Linear Discriminant Analysis

Nik Bear Brown

In this lesson we'll learn the theory behind using Linear Discriminant Analysis (LDA) as a supervised classification technique. We'll then use LDA to classify the UCI wine dataset in R.

Additional packages needed

To run the code you may need additional packages.

- If necessary install the followings packages.

```
install.packages("ggplot2");
install.packages("MASS");
install.packages("car");
```

```
require(ggplot2)

## Loading required package: ggplot2

require(MASS)

## Loading required package: MASS

require(car)

## Loading required package: car
```

Data

We will be using the UCI Machine Learning Repository: Wine Data Set. These data are the results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The analysis determined the quantities of 13 constituents found in each of the three types of wines.

The attributes are:
1) Alcohol
2) Malic acid
3) Ash
4) Alcalinity of ash
5) Magnesium
6) Total phenols
7) Flavanoids
8) Nonflavanoid phenols

9) Proanthocyanins
10) Color intensity
11) Hue
12) OD280/OD315 of diluted wines
13) Proline

Feel free to tweet questions to [@NikBearBrown](https://twitter.com/NikBearBrown)

```
# Load our data
data_url <-
'http://nikbearbrown.com/YouTube/MachineLearning/M07/wine.csv'
wn <- read.csv(url(data_url))
head(wn)
```

```
##   Cultivar Alcohol Malic.acid  Ash Alcalinity.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 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
```

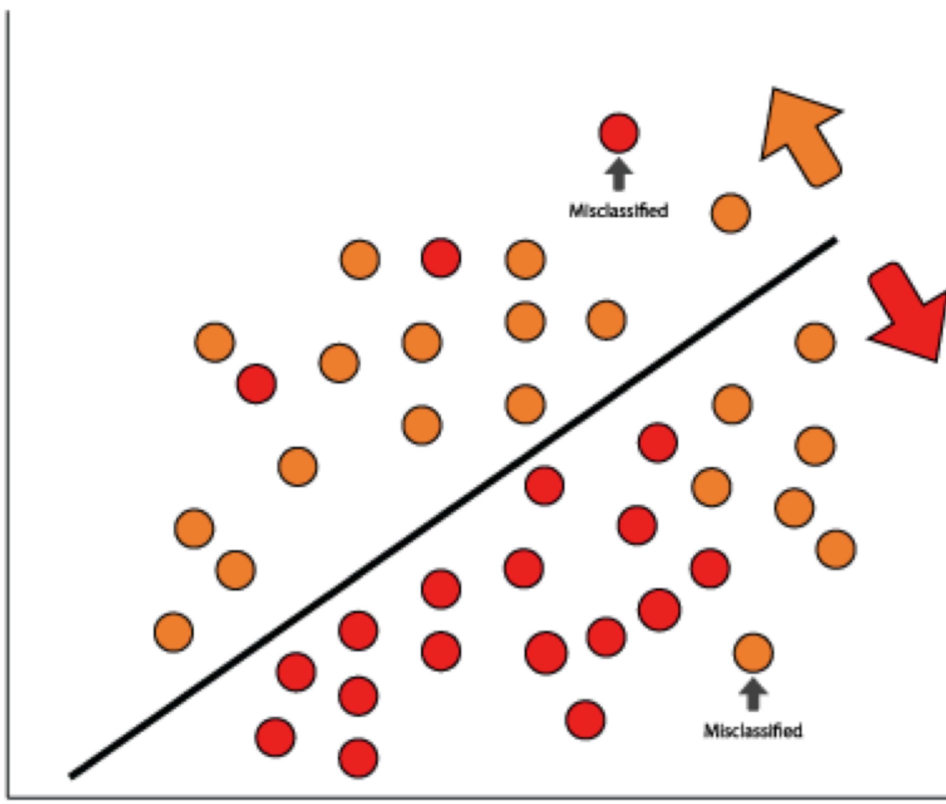
Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a generalization of Fisher's linear discriminant to find a linear combination of features that characterizes or separates two or more classes of objects or events. Discriminant analysis seeks to generate lines that are efficient for discrimination.

LDA is also closely related to principal component analysis (PCA) and factor analysis in that they both look for linear combinations of variables which best explain the data. In the case of LDA, we are maximizing the linear compenent axes for class discrimination. In the case of PCA, we are finding basis that maximize the variance.

LDA can also be used as a supervised technique by finding a discriminant projection that maximizing between-class distance and minimizing within-class distance.

LDA classifies n items $X = x_1, \ldots x_n$ to one of G groups based on measurements on p predictors. Similar to linear regression except our line(s) act to seperate groups.



Linear Discriminants seperate groups

Linear Discriminants seperate groups

LDA for two classes

Consider a set of observations $\vec{x}$ and a known class y

LDA approaches the problem by assuming that the conditional probability density functions $p(\vec{x}\,|y=0)$ and $p(\vec{x}\,|y=1)$ are both normally distributed with mean and covariance parameters $(\vec{\mu}_0, \Sigma_0)$ and $(\vec{\mu}_1, \Sigma_1)$

LDA instead makes the additional simplifying homoscedasticity assumption (i.e. that the class covariances are identical, so $(\Sigma_0 = \Sigma_1 = \Sigma)$ and that the covariances have full rank.

