called PC1 to PC13. Each of these components plays a crucial role in explaining a certain proportion of the variance present in the dataset. As we use normalized values for principal component analysis (PCA), we utilize the function *scale()* to standardize our data and ensure accurate analysis.

```
library(stats)
library(factoextra)
# normalized data
wine_s <- scale(winedata[,2:14])
head(wine_s)
```

```
##        Alcohol  Malic acid        Ash Alcalinity of ash  Magnesium
## [1,] 1.5143408 -0.56066822  0.2313998        -1.1663032 1.90852151
## [2,] 0.2455968 -0.49800856 -0.8256672        -2.4838405 0.01809398
## [3,] 0.1963252  0.02117152  1.1062139        -0.2679823 0.08810981
## [4,] 1.6867914 -0.34583508  0.4865539        -0.8069748 0.92829983
## [5,] 0.2948684  0.22705328  1.8352256         0.4506745 1.27837900
## [6,] 1.4773871 -0.51591132  0.3043010        -1.2860793 0.85828399
##      Total phenols Flavanoids Nonflavanoid phenols Proanthocyanins
## [1,]     0.8067217  1.0319081           -0.6577078       1.2214385
## [2,]     0.5670481  0.7315653           -0.8184106      -0.5431887
## [3,]     0.8067217  1.2121137           -0.4970050       2.1299594
## [4,]     2.4844372  1.4623994           -0.9791134       1.0292513
## [5,]     0.8067217  0.6614853            0.2261576       0.4002753
## [6,]     1.5576991  1.3622851           -0.1755994       0.6623487
##      Color intensity        Hue OD280/OD315 of diluted wines     Proline
## [1,]       0.2510088  0.3611585                    1.8427215  1.01015939
## [2,]      -0.2924962  0.4049085                    1.1103172  0.96252635
## [3,]       0.2682629  0.3174085                    0.7863692  1.39122370
## [4,]       1.1827317 -0.4263410                    1.1807407  2.32800680
## [5,]      -0.3183774  0.3611585                    0.4483365 -0.03776747
## [6,]       0.7298108  0.4049085                    0.3356589  2.23274072
```

```
wine.pc <- prcomp(wine_s)
wine.pc
```

```
## Standard deviations (1, .., p=13):
##  [1] 2.1692972 1.5801816 1.2025273 0.9586313 0.9237035 0.8010350 0.7423128
##  [8] 0.5903367 0.5374755 0.5009017 0.4751722 0.4108165 0.3215244
##
## Rotation (n x k) = (13 x 13):
##                                       PC1          PC2         PC3         PC4
## Alcohol                      -0.144329395 -0.483651548 -0.20738262 -0.01785630
## Malic acid                    0.245187580 -0.224930935  0.08901289  0.53689028
## Ash                           0.002051061 -0.316068814  0.62622390 -0.21417556
## Alcalinity of ash             0.239320405  0.010590502  0.61208035  0.06085941
## Magnesium                    -0.141992042 -0.299634003  0.13075693 -0.35179658
## Total phenols                -0.394660845 -0.065039512  0.14617896  0.19806835
## Flavanoids                   -0.422934297  0.003359812  0.15068190  0.15229479
## Nonflavanoid phenols          0.298533103 -0.028779488  0.17036816 -0.20330102
## Proanthocyanins              -0.313429488 -0.039301722  0.14945431  0.39905653
## Color intensity               0.088616705 -0.529995672 -0.13730621  0.06592568
## Hue                          -0.296714564  0.279235148  0.08522192 -0.42777141
```

```
## OD280/OD315 of diluted wines -0.376167411  0.164496193  0.16600459  0.18412074
## Proline                      -0.286752227 -0.364902832 -0.12674592 -0.23207086
##                                      PC5         PC6         PC7         PC8
## Alcohol                       0.26566365 -0.21353865 -0.05639636 -0.39613926
## Malic acid                   -0.03521363 -0.53681385  0.42052391 -0.06582674
## Ash                           0.14302547 -0.15447466 -0.14917061  0.17026002
## Alcalinity of ash            -0.06610294  0.10082451 -0.28696914 -0.42797018
## Magnesium                    -0.72704851 -0.03814394  0.32288330  0.15636143
## Total phenols                 0.14931841  0.08412230 -0.02792498  0.40593409
## Flavanoids                    0.10902584  0.01892002 -0.06068521  0.18724536
## Nonflavanoid phenols          0.50070298  0.25859401  0.59544729  0.23328465
## Proanthocyanins              -0.13685982  0.53379539  0.37213935 -0.36822675
## Color intensity               0.07643678  0.41864414 -0.22771214  0.03379692
## Hue                           0.17361452 -0.10598274  0.23207564 -0.43662362
## OD280/OD315 of diluted wines  0.10116099 -0.26585107 -0.04476370  0.07810789
## Proline                       0.15786880 -0.11972557  0.07680450 -0.12002267
##                                      PC9        PC10        PC11        PC12
## Alcohol                      -0.50861912 -0.21160473  0.22591696  0.26628645
## Malic acid                    0.07528304  0.30907994 -0.07648554 -0.12169604
## Ash                           0.30769445  0.02712539  0.49869142  0.04962237
## Alcalinity of ash            -0.20044931 -0.05279942 -0.47931378  0.05574287
## Magnesium                    -0.27140257 -0.06787022 -0.07128891 -0.06222011
## Total phenols                -0.28603452  0.32013135 -0.30434119  0.30388245
## Flavanoids                   -0.04957849  0.16315051  0.02569409  0.04289883
## Nonflavanoid phenols         -0.19550132 -0.21553507 -0.11689586 -0.04235219
## Proanthocyanins               0.20914487 -0.13418390  0.23736257  0.09555303
## Color intensity              -0.05621752  0.29077518 -0.03183880 -0.60422163
## Hue                          -0.08582839  0.52239889  0.04821201 -0.25921400
## OD280/OD315 of diluted wines -0.13722690 -0.52370587 -0.04642330 -0.60095872
## Proline                       0.57578611 -0.16211600 -0.53926983  0.07940162
##                                     PC13
## Alcohol                      -0.01496997
## Malic acid                   -0.02596375
## Ash                           0.14121803
## Alcalinity of ash            -0.09168285
## Magnesium                    -0.05677422
## Total phenols                 0.46390791
## Flavanoids                   -0.83225706
## Nonflavanoid phenols         -0.11403985
## Proanthocyanins               0.11691707
## Color intensity               0.01199280
## Hue                           0.08988884
## OD280/OD315 of diluted wines  0.15671813
## Proline                      -0.01444734
```

```r
summary(wine.pc)
```

```
## Importance of components:
##                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
## Standard deviation     2.169 1.5802 1.2025 0.95863 0.92370 0.80103 0.74231
## Proportion of Variance 0.362 0.1921 0.1112 0.07069 0.06563 0.04936 0.04239
## Cumulative Proportion  0.362 0.5541 0.6653 0.73599 0.80162 0.85098 0.89337
##                            PC8     PC9   PC10    PC11    PC12    PC13
## Standard deviation     0.59034 0.53748 0.5009 0.47517 0.41082 0.32152
## Proportion of Variance 0.02681 0.02222 0.0193 0.01737 0.01298 0.00795
```

```
## Cumulative Proportion  0.92018 0.94240 0.9617 0.97907 0.99205 1.00000
```

