Biplot

Lino Oswaldo Sánchez Juárez

2022-06-01

Introducción

El prefijo "bi" hace referencia ala superposición de individuos y variables en una misma representación. Son útiles paradescribir gráficamente los datos o para mostrar los resultados obtenidos de modelos más formales.

Librerías necesesarias:

Matriz de datos:

```
vinos <- read_excel("C:/Users/Usuario/Documents/Esatadistica multivariada/BIPLOT/vinos.xls")
BD<-as.data.frame(vinos)[,-1]
```

Usaremos uan base de dtaos que contien varaibles basdas en las caracteristicas de dos varidades de vinos **TORO** y **Ribera**, de los años 1986 y 1987.

Exploracion de matriz

```
dim(BD)
```

```
## [1] 45 21
```

Tenemos 45 observaciones y 21 variables.

```
colnames(BD)
```

```
##  [1] "a_o"      "denomina" "grupo"    "grado"    "avol"     "atot"
##  [7] "acfi"     "ph"       "folin"    "somers"   "srv"      "procian"
## [13] "acrg"     "acse"     "achplc"   "ic"       "ic2"      "tono"
## [19] "iim"      "eq1"      "vla"
```

Con esto sabemos los nobres de nuestra 21 variables

```
str(BD)
```

```
## 'data.frame':    45 obs. of  21 variables:
##  $ a_o     : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ denomina: num  1 1 1 1 1 1 1 1 1 1 ...
##  $ grupo   : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ grado   : num  12.8 12.8 12.5 11.9 12.5 12.1 12.2 12.6 13 12.4 ...
##  $ avol    : num  1.2 0.75 1 0.7 0.95 0.5 0.8 0.4 0.4 0.35 ...
##  $ atot    : num  6.7 6.9 7.2 7.7 7.7 5.8 5.9 5.4 4.6 5.5 ...
##  $ acfi    : num  5.2 6 6 6.8 6.3 5.2 4.9 4.9 4.1 5 ...
##  $ ph      : num  3.7 3.5 3.6 3.3 3.6 3.2 3.4 3.3 3.6 3.3 ...
##  $ folin   : num  2827 1818 1459 2054 2930 ...
##  $ somers  : num  50.8 37.8 35.1 32.1 49.6 30.6 35.6 30.6 41.7 30 ...
##  $ srv     : num  811 968 866 978 1128 ...
##  $ procian : num  3794 1736 2306 3420 3158 ...
##  $ acrg    : num  386 144 225 204 214 167 252 315 293 152 ...
##  $ acse    : num  287 141 132 110 148 95 160 124 170 67 ...
##  $ achplc  : num  181 69 78 84 75 74 101 101 137 56 ...
##  $ ic      : num  7.81 4.88 5.52 4.64 6.99 3.98 7.6 6.15 6.6 5.49 ...
##  $ ic2     : num  8.95 5.55 6.35 5.15 7.87 4.36 8.84 7.11 7.85 6.23 ...
##  $ tono    : num  0.72 0.755 0.456 0.675 0.672 0.716 0.716 0.74 0.93 0.75 ...
##  $ iim     : num  18.4 23.6 36.8 36.4 34.2 38.1 28.5 27.7 21.6 30.3 ...
##  $ eq1     : num  0.489 0.48 0.598 0.42 0.45 0.434 0.501 0.566 0.557 0.689 ...
##  $ vla     : num  0.21 0.56 0.38 0.29 0.36 0.3 0.24 0.4 0.28 0.26 ...
```

Es necesario que para nuestros biplot la variable sean numéricas en esta base las tres primeras son reconocidas como numéricas pero son categoricas.