In this case, several terms cancel:

$$\vec{x}^T\,\Sigma_0^{-1}\,\vec{x} = \vec{x}^T\,\Sigma_1^{-1}\,\vec{x}\,\vec{x}^T\,\Sigma_i{}^{-1}\,\overrightarrow{\mu_i} = \overrightarrow{\mu_i}^T\,\Sigma_i{}^{-1}\,\vec{x}$$

because Σ_i is Hermitian (i.e. a square matrix with complex entries that is equal to its own conjugate transpose) and the above decision criterion becomes a threshold on the dot product $\vec{w}\cdot\vec{x} > c$ for some threshold constant c, where $\vec{w} = \Sigma^{-1}(\vec{\mu}_1 - \vec{\mu}_0)c = \frac{1}{2}(T - \overrightarrow{\mu_0}^T\,\Sigma_0^{-1}\,\overrightarrow{\mu_0} + \overrightarrow{\mu_1}^T\,\Sigma_1^{-1}\,\overrightarrow{\mu_1})$

This means that the criterion of an input $\vec{x}$ being in a class y is purely a function of this linear combination of the known observations. That is the $\vec{x}$ postion is classified by n-lines and its postion in n-dimensional space determines its class.

Fisher's linear discriminant

Suppose two classes of observations have means $\vec{\mu}_0, \vec{\mu}_1$ and covariances Σ_0, Σ_1 . Then the linear combination of features \$w x \$ will have means $\vec{w}\cdot\overrightarrow{\mu_i}$ and variances $\vec{w}^T\,\Sigma_i\,\vec{w}$ for i=0,1 . Fisher defined the separation between these two distributions to be the ratio of the variance between the classes to the variance within the classes:

$$S = \frac{\sigma^2_{\text{between}}}{\sigma^2_{\text{within}}} = \frac{(\vec{w}\cdot\vec{\mu}_1 - \vec{w}\cdot\vec{\mu}_0)^2}{\vec{w}^T\,\Sigma_1\,\vec{w} + \vec{w}^T\,\Sigma_0\,\vec{w}} = \frac{(\vec{w}\cdot(\vec{\mu}_1 - \vec{\mu}_0))^2}{\vec{w}^T\,(\Sigma_0 + \Sigma_1)\,\vec{w}}$$

This measure is, in some sense, a measure of the signal-to-noise ratio for the class labelling. It can be shown that the maximum separation occurs when $\vec{w} \propto (\Sigma_0 + \Sigma_1)^{-1}(\vec{\mu}_1 - \vec{\mu}_0)$

When the assumptions of LDA are satisfied, the above equation is equivalent to LDA.

Multiclass LDA

In the case where there are more than two classes, the analysis used in the derivation of the Fisher discriminant can be extended to find a subspace which appears to contain all of the class variability. This generalization is due to CR. Rao.Suppose that each of C classes has a mean μ_i and the same covariance Σ. Then the scatter between class variability may be defined by the sample covariance of the class means $\Sigma_b = \frac{1}{C}\sum_{i=1}^{C}(\mu_i - \mu)(\mu_i - \mu)^T$ where μ is the mean of the class means. The class separation in a direction $\vec{w}$ in this case will be given by

$$S = \frac{\vec{w}^T \, \Sigma_b \, \vec{w}}{\vec{w}^T \, \Sigma \, \vec{w}}$$

This means that when $\vec{w}$ is an eigenvector of $\Sigma^{-1}\Sigma_b$ the separation will be equal to the corresponding eigenvalue. If $\Sigma^{-1}\Sigma_b$ is diagonalizable, the variability between features will be contained in the subspace spanned by the eigenvectors corresponding to the C − 1 largest eigenvalues (since Σ_b is of rank C − 1 at most). These eigenvectors are primarily used in feature reduction, as in PCA. The eigenvectors corresponding to the smaller eigenvalues will tend to be very sensitive to the exact choice of training data, and it is often necessary to use regularization.

Linear Discriminant Analysis in R

LDA function ... outcome must be categories

```
head(wn)
```

```
##   Cultivar Alcohol Malic.acid  Ash Alcalinity.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 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
summary(wn)
```

```
##     Cultivar        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.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       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
```

```r
length(wn)
```