Question 1 - part 2 - Reduce to Five PCs

The first 5 PCs, explain nearly 80.2% of the total variance. This means that the first five principal components can accurately represent the data. While **Cumulative Proportion** of PC1 to PC5 accounts for just over 80% of the total variance present in the dataset. This implies that the first five principal components are capable of accurately representing the data. Therefore, it can be concluded that these components are highly relevant and critical in understanding the dataset.

```
wine.5pc <- wine.pc$x[,1:5]
head(wine.5pc)
```

```
##            PC1        PC2        PC3        PC4        PC5
## [1,] -3.307421 -1.4394023 -0.1652728 -0.2150246 -0.6910933
## [2,] -2.203250  0.3324551 -2.0207571 -0.2905387  0.2569299
## [3,] -2.509661 -1.0282507  0.9800541  0.7228632  0.2503270
## [4,] -3.746497 -2.7486184 -0.1756962  0.5663856  0.3109644
## [5,] -1.006070 -0.8673840  2.0209873 -0.4086131 -0.2976180
## [6,] -3.041674 -2.1164309 -0.6276254 -0.5141870  0.6302409
```

Graph of Variables (optional)

In a **graph of variables**, the positioning of each variable can give us valuable insights into their relationships. When two variables are positively correlated, they tend to move in the same direction as each other. This results in their respective positions on the graph being closer together. Conversely, when two variables are negatively correlated, they move in opposite directions and are positioned farther apart on the graph. In the dataset, we can observe that *Proline* and *Total phenols* have a direct relationship, meaning when one variable increases, so does the other. On the other hand, *Flavanoids* and *Alcalinity of ash* have an inverse relationship, indicating that as one variable increases, the other decreases.

```
fviz_pca_var(wine.pc,
             col.var = "contrib",
             gradient.cols = c("deeppink1", "cyan3", "#FC4E07"),
             repel = TRUE)
```

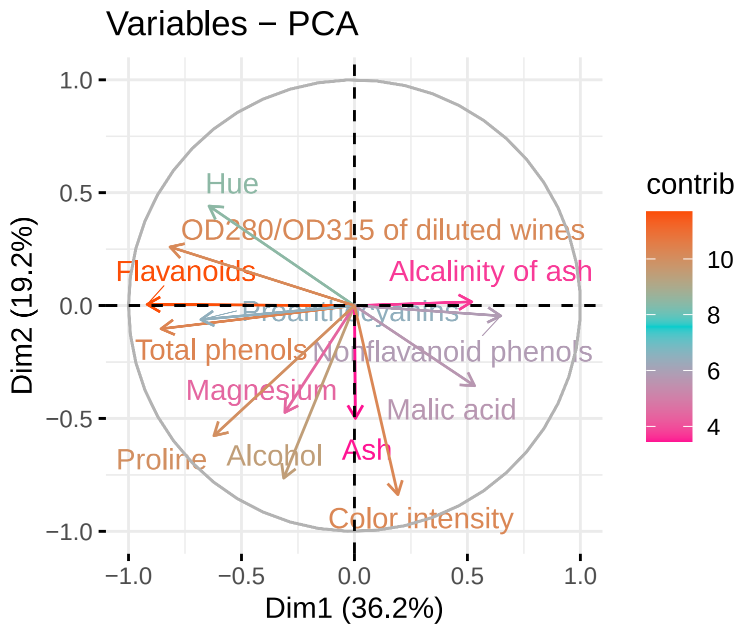

Figure 1: Graph of Variables

Biplot (optional)

A biplot is a combination of a score plot and a loadings (eigenvalues) plot, which are two common plots used in principal component analysis. Although these plots are on different scales, it is possible to rescale them and overlay them on a single plot. By selecting the appropriate scaling, the biplot can accurately show the relationship between variables or observations. Additionally, it can also provide approximate relationships between variables and observations.