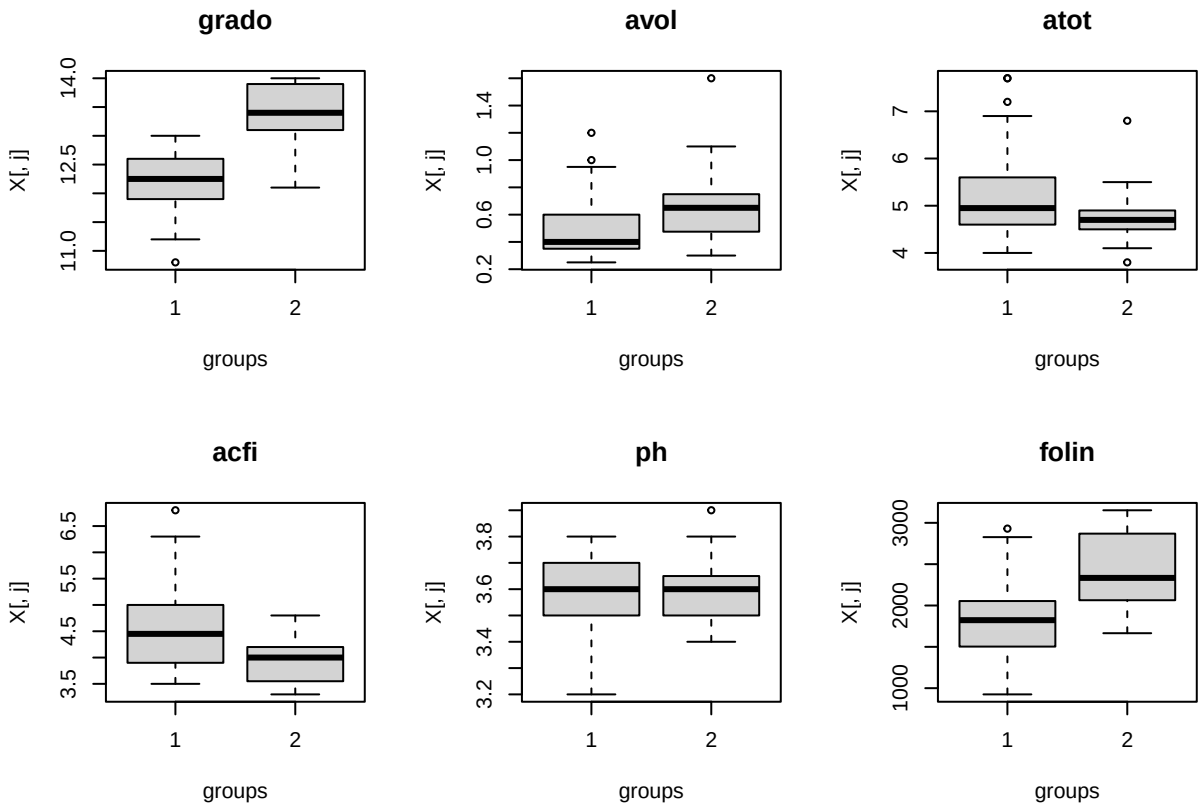
```
BD$denomina = as.factor(BD$denomina)
BD$grupo = as.factor(BD$grupo)
```

Convertimos a factor las varibles de denominación y grupo.

Gráficos de exploración

```
BX1<-BoxPlotPanel(BD[,4:9], nrows=2, groups=BD$denomina)
```

```
## [1] 2
```

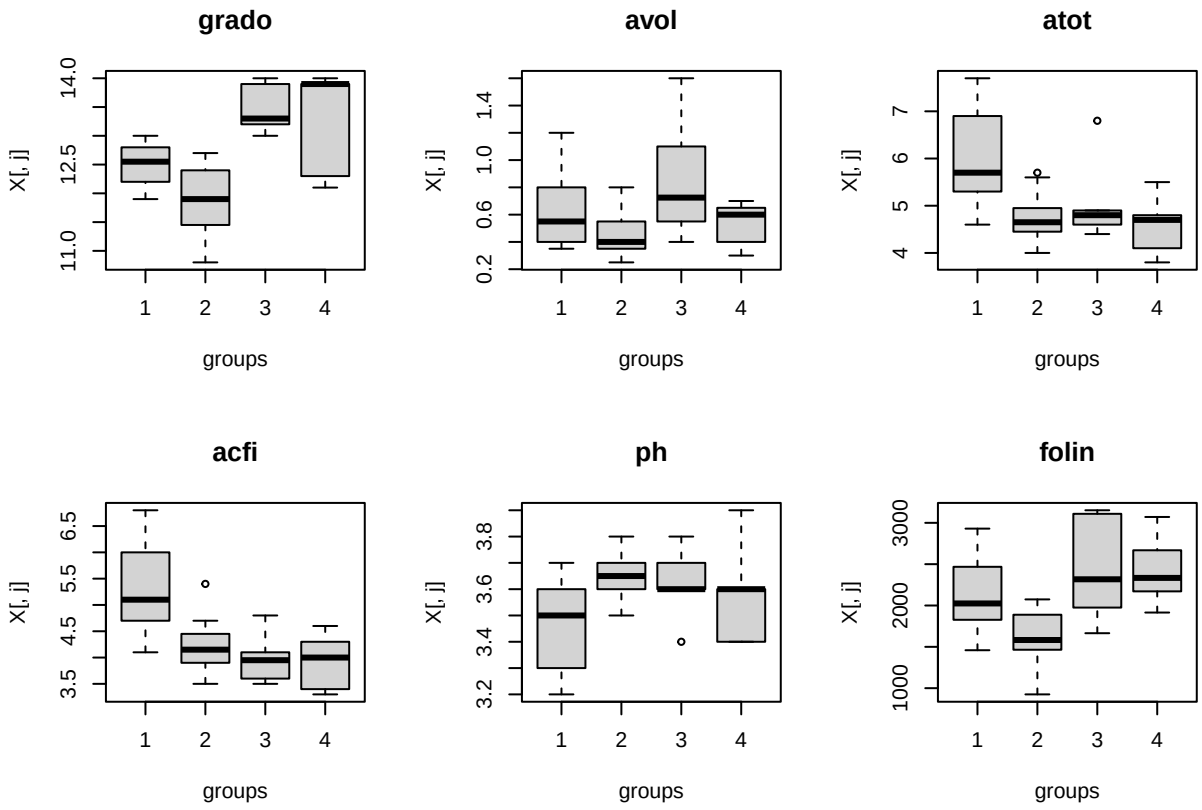



```
BX1
```

```
## $mfrow
## [1] 2 3
```

```
BX2<-BoxPlotPanel(BD[,4:9], nrows=2, groups=BD$grupo)
```

```
## [1] 2
```




```
BX2
```

```
## $mfrow
## [1] 2 3
```