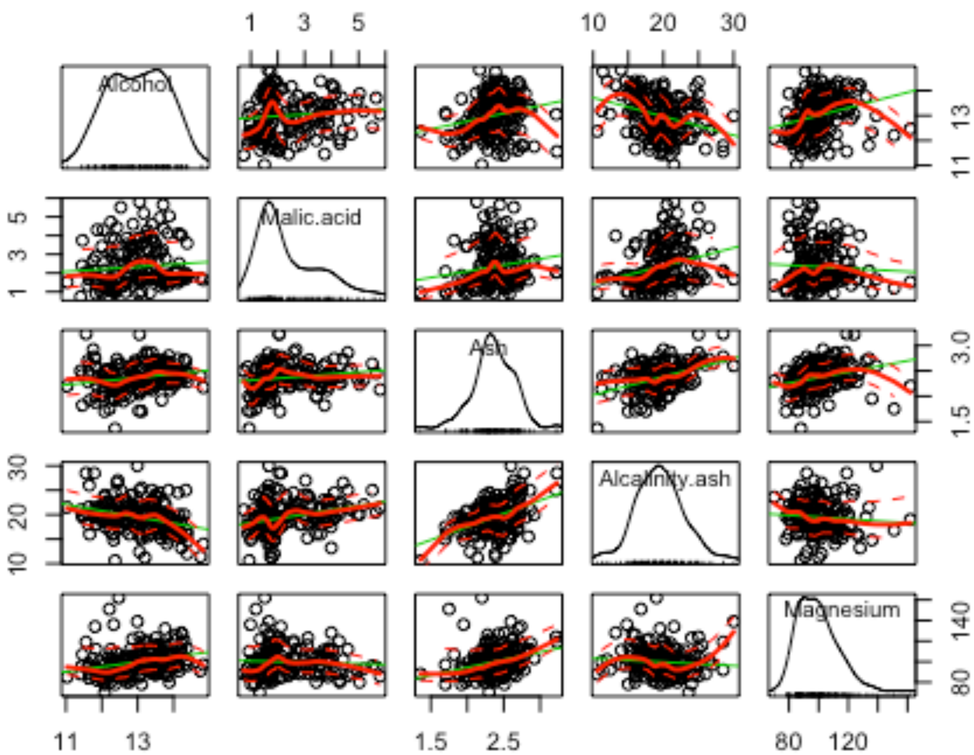
```
## [1] 14
```

You can also embed plots, for example:

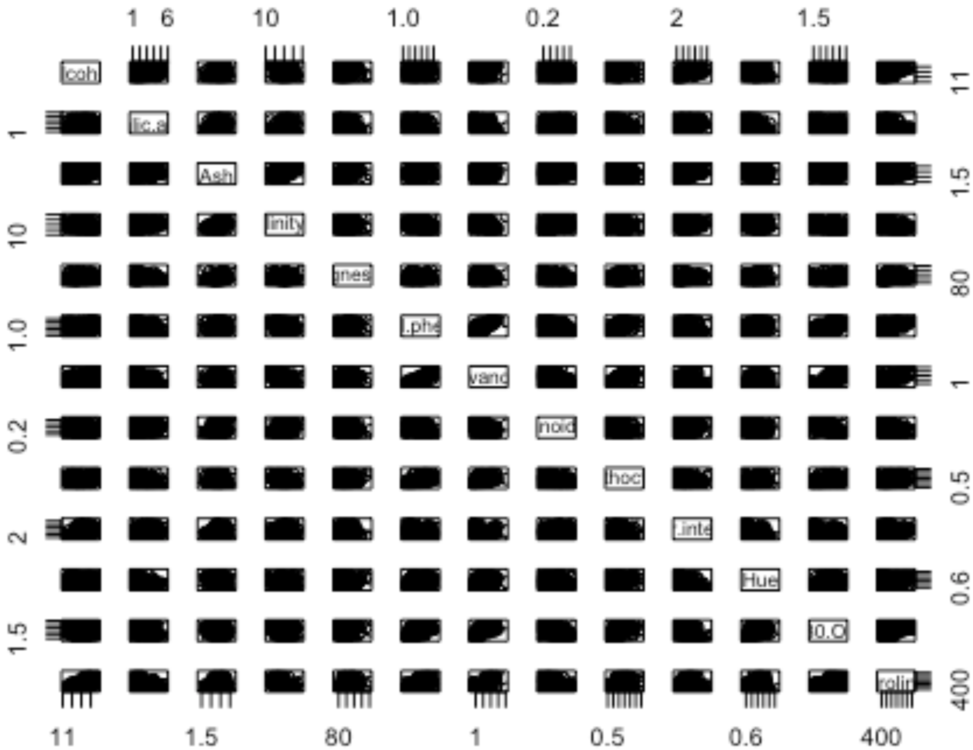
```r
names(wn)
```

```
##  [1] "Cultivar"             "Alcohol"              "Malic.acid"
##  [4] "Ash"                  "Alcalinity.ash"       "Magnesium"
##  [7] "Total.phenols"        "Flavanoids"
"Nonflavanoid.phenols"
## [10] "Proanthocyanins"      "Color.intensity"      "Hue"
## [13] "OD280.OD315"          "Proline"
```