The first principal component (PC1) explains 36.2% of the variability in the data, while the second principal component (PC2) explains 19.2%. Closer arrows indicate a stronger correlation between variables. For instance, the correlation between *Flavanoids* and *Malic acid* is weak.

```
library(devtools)
#install_github("vqv/ggbiplot")
library(ggbiplot)
```

```r
ggbiplot(wine.pc,
              obs.scale = 3,
              var.scale = 3,
              groups = winedata[,1],
              ellipse = TRUE,
              circle = TRUE,
              ellipse.prob = .68) +
  theme(legend.position = "none")
```

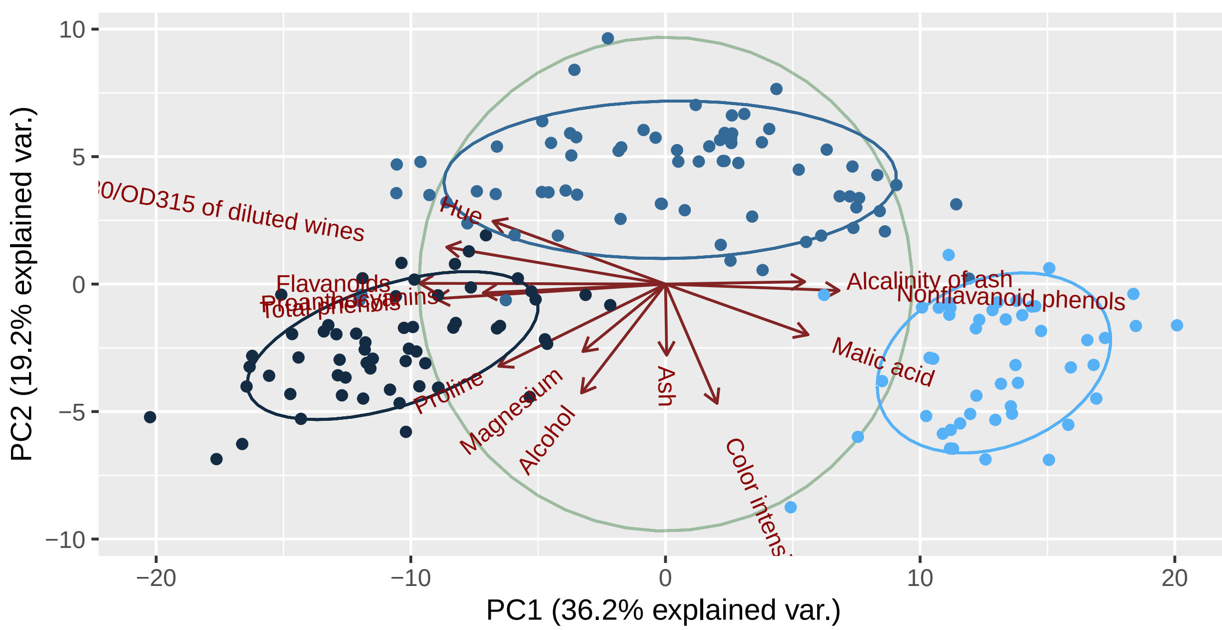

Figure 2: Biplot

Question 1-a - Scatterplot Matrix - PCA

The scatterplot matrix provides a visual representation of the pairwise relationship between the first 5 principal components, highlighting the distinction between the 3 cultivars through the use of different colors. Upon closer inspection, it is apparent that the projection along the first PC is particularly effective in accurately separating the three classes. So, yes we can see separation of the 3 cultivars. See figure 3.

```
pairs(wine.pc$x[,1:5],
      col=c("deeppink1", "cyan3", "tan2")[winedata[,1]], asp=1, pch = 21, cex=.5, lower.panel = NULL)
```

