

Multivariate Statistical Analysis

Homework Assignment 9: Linear Discriminant Analysis

Ari Viitala

```
In [149]: #importing Maas-library
library(MASS)

In [150]: #importing the data
data <- read.table("ALCOHOL.txt", header = T, sep = "\t")
#doing the linear discriminant analysis
L = lda(TYPE~., data)
```

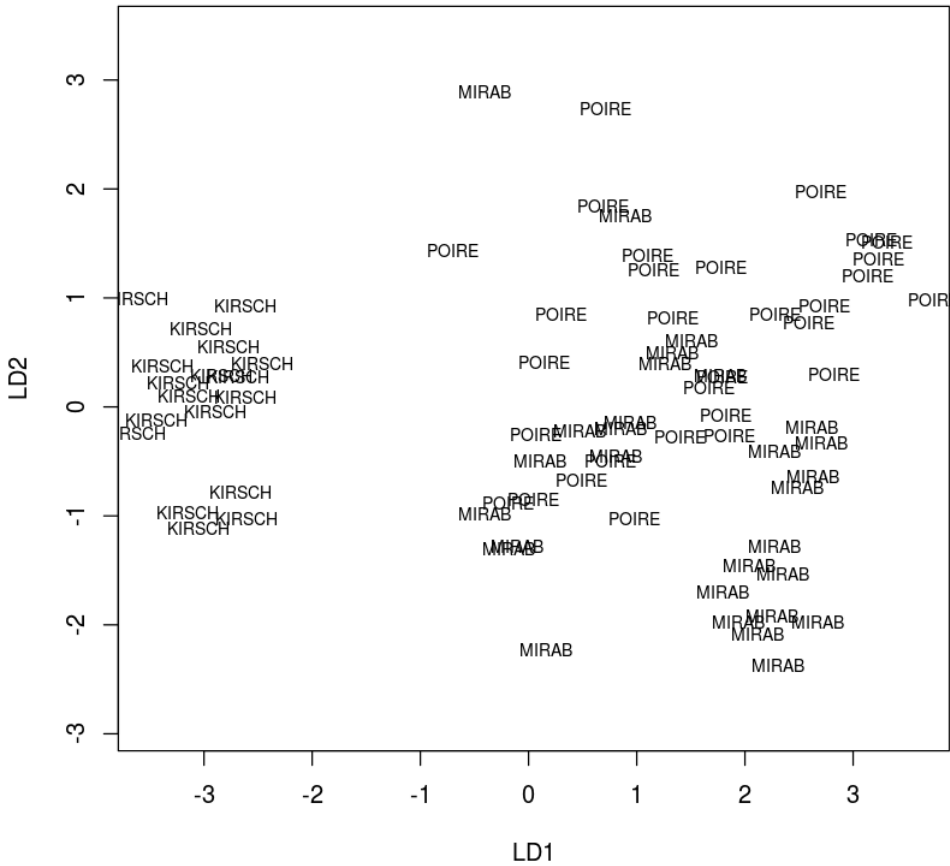
(a) Give the vector **a**, such that $\|a\|_2 = 1$

```
In [151]: #extracting the vector a from the LDA fit
a = L$scaling
#scaling the first column which is the Fischer
scaled_a = a[,1] / norm(a[,1], type = "2")
#printing the vector
scaled_a
```

MEOH	0.0108184342367015
ACET	-0.000148717517930887
BU1	0.422889085014792
MEPR	-0.081959976483502
ACAL	-0.129509361437617
LNPRO1	-0.893060831463324