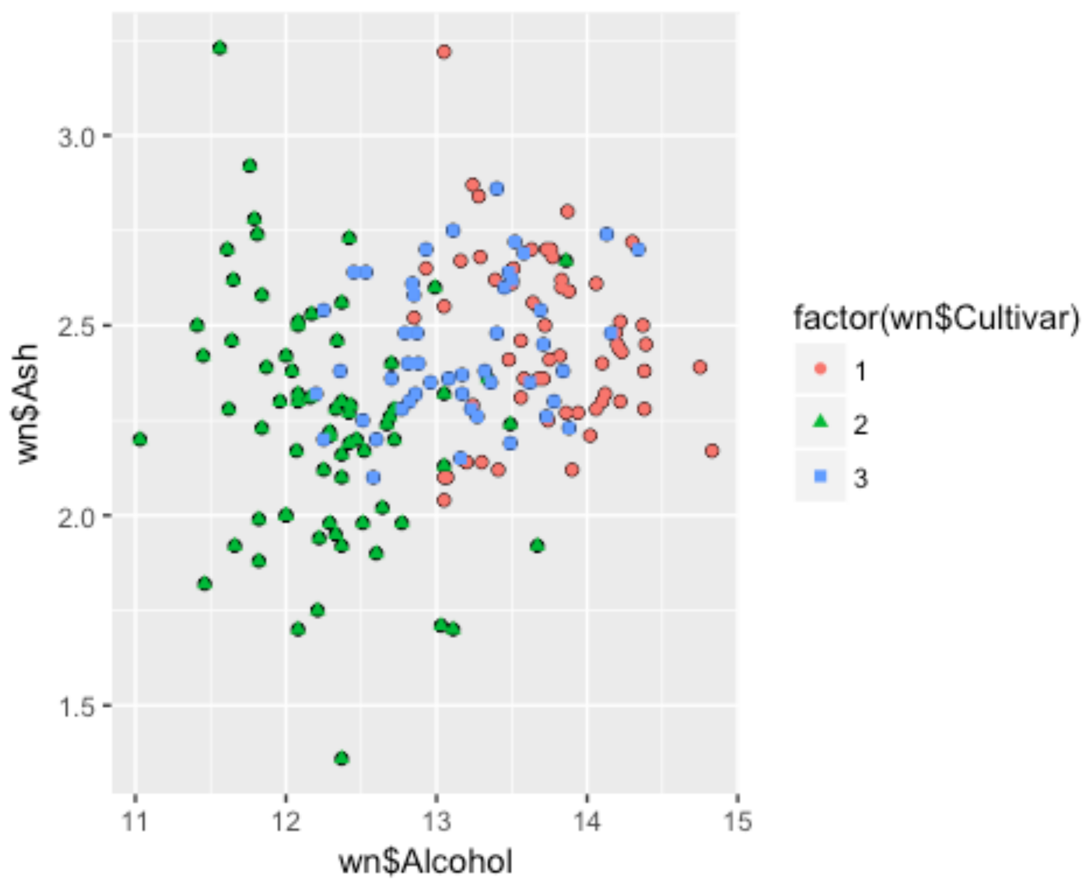
```r
scatterplotMatrix(wn[2:6])
```



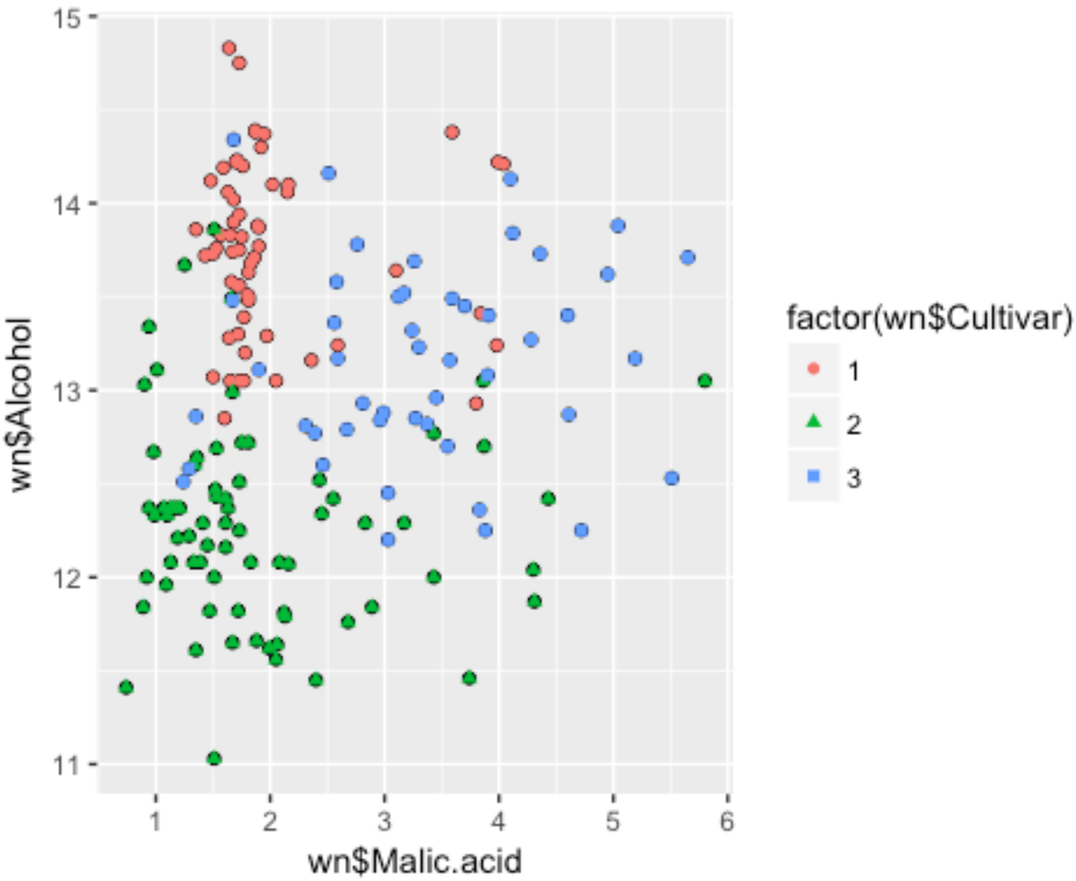

```r
pairs(wn[,2:14])
```



```
qplot(wn$Alcohol,wn$Ash,data=wn)+geom_point(aes(colour =
factor(wn$Cultivar),shape = factor(wn$Cultivar)))
```




```
qplot(wn$Malic.acid,wn$Alcohol,data=wn)+geom_point(aes(colour =
factor(wn$Cultivar),shape = factor(wn$Cultivar)))
```



You can also embed plots, for example:

```
lsa.m1<-lda(Cultivar ~ Malic.acid + Alcohol, data=wn)
lsa.m1

## Call:
## lda(Cultivar ~ Malic.acid + Alcohol, data = wn)
##
## Prior probabilities of groups:
##         1         2         3
## 0.3314607 0.3988764 0.2696629
##
## Group means:
##   Malic.acid  Alcohol
## 1   2.010678 13.74475
## 2   1.932676 12.27873
## 3   3.333750 13.15375
##
## Coefficients of linear discriminants:
##                   LD1        LD2
## Malic.acid -0.1258716  1.0541258
## Alcohol    -1.9357609 -0.2644917
##
```