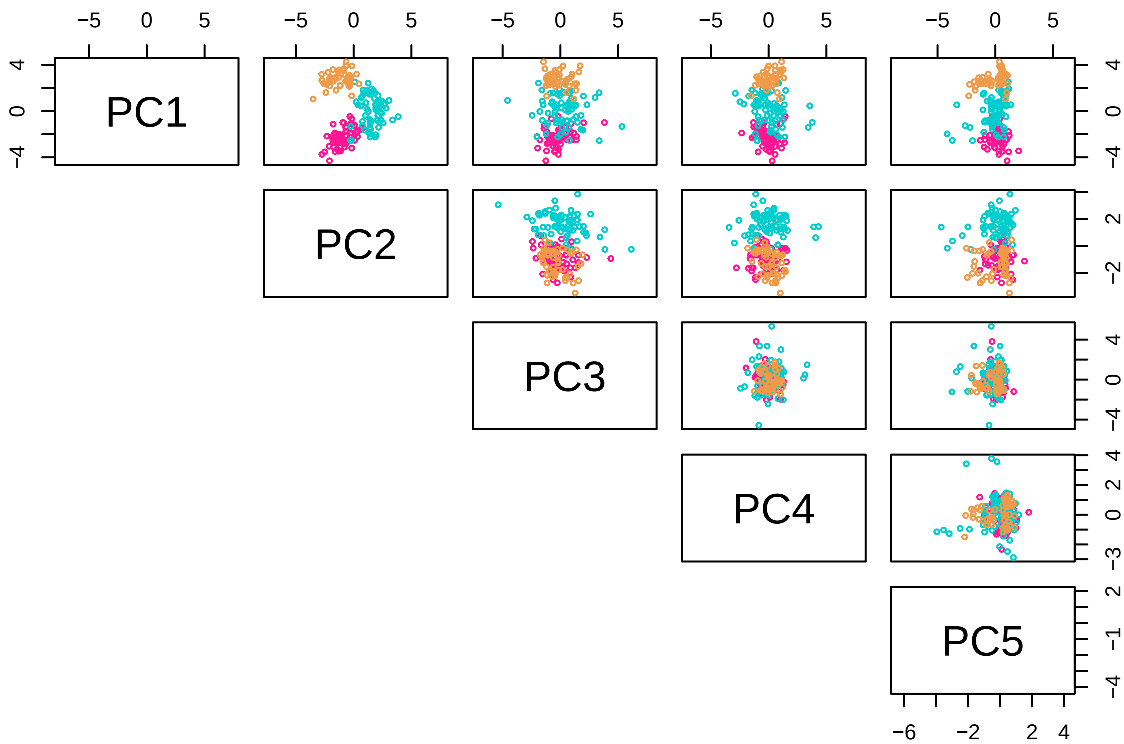

Figure 3: Scatter Plot of First 5 PCs - PCA

3D Scatterplot (optional)

```
library(scatterplot3d)
d1 <- scatterplot3d(wine.pc$x[,c(1,2,3)],
              color=c("deeppink1", "cyan3", "tan2")[winedata[,1]])
```

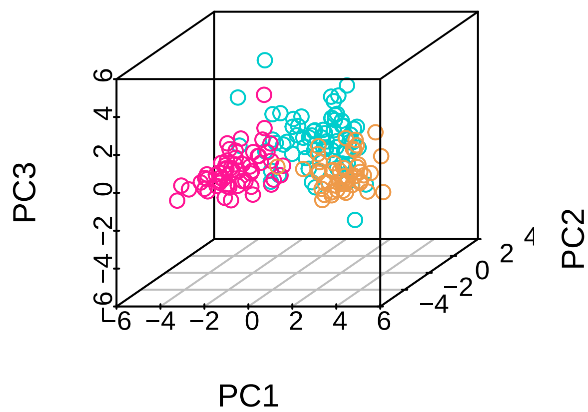

Figure 4: 3D Scatter Plot of PCs 1, 2 and 3

```r
d2 <- scatterplot3d(wine.pc$x[,c(1,2,4)],
             color=c("deeppink1", "cyan3", "tan2")[winedata[,1]])
```

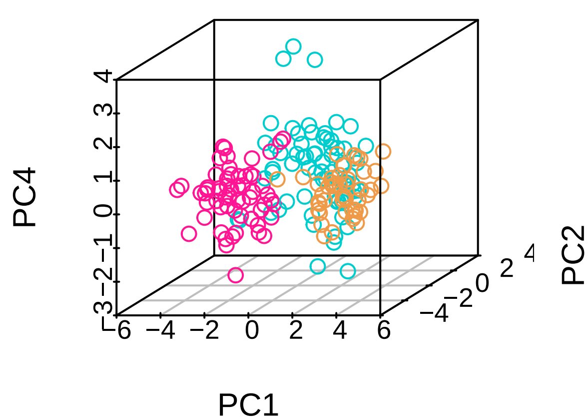

Figure 5: 3D Scatter Plot of PCs 1, 2 and 4

```
d3 <- scatterplot3d(wine.pc$x[,c(1,3,4)],
              color=c("deeppink1", "cyan3", "tan2")[winedata[,1]])
```

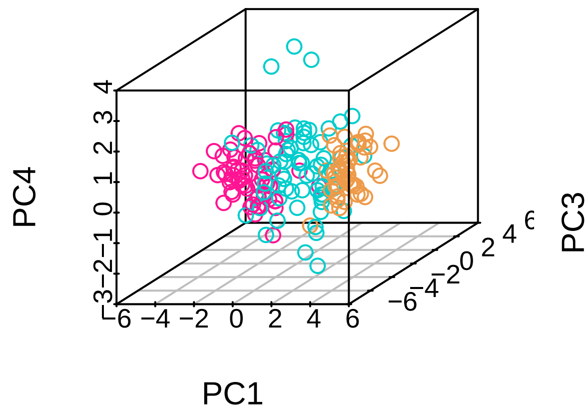

Figure 6: 3D Scatter Plot of PCs 1, 3 and 4

Scree Plot (optional)