Graficamoslas dos varibles enncontra de las otras variobles y observamos la distribución y la relación que puede haber entre ellas.

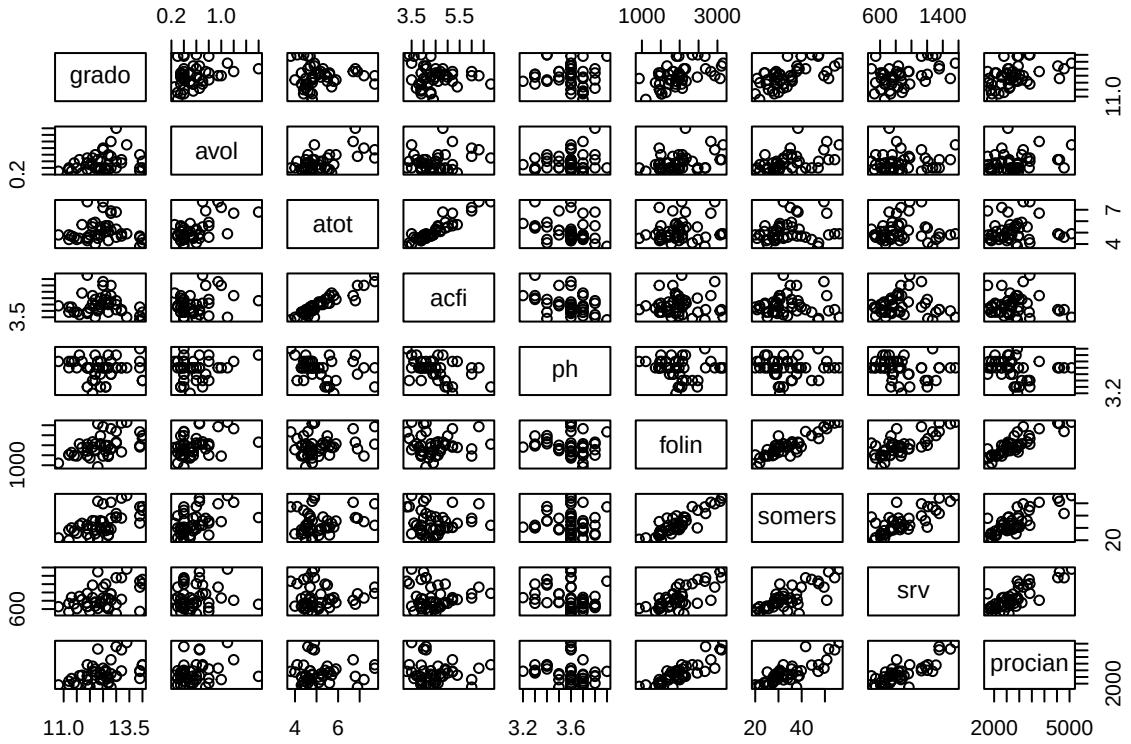
Filtrado de variables

1.- Selección de variables numéricas

```
X<-BD[,4:21]
```

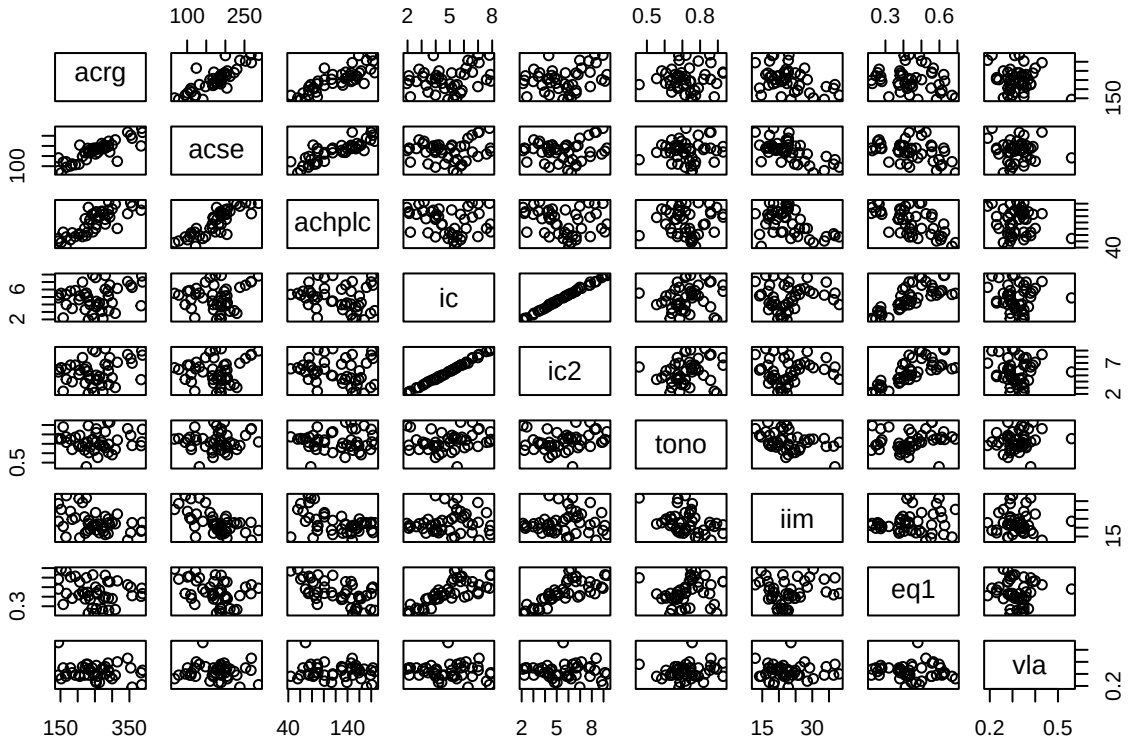
Usamos de la 4 a la 21 ## 2.- Generación del scatter plot

```
PL1<-plot(X[,1:9])
```



Generamos el scatter plot de las variables 1 a la 9; Observamos relación en algunas variables no entodas hay se observan más en la parte inferior derecha.

```
PL2<-plot(X[,10:18])
```



Generamos de la 10 a la 18 y en este gráfico obsevamos correlaciones en la parte superior izquierda, y en las demás no hay correlación aparente.

Reducción de la dimensionalidad

1.- ACP

```
acpvino<-PCA.Analysis(X,Scaling = 5)
summary(acpvino)
```

```