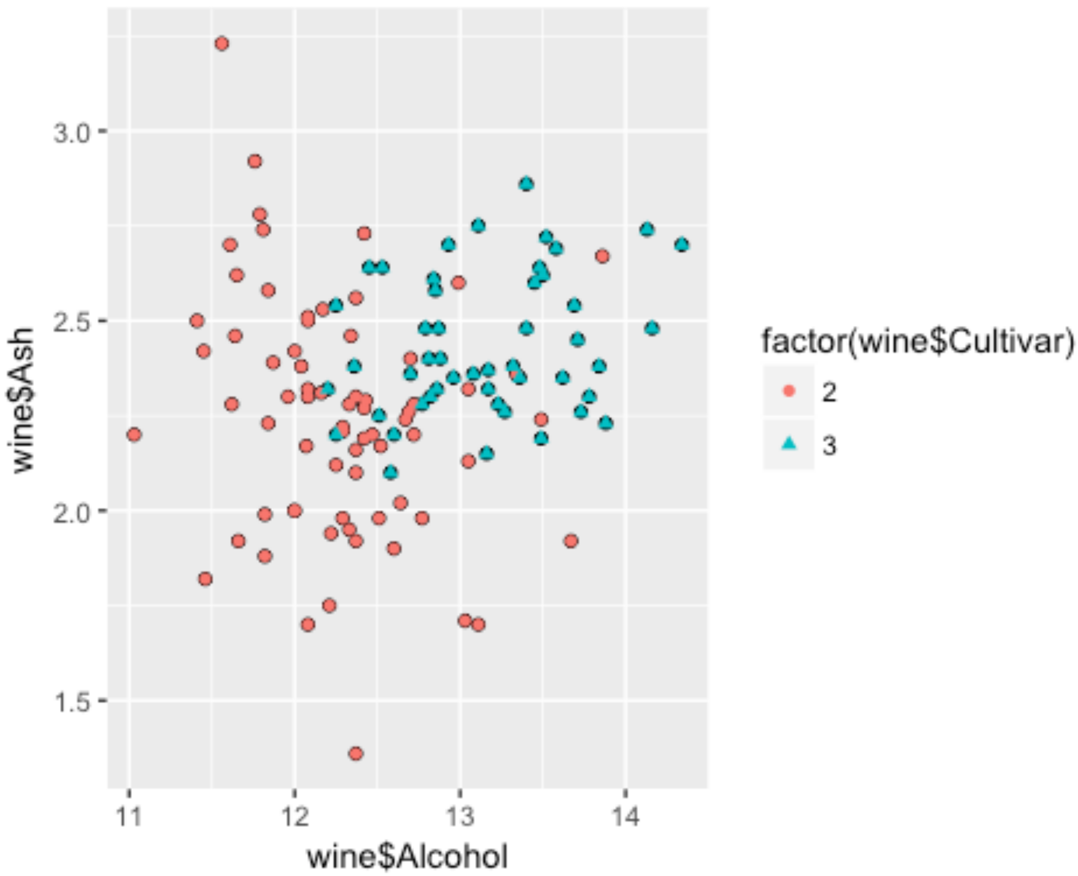
```
## Proportion of trace:
##    LD1    LD2
## 0.7955 0.2045
```

You can also embed plots, for example:

```
wine<-wn[which(wn$Cultivar!=1),]
head(wine)

##    Cultivar Alcohol Malic.acid  Ash Alcalinity.ash Magnesium
Total.phenols
## 60        2   12.37       0.94 1.36           10.6        88
1.98
## 61        2   12.33       1.10 2.28           16.0       101
2.05
## 62        2   12.64       1.36 2.02           16.8       100
2.02
## 63        2   13.67       1.25 1.92           18.0        94
2.10
## 64        2   12.37       1.13 2.16           19.0        87
3.50
## 65        2   12.17       1.45 2.53           19.0       104
1.89
##    Flavanoids Nonflavanoid.phenols Proanthocyanins Color.intensity
Hue
## 60       0.57                 0.28            0.42            1.95
1.05
## 61       1.09                 0.63            0.41            3.27
1.25
## 62       1.41                 0.53            0.62            5.75
0.98
## 63       1.79                 0.32            0.73            3.80
1.23
## 64       3.10                 0.19            1.87            4.45
1.22
## 65       1.75                 0.45            1.03            2.95
1.45
##    OD280.OD315 Proline
## 60        1.82     520
## 61        1.67     680
## 62        1.59     450
## 63        2.46     630
## 64        2.87     420
## 65        2.23     355

summary(wine)

##     Cultivar        Alcohol        Malic.acid        Ash
##  Min.   :2.000   Min.   :11.03   Min.   :0.740   Min.   :1.360
##  1st Qu.:2.000   1st Qu.:12.16   1st Qu.:1.490   1st Qu.:2.195
##  Median :2.000   Median :12.52   Median :2.160   Median :2.320
```

```
##  Mean   :2.403   Mean   :12.63   Mean   :2.498   Mean   :2.322
##  3rd Qu.:3.000   3rd Qu.:13.11   3rd Qu.:3.400   3rd Qu.:2.500
##  Max.   :3.000   Max.   :14.34   Max.   :5.800   Max.   :3.230
##  Alcalinity.ash    Magnesium      Total.phenols     Flavanoids
##  Min.   :10.60   Min.   : 70.00   Min.   :0.980   Min.   :0.340
##  1st Qu.:18.90   1st Qu.: 86.00   1st Qu.:1.615   1st Qu.:0.760
##  Median :20.50   Median : 93.00   Median :1.950   Median :1.500
##  Mean   :20.71   Mean   : 96.47   Mean   :2.025   Mean   :1.557
##  3rd Qu.:22.50   3rd Qu.:102.50   3rd Qu.:2.420   3rd Qu.:2.135
##  Max.   :30.00   Max.   :162.00   Max.   :3.520   Max.   :5.080
##  Nonflavanoid.phenols Proanthocyanins Color.intensity       Hue
##  Min.   :0.1300       Min.   :0.410   Min.   : 1.280   Min.
:0.4800
##  1st Qu.:0.2900       1st Qu.:1.035   1st Qu.: 2.800   1st
Qu.:0.7000
##  Median :0.4000       Median :1.400   Median : 3.800   Median
:0.8900
##  Mean   :0.3975       Mean   :1.438   Mean   : 4.825   Mean
:0.9056
##  3rd Qu.:0.5000       3rd Qu.:1.735   3rd Qu.: 5.940   3rd
Qu.:1.0650
##  Max.   :0.6600       Max.   :3.580   Max.   :13.000   Max.
:1.7100
##   OD280.OD315       Proline
##  Min.   :1.270   Min.   :278.0
##  1st Qu.:1.720   1st Qu.:450.0
##  Median :2.300   Median :560.0
##  Mean   :2.341   Mean   :564.0
##  3rd Qu.:2.960   3rd Qu.:673.5
##  Max.   :3.690   Max.   :985.0
```

```
qplot(wine$Alcohol,wine$Ash,data=wine)+geom_point(aes(colour =
factor(wine$Cultivar),shape = factor(wine$Cultivar)))
```