This plot shows the eigenvalues in a downward curve, from highest to lowest.

```
fviz_eig(wine.pc, addlabels = TRUE)
```

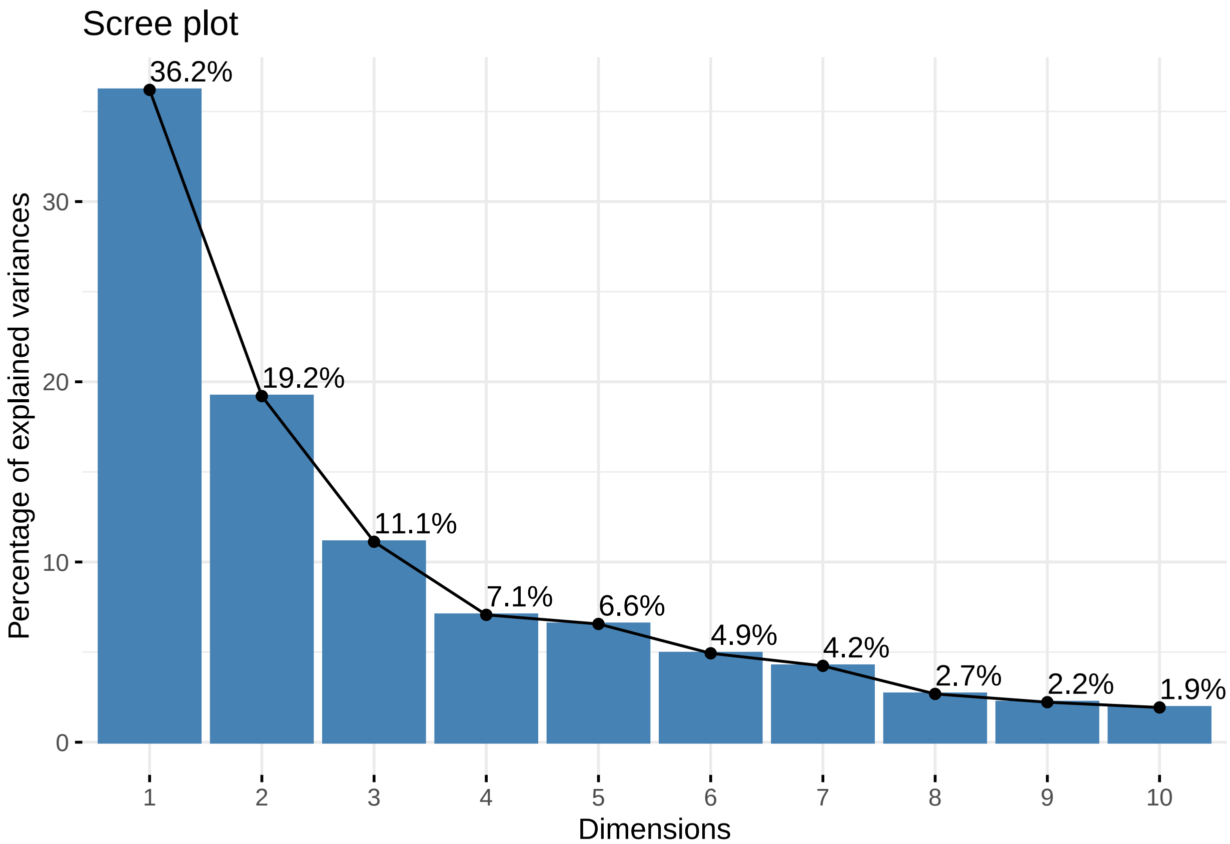

Figure 7: Scree plot

Question 1b - Interpert PC1 and PC2

The **Proportion of Variance** section reveals that the first principal component (PC1) is responsible for explaining almost 36% of the total variance while the second component explains 19.21% of the total variance. The third component explains approximately 11% and so on. The **Cumulative Proportion** of PC1 to PC2 accounts for just over 55% of the total variance present in the dataset.

```
summary(wine.pc)
```

```
## Importance of components:
##                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
## Standard deviation     2.169 1.5802 1.2025 0.95863 0.92370 0.80103 0.74231
## Proportion of Variance 0.362 0.1921 0.1112 0.07069 0.06563 0.04936 0.04239
## Cumulative Proportion  0.362 0.5541 0.6653 0.73599 0.80162 0.85098 0.89337
##                            PC8     PC9   PC10    PC11    PC12    PC13
## Standard deviation     0.59034 0.53748 0.5009 0.47517 0.41082 0.32152
## Proportion of Variance 0.02681 0.02222 0.0193 0.01737 0.01298 0.00795
## Cumulative Proportion  0.92018 0.94240 0.9617 0.97907 0.99205 1.00000
```

A higher absolute coefficient value indicates that the variable has a greater impact on the principal component. A coefficient value close to 1 or -1 indicates that the variable is closely related to the component. The sign (positive/negative) indicates the direction of the relationship between the variable and the component.

The first principal component has a relative value of coefficient with three variables (*Total phenols*, *Flavanoids*, and *OD280/OD315 of diluted wines*), in the negative direction. However, these values are lower than 0.5, indicating a relatively weak influence. The remaining variables have a small amount of coefficient, with most value values being less than 0.3, which implies that the PC1 component has no remarkable influence by the other variables.