Plotting the LDA model to see how it has divided the drinks

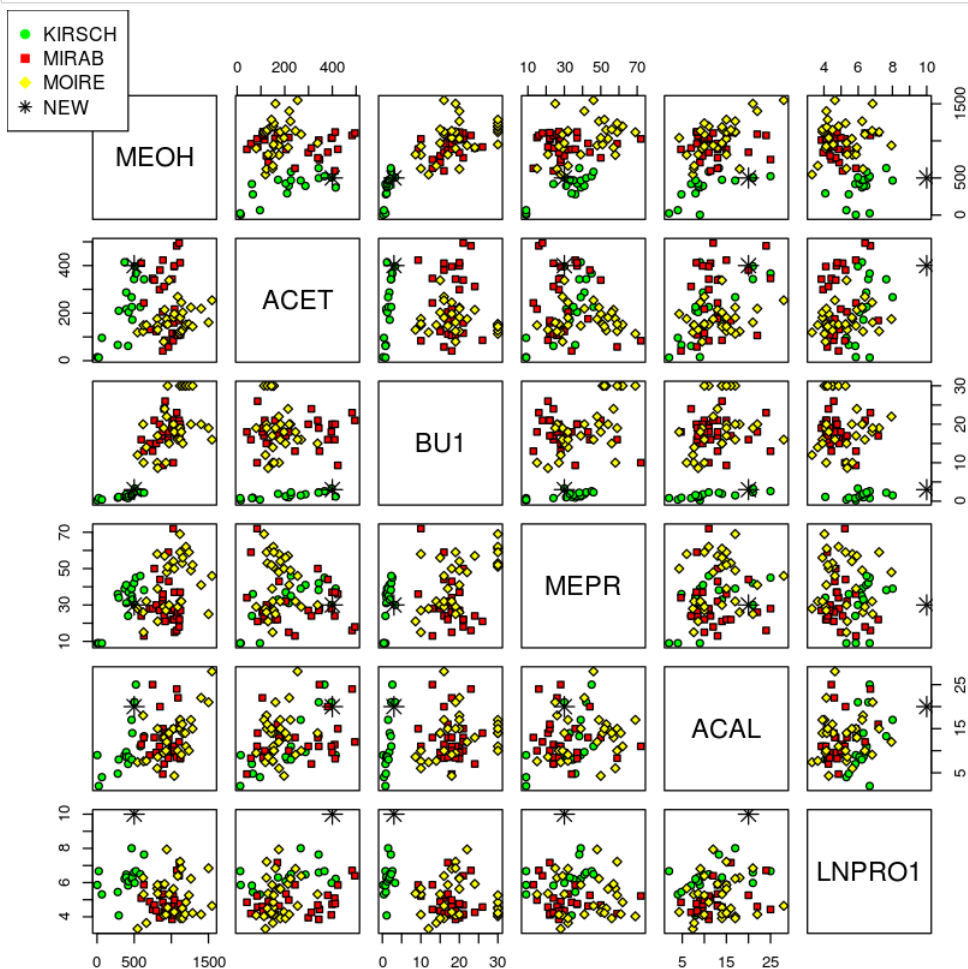
```
In [152]: plot(L)
```



Kirsch seems to be the most distinguished group

(b) Visualize the original data using a pairwise scatterplot

```
In [153]: #Importing the new sample from a text file since it was more convinient
new = read.table("NEW.txt", header = T, sep = "\t")
#Concatenating the data frames
all = rbind(data, new)
#doing the pairs plot
pairs(all[,2:7], pch = c(21, 22, 23, 8)[unclass(lors$TYPE)],
      bg = c("green", "red", "yellow", "black")[unclass(all$TYPE)],
      cex = c(1, 1, 1, 2)[unclass(all$TYPE)])
par(fig=c(0, 1, 0, 1), oma=c(0, 0, 0, 0), mar=c(0, 0, 0, 0), new=TRUE)
plot(0, 0, type='n', bty='n', xaxt='n', yaxt='n')
legend("topleft", legend = c("KIRSCH", "MIRAB", "MOIRE", "NEW"),
      pch = c(21, 22, 23, 8), col = c("green", "red", "yellow", "black"),
      pt.bg = c("green", "red", "yellow", "black"), bty = 'o', cex = 0.8)
```



Based on the plots it would seem that the new drink belongs to the Kirsch group.

(c) In which group will the new drink be classified to?

```
In [154]: #predicting with the model
predict(L, new)
```

\$class

KIRSCH

\$posterior

	KIRSCH	MIRAB	POIRE
1	0.9999964	1.81432e-06	1.806996e-06

\$x

	LD1	LD2
1	-4.198017	0.4860258

The model would also suggest that the new sample belongst to the Kirsch group with a really high probability.