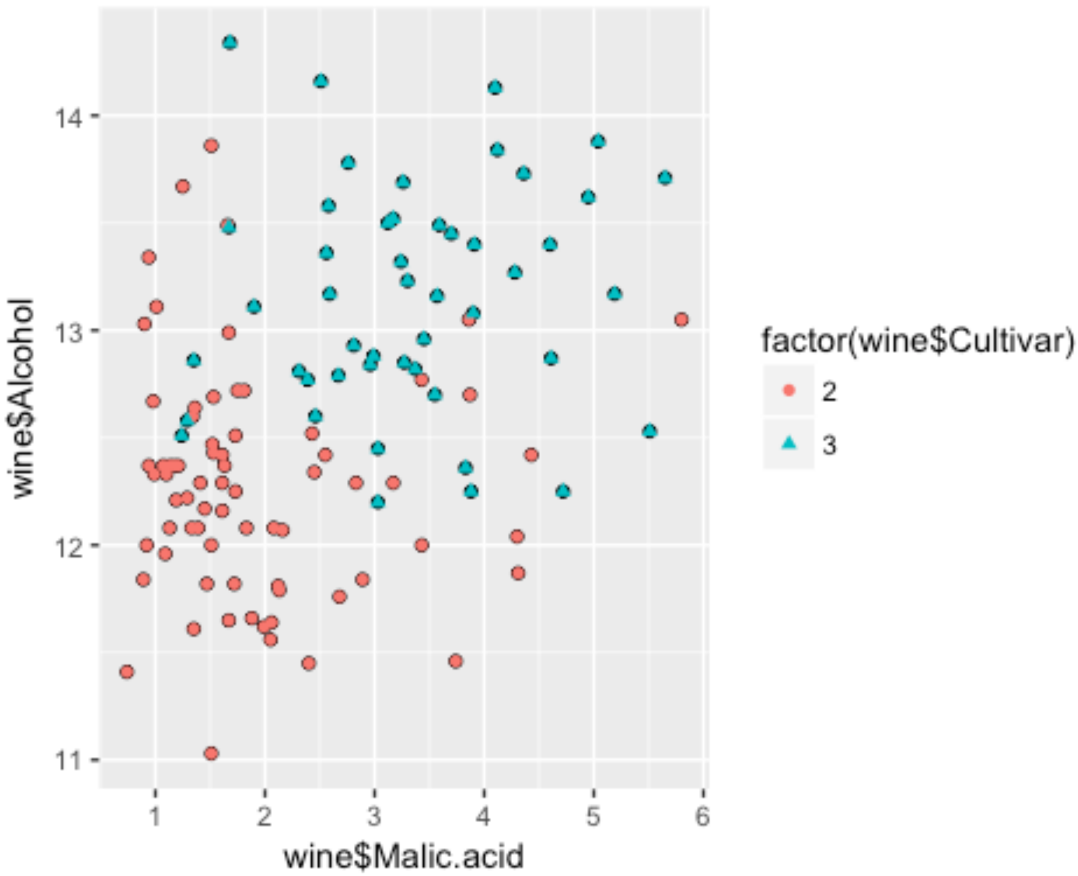


You can also embed plots, for example:

```
lsa.m2<-lda(Cultivar ~ Alcohol + Ash, data=wine)
lsa.m2

## Call:
## lda(Cultivar ~ Alcohol + Ash, data = wine)
##
## Prior probabilities of groups:
##         2         3
## 0.5966387 0.4033613
##
## Group means:
##    Alcohol      Ash
## 2 12.27873 2.244789
## 3 13.15375 2.437083
##
## Coefficients of linear discriminants:
##              LD1
## Alcohol 1.731380
## Ash     1.711451
```