In general:

PC1 = -0.144329395 * *Alcohol* +0.245187580 * *Malic acid* +0.002051061 * *Ash* +0.239320405 * *Alcalinity of ash* -0.141992042 * *Magnesium* -0.394660845 * *Total phenols* -0.422934297 * *Flavanoids* +0.298533103 * *Nonflavanoid phenols* -0.313429488 * *Proanthocyanins* +0.088616705 * *Color intensity* -0.296714564 * *Hue* -0.376167411 * *OD280/OD315 of diluted wines* -0.286752227 * *Proline*

In relation to the PC2, it shows a moderate coefficient value when considering two variables - *Alcohol* and *Color intensity* - in the positive direction. This implies that the PC2 component is not significantly influenced by the other variables.

In general:

PC1 = -0.483651548 * *Alcohol* -0.224930935 * *Malic acid* -0.316068814 * *Ash* +0.010590502 * *Alcalinity of ash* -0.299634003 * *Magnesium* -0.065039512 * *Total phenols* +0.003359812 * *Flavanoids* -0.028779488 * *Nonflavanoid phenols* -0.039301722 * *Proanthocyanins* -0.529995672 * *Color intensity* +0.279235148 * *Hue* +0.164496193 * *OD280/OD315 of diluted wines* -0.364902832 * *Proline*

```
wine.pc$rotation[,1:2]
```

```
##                                      PC1          PC2
## Alcohol                     -0.144329395 -0.483651548
## Malic acid                   0.245187580 -0.224930935
## Ash                          0.002051061 -0.316068814
## Alcalinity of ash            0.239320405  0.010590502
## Magnesium                   -0.141992042 -0.299634003
## Total phenols               -0.394660845 -0.065039512
## Flavanoids                  -0.422934297  0.003359812
## Nonflavanoid phenols         0.298533103 -0.028779488
## Proanthocyanins             -0.313429488 -0.039301722
```

```
## Color intensity               0.088616705 -0.529995672
## Hue                          -0.296714564  0.279235148
## OD280/OD315 of diluted wines -0.376167411  0.164496193
## Proline                      -0.286752227 -0.364902832
```

Question 2a - FastICA for 5 Independent Component (Independent Component Analysis)

FastICA is an algorithm that extracts information from complex data by performing independent component analysis. It identifies the independent components and represents the data accurately by finding an orthogonal rotation that maximizes non-Gaussianity.