##  ###### Principal Components Analysis #######
##
## Transformation of the raw data:
## [1] "Standardize columns"
##
##  Eigenvalues & Explained Variance (Inertia)
##      Eigenvalue Exp. Var Cummulative
## [1,]  277.12688   34.991      34.991
## [2,]  199.36534   25.172      60.163
## [3,]   85.42317   10.786      70.949
##
##
##  STRUCTURE OF THE PRINCIPAL COMPONENTS
```

```
##          Dim 1  Dim 2  Dim 3
## grado   -0.676 -0.142  0.188
## avol    -0.450  0.204 -0.519
## atot    -0.225  0.738 -0.526
## acfi    -0.063  0.797 -0.397
## ph       0.191 -0.593 -0.193
## folin   -0.910 -0.094 -0.072
## somers  -0.920 -0.154 -0.090
## srv     -0.798 -0.088  0.277
## procian -0.873 -0.102  0.036
## acrg    -0.301 -0.726 -0.441
## acse    -0.213 -0.856 -0.372
## achplc   0.119 -0.830 -0.355
## ic      -0.926  0.117 -0.074
## ic2     -0.932  0.095 -0.048
## tono    -0.351 -0.290  0.612
## iim      0.021  0.810 -0.179
## eq1     -0.688  0.416  0.255
## vla      0.006  0.071  0.368
```

Reducimos la dimencionalidad con un análisis de compoenetes pricipales. y oedimos un resumen.