You can also embed plots, for example:

```r
qplot(wine$Malic.acid,wine$Alcohol,data=wine)+geom_point(aes(colour =
factor(wine$Cultivar),shape = factor(wine$Cultivar)))
```




```r
lsa.m3<-lda(Cultivar ~ Malic.acid + Alcohol, data=wine)
lsa.m3

## Call:
## lda(Cultivar ~ Malic.acid + Alcohol, data = wine)
##
## Prior probabilities of groups:
##         2         3
## 0.5966387 0.4033613
##
## Group means:
##   Malic.acid  Alcohol
## 2   1.932676 12.27873
## 3   3.333750 13.15375
##
## Coefficients of linear discriminants:
##                  LD1
## Malic.acid 0.5917897
## Alcohol    1.4310158
```

You can also embed plots, for example:

```r
names(wine) # Alcohol (2) + Malic.acid(3) + Ash (4)
```

```
##  [1] "Cultivar"             "Alcohol"              "Malic.acid"
##  [4] "Ash"                  "Alcalinity.ash"       "Magnesium"
##  [7] "Total.phenols"        "Flavanoids"
"Nonflavanoid.phenols"
## [10] "Proanthocyanins"      "Color.intensity"      "Hue"
## [13] "OD280.OD315"          "Proline"
```

```r
lsa.m2.p<-predict(lsa.m2, newdata = wine[,c(2,4)])
lsa.m2.p
```

```
## $class
##   [1] 2 2 2 3 2 2 2 2 2 3 2 2 3 3 3 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2
2 2 2 2
##  [36] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2
2 2 2 2
##  [71] 2 3 3 3 2 2 2 2 2 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 2 3 3 2 3 3
2 3 3 3
## [106] 3 3 3 2 3 3 2 2 3 3 3 3 3 3
## Levels: 2 3
##
## $posterior
##              2           3
## 60  0.989977684 0.010022316
## 61  0.860205334 0.139794666
## 62  0.838593501 0.161406499
## 63  0.209954614 0.790045386
## 64  0.887755020 0.112244980
## 65  0.823294624 0.176705376
## 66  0.691167667 0.308832333
## 67  0.760808366 0.239191634
## 68  0.944036698 0.055963302
## 69  0.159729938 0.840270062
## 70  0.979624822 0.020375178
## 71  0.897120725 0.102879275
## 72  0.013401616 0.986598384
## 73  0.146736039 0.853263961
## 74  0.214129980 0.785870020
## 75  0.949560229 0.050439771
## 76  0.993894082 0.006105918
## 77  0.799148583 0.200851417
## 78  0.971785045 0.028214955
## 79  0.945756380 0.054243620
## 80  0.563880013 0.436119987
## 81  0.977119573 0.022880427
## 82  0.695144470 0.304855530
## 83  0.868676341 0.131323659
## 84  0.352497800 0.647502200
## 85  0.919432248 0.080567752
```

```
## 86  0.702175704 0.297824296
## 87  0.905944583 0.094055417
## 88  0.948585474 0.051414526
## 89  0.969289259 0.030710741
## 90  0.927713154 0.072286846
## 91  0.923364171 0.076635829
## 92  0.918995588 0.081004412
## 93  0.674959331 0.325040669
## 94  0.894171102 0.105828898
## 95  0.983437283 0.016562717
## 96  0.835137841 0.164862159
## 97  0.883445568 0.116554432
## 98  0.947426162 0.052573838
## 99  0.905284898 0.094715102
## 100 0.897120725 0.102879275
## 101 0.988407338 0.011592662
## 102 0.896064451 0.103935549
## 103 0.771530871 0.228469129
## 104 0.991055782 0.008944218
## 105 0.899267283 0.100732717
## 106 0.826525939 0.173474061
## 107 0.929392410 0.070607590
## 108 0.639177508 0.360822492
## 109 0.962358167 0.037641833
## 110 0.942148220 0.057851780
## 111 0.997639559 0.002360441
## 112 0.825998365 0.174001635
## 113 0.834383702 0.165616298
## 114 0.983044410 0.016955590
## 115 0.872235013 0.127764987
## 116 0.998015001 0.001984999
## 117 0.987390259 0.012609741
## 118 0.859807113 0.140192887
## 119 0.795593234 0.204406766
## 120 0.977119573 0.022880427
## 121 0.985003517 0.014996483
## 122 0.781973158 0.218026842
## 123 0.527325472 0.472674528
## 124 0.497893897 0.502106103
## 125 0.949718413 0.050281587
## 126 0.952315853 0.047684147
## 127 0.812472048 0.187527952
## 128 0.876866423 0.123133577
## 129 0.835643218 0.164356782
## 130 0.918886086 0.081113914
## 131 0.499639567 0.500360433
## 132 0.421215952 0.578784048
## 133 0.476445877 0.523554123
## 134 0.594636691 0.405363309
## 135 0.791991332 0.208008668
```

```