Now we do fastICA for 5 independent component (n.comp = 5).

```
library(fastICA)
set.seed(777)
wine.ic <- fastICA(wine_s, n.comp = 5, alg.typ = "parallel", fun = "logcosh", alpha = 1,
method = "R", row.norm = FALSE, maxit=200, tol=0.0001, verbose=TRUE)
head(wine.ic$X)
```

```
##        Alcohol  Malic acid        Ash Alcalinity of ash  Magnesium
## [1,] 1.5143408 -0.56066822  0.2313998        -1.1663032 1.90852151
## [2,] 0.2455968 -0.49800856 -0.8256672        -2.4838405 0.01809398
## [3,] 0.1963252  0.02117152  1.1062139        -0.2679823 0.08810981
## [4,] 1.6867914 -0.34583508  0.4865539        -0.8069748 0.92829983
## [5,] 0.2948684  0.22705328  1.8352256         0.4506745 1.27837900
## [6,] 1.4773871 -0.51591132  0.3043010        -1.2860793 0.85828399
##      Total phenols Flavanoids Nonflavanoid phenols Proanthocyanins
## [1,]     0.8067217  1.0319081           -0.6577078       1.2214385
## [2,]     0.5670481  0.7315653           -0.8184106      -0.5431887
## [3,]     0.8067217  1.2121137           -0.4970050       2.1299594
## [4,]     2.4844372  1.4623994           -0.9791134       1.0292513
## [5,]     0.8067217  0.6614853            0.2261576       0.4002753
## [6,]     1.5576991  1.3622851           -0.1755994       0.6623487
##      Color intensity        Hue OD280/OD315 of diluted wines     Proline
## [1,]       0.2510088  0.3611585                    1.8427215  1.01015939
## [2,]      -0.2924962  0.4049085                    1.1103172  0.96252635
## [3,]       0.2682629  0.3174085                    0.7863692  1.39122370
## [4,]       1.1827317 -0.4263410                    1.1807407  2.32800680
## [5,]      -0.3183774  0.3611585                    0.4483365 -0.03776747
## [6,]       0.7298108  0.4049085                    0.3356589  2.23274072
```

```
wine.ic$K
```

```
##                [,1]         [,2]        [,3]        [,4]        [,5]
##  [1,] -0.066720472 -0.306936801  0.17294212 -0.01867942 -0.28841837
##  [2,]  0.113345109 -0.142746533 -0.07423031  0.56163908  0.03822976
##  [3,]  0.000948163 -0.200584804 -0.52222547 -0.22404832 -0.15527594
##  [4,]  0.110632836  0.006720985 -0.51043078  0.06366482  0.07176481
##  [5,] -0.065639962 -0.190154881 -0.10904183 -0.36801319  0.78932192
##  [6,] -0.182443484 -0.041275625 -0.12190269  0.20719862 -0.16210788
##  [7,] -0.195513712  0.002132217 -0.12565781  0.15931506 -0.11836416
##  [8,]  0.138005632 -0.018264149 -0.14207473 -0.21267250 -0.54358937
##  [9,] -0.144891921 -0.024941810 -0.12463409  0.41745167  0.14858219
## [10,]  0.040965656 -0.336347887  0.11450346  0.06896463 -0.08298377
## [11,] -0.137164960  0.177209281 -0.07106892 -0.44749020 -0.18848501
## [12,] -0.173894355  0.104393205 -0.13843583  0.19260808 -0.10982566
## [13,] -0.132559579 -0.231576036  0.10569693 -0.24276853 -0.17139063
```

```
wine.ic$W
```

```
##             [,1]        [,2]         [,3]        [,4]        [,5]
```

```
## [1,] -0.02182777  0.04144909 -0.002307059  0.03320545  0.99834743
## [2,] -0.05312866  0.95838271  0.270179335 -0.06501162 -0.03816478
## [3,]  0.02857207 -0.26763252  0.962698146  0.02441017  0.01314897
## [4,] -0.23530551 -0.08000661  0.008873717 -0.96810513  0.03039707
## [5,]  0.96980196  0.04190871 -0.011460419 -0.23842713  0.02736741
```

```
wine.ic$A
```

```
##             [,1]        [,2]         [,3]        [,4]       [,5]       [,6]
## [1,] -0.17888804 -0.08510683 -0.074679524  0.01220166  0.7559058 -0.1588659
## [2,] -0.81877642 -0.32883299 -0.265296793  0.23178355 -0.3682909 -0.1076055
## [3,]  0.03687303 -0.19555902 -0.857799513 -0.70346154 -0.2884618 -0.1912145
## [4,]  0.12011819 -0.46654762  0.243813906 -0.07265030  0.1825648 -0.1764732
## [5,] -0.28655689  0.55812379  0.003734297  0.50998809 -0.2825861 -0.8487198
##               [,7]        [,8]        [,9]       [,10]       [,11]      [,12]
## [1,] -0.1171222817 -0.41906681  0.04546203 -0.03821965 -0.07115292 -0.1334785
## [2,] -0.0003454507  0.03420228 -0.06472748 -0.84449038  0.44845000  0.2499763
## [3,] -0.1679654934 -0.20685665 -0.18575963 -0.06620371  0.02018263 -0.1170998
## [4,] -0.1521307819  0.31751148 -0.42222386  0.02045401  0.38159494 -0.1969075
## [5,] -0.9142784261  0.62523329 -0.66183605  0.22540491 -0.67573360 -0.8220886
##             [,13]
## [1,] -0.040390453
## [2,] -0.605793750
## [3,] -0.007904322
## [4,]  0.269933200
## [5,] -0.606056071
```

```
head(wine.ic$S)
```

```
##            [,1]        [,2]       [,3]        [,4]       [,5]
## [1,]  0.8664022 -0.92628276 -0.1211859  0.05085497 -1.4760541
## [2,] -0.1398251 -0.27839163  1.6821585  0.35434221 -1.0195997
## [3,] -0.4048544 -0.52663740 -0.9506125 -0.68331996 -1.1285404
## [4,] -0.3321567 -1.84429567 -0.3171094 -0.43363430 -1.6517957
## [5,]  0.4051764 -0.04803493 -1.7776370  0.31598052 -0.4696198
## [6,] -0.4199864 -1.47123892  0.1472950  0.73727602 -1.3807117
```

Question 2b - Scatterplot Matrix - ICA

The scatterplot matrix provides a visual representation of the pairwise relationship between the first 5 independent variables, highlighting the distinction between the 3 cultivars through the use of different colors. Upon closer inspection, it is apparent that the projection along the *var 5* is particularly effective in accurately separating the three classes. So, yes we can see separation of the 3 cultivars. See figure 8.

```
pairs(wine.ic$S, col=c("deeppink1", "cyan3", "tan2")[winedata[,1]], lower.panel = NULL)
```