## 136 0.769842178 0.230157822
## 137 0.777618886 0.222381114
## 138 0.510560146 0.489439854
## 139 0.167614340 0.832385660
## 140 0.298836253 0.701163747
## 141 0.193996818 0.806003182
## 142 0.155442810 0.844557190
## 143 0.033209945 0.966790055
## 144 0.074284258 0.925715742
## 145 0.910918990 0.089081010
## 146 0.395857398 0.604142602
## 147 0.048610995 0.951389005
## 148 0.368573548 0.631426452
## 149 0.159828615 0.840171385
## 150 0.303622813 0.696377187
## 151 0.047803683 0.952196317
## 152 0.429739155 0.570260845
## 153 0.103701465 0.896298535
## 154 0.257970672 0.742029328
## 155 0.830177214 0.169822786
## 156 0.270671886 0.729328114
## 157 0.034899121 0.965100879
## 158 0.573873462 0.426126538
## 159 0.002661702 0.997338298
## 160 0.047837151 0.952162849
## 161 0.803071776 0.196928224
## 162 0.034032344 0.965967656
## 163 0.312148962 0.687851038
## 164 0.397616578 0.602383422
## 165 0.053367143 0.946632857
## 166 0.069795677 0.930204323
## 167 0.059027745 0.940972255
## 168 0.547207109 0.452792891
## 169 0.030236009 0.969763991
## 170 0.031375696 0.968624304
## 171 0.891467608 0.108532392
## 172 0.601604497 0.398395503
## 173 0.009404846 0.990595154
## 174 0.042062372 0.957937628
## 175 0.097041615 0.902958385
## 176 0.245800940 0.754199060
## 177 0.240668148 0.759331852
## 178 0.004578137 0.995421863
##
## $x
##              LD1
## 60  -2.100088299
## 61  -0.594808825
## 62  -0.503058244
## 63   1.109117982
```

```
## 64  -0.730927716
## 65  -0.443966927
## 66  -0.046347425
## 67  -0.236973923
## 68  -1.141675891
## 69   1.290800935
## 70  -1.709643299
## 71  -0.783865572
## 72   2.721668210
## 73   1.345133833
## 74   1.095566144
## 75  -1.201190374
## 76  -2.370955621
## 77  -0.358369808
## 78  -1.528757513
## 79  -1.159587565
## 80   0.251175826
## 81  -1.645370396
## 82  -0.056486721
## 83  -0.634020133
## 84   0.720242734
## 85  -0.929749759
## 86  -0.074597687
## 87  -0.837799886
## 88  -1.190253911
## 89  -1.481399826
## 90  -0.993424786
## 91  -0.959195771
## 92  -0.926561091
## 93  -0.005741075
## 94  -0.766751065
## 95  -1.824088555
## 96  -0.489331697
## 97  -0.707859039
## 98  -1.177499239
## 99  -0.833614760
## 100 -0.783865572
## 101 -2.020295223
## 102 -0.777687528
## 103 -0.269433895
## 104 -2.162392866
## 105 -0.796595661
## 106 -0.456099141
## 107 -1.007151334
## 108  0.080429337
## 109 -1.367153862
## 110 -1.122593049
## 111 -2.888376675
## 112 -0.454106224
## 113 -0.486366903
```

```
## 114 -1.811159175
## 115 -0.651134640
## 116 -2.982518756
## 117 -1.974133286
## 118 -0.593015199
## 119 -0.346436886
## 120 -1.645370396
## 121 -1.878820037
## 122 -0.302093158
## 123  0.331168194
## 124  0.395067096
## 125 -1.202984000
## 126 -1.233227180
## 127 -0.404556327
## 128 -0.674028608
## 129 -0.491324614
## 130 -0.925763924
## 131  0.391280552
## 132  0.562824209
## 133  0.441627616
## 134  0.182717797
## 135 -0.334503964
## 136 -0.264252310
## 137 -0.288342028
## 138  0.367589418
## 139  1.259561296
## 140  0.852973666
## 141  1.162828422
## 142  1.308314025
## 143  2.218571579
## 144  1.758472800
## 145 -0.870235275
## 146  0.619747899
## 147  2.003257487
## 148  0.682426468
## 149  1.290402351
## 150  0.840642160
## 151  2.012798908
## 152  0.543916076
## 153  1.560049341
## 154  0.963433087
## 155 -0.470024980
## 156  0.928008322
## 157  2.190719900
## 158  0.229079025
## 159  3.604074084
## 160  2.012400325
## 161 -0.371722355
## 162  2.204845031
## 163  0.818943943
```

```
## 164  0.615762065
## 165  1.949921048
## 166  1.794894024
## 167  1.892000899
## 168  0.287796342
## 169  2.271110851
## 170  2.250409093
## 171 -0.751430183
## 172  0.166998332
## 173  2.915906542
## 174  2.085442064
## 175  1.600057816
## 176  0.998459269
## 177  1.013580859
## 178  3.308942334
```

```
lsa.m2.p$class
```

```
##   [1] 2 2 2 3 2 2 2 2 2 3 2 2 3 3 3 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2
2 2 2 2
##  [36] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2
2 2 2 2
##  [71] 2 3 3 3 2 2 2 2 2 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 2 3 3 2 3 3
2 3 3 3
## [106] 3 3 3 2 3 3 2 2 3 3 3 3 3 3
## Levels: 2 3
```

You can also embed plots, for example:

```
lsa.m3.p<-predict(lsa.m3, newdata = wine[,c(2,3)])
lsa.m1.p<-predict(lsa.m1, newdata = wn[,c(2,3)])
```

You can also embed plots, for example:

```
cm.m1<-table(lsa.m1.p$class,wn[,c(1)])
cm.m1
```

```
##
##      1  2  3
##   1 51  5  7
##   2  1 61  9
##   3  7  5 32
```

```
cm.m2<-table(lsa.m2.p$class,wine[,c(1)])
cm.m2
```

```
##
##      2  3
##   2 64 12
##   3  7 36
```

```
cm.m3<-table(lsa.m3.p$class,wine[,c(1)])
cm.m3

##
##      2  3
##   2 63 11
##   3  8 37
```

Resources

- Discriminant Function Analysis
- Computing and visualizing LDA in R
- Computing and visualizing LDA in R