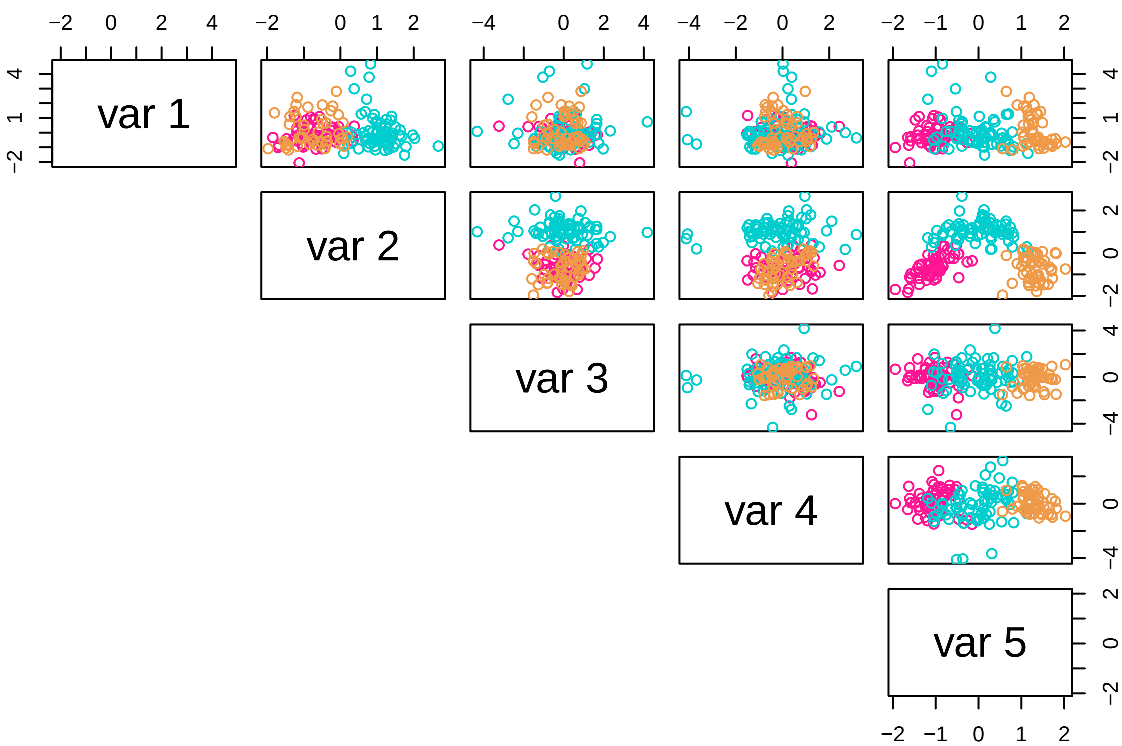

Figure 8: Scatter Plot of 5 PCs - ICA

3D Scatterplot (optional)

```
library(scatterplot3d)
dd1 <- scatterplot3d(wine.ic$S[,c(5,3,1)],
              color=c("deeppink1", "cyan3", "tan2")[winedata[,1]],
              xlab = "var 5", ylab = "var 3", zlab = "var 1")
```

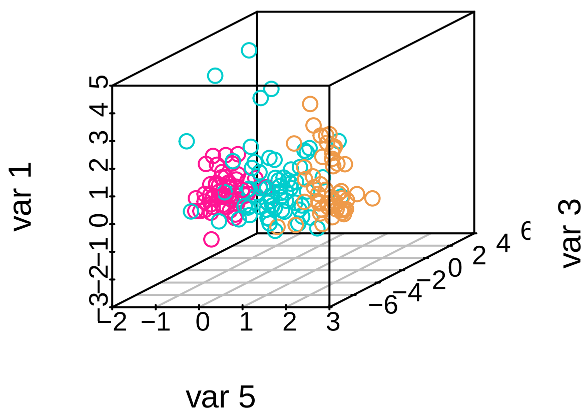

Figure 9: 3D Scatter Plot of Variables 1, 3 and 5

```
dd2 <- scatterplot3d(wine.ic$S[,c(5,2,3)],
              color=c("deeppink1", "cyan3", "tan2")[winedata[,1]],
              xlab = "var 5", ylab = "var 2", zlab = "var 3")
```

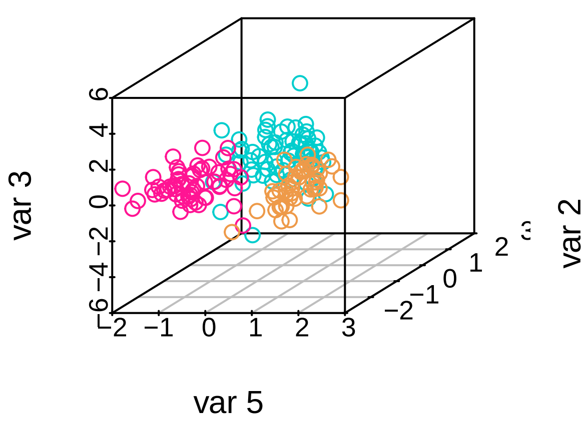

Figure 10: 3D Scatter Plot of Variables 2, 3 and 5

```
dd3 <- scatterplot3d(wine.ic$S[,c(5,4,3)],
              color=c("deeppink1", "cyan3", "tan2")[winedata[,1]],
              xlab = "var 5", ylab = "var 4", zlab = "var 3")
```

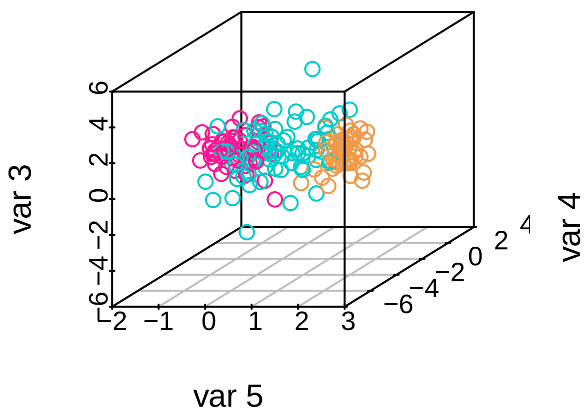

Figure 11: 3D Scatter Plot of Variables 3, 4 and 5