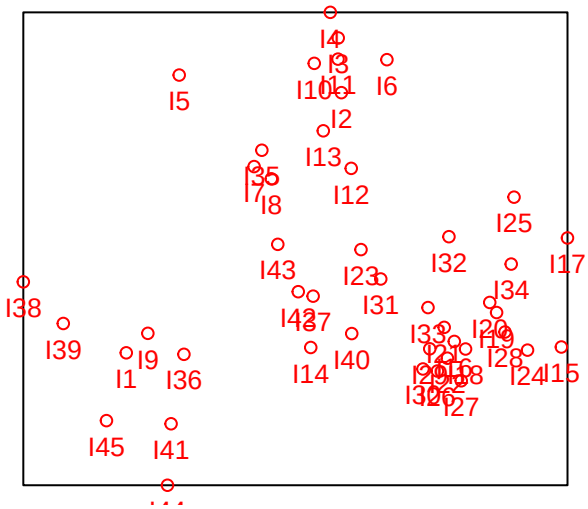
2.- Contenido del objeto acpvino

```
names(acpvino)
```

```
##  [1] "Title"                 "Type"                  "call"
##  [4] "Non_Scaled_Data"       "alpha"                 "Dimension"
##  [7] "Means"                 "Medians"               "Deviations"
## [10] "Minima"                "Maxima"                "P25"
## [13] "P75"                   "GMean"                 "Initial_Transformation"
## [16] "Scaled_Data"           "nrows"                 "ncols"
## [19] "nrowsSup"              "ncolsSup"              "dim"
## [22] "EigenValues"           "Inertia"               "CumInertia"
## [25] "EV"                    "Structure"             "RowCoordinates"
## [28] "ColCoordinates"        "RowContributions"      "ColContributions"
## [31] "Scale_Factor"          "ClusterType"           "Clusters"
## [34] "ClusterColors"         "ClusterNames"
```

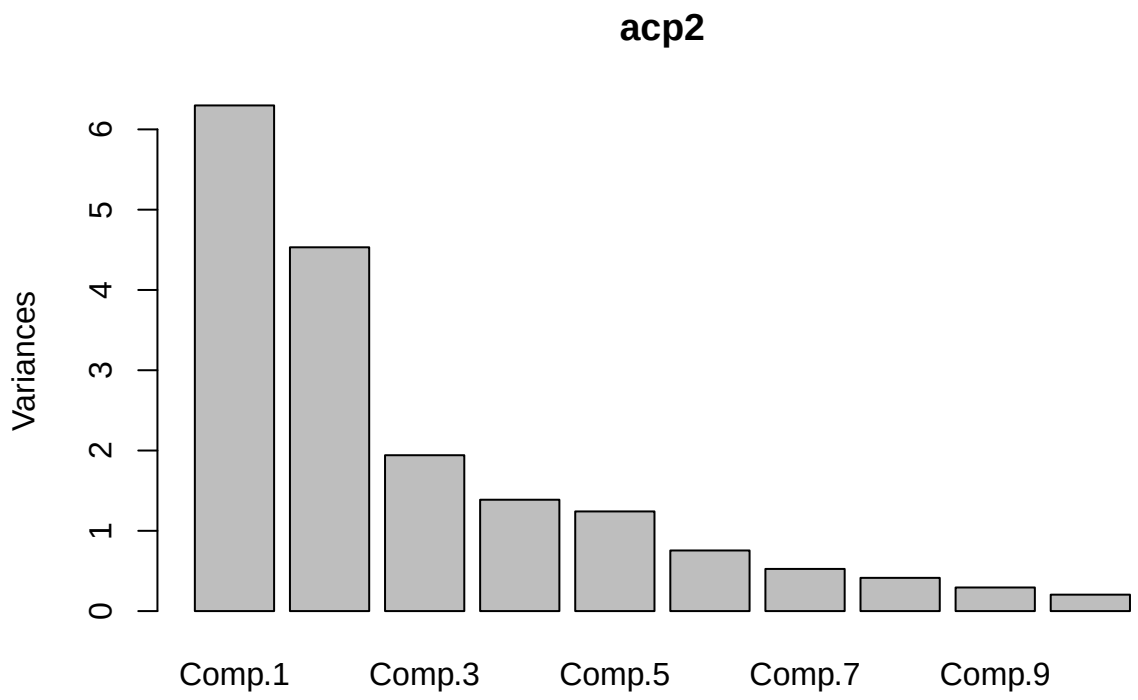
3.- Generación del gráfico sin caja

```
acp1<-plot(acpvino, ShowBox=TRUE)
```



Screeplot con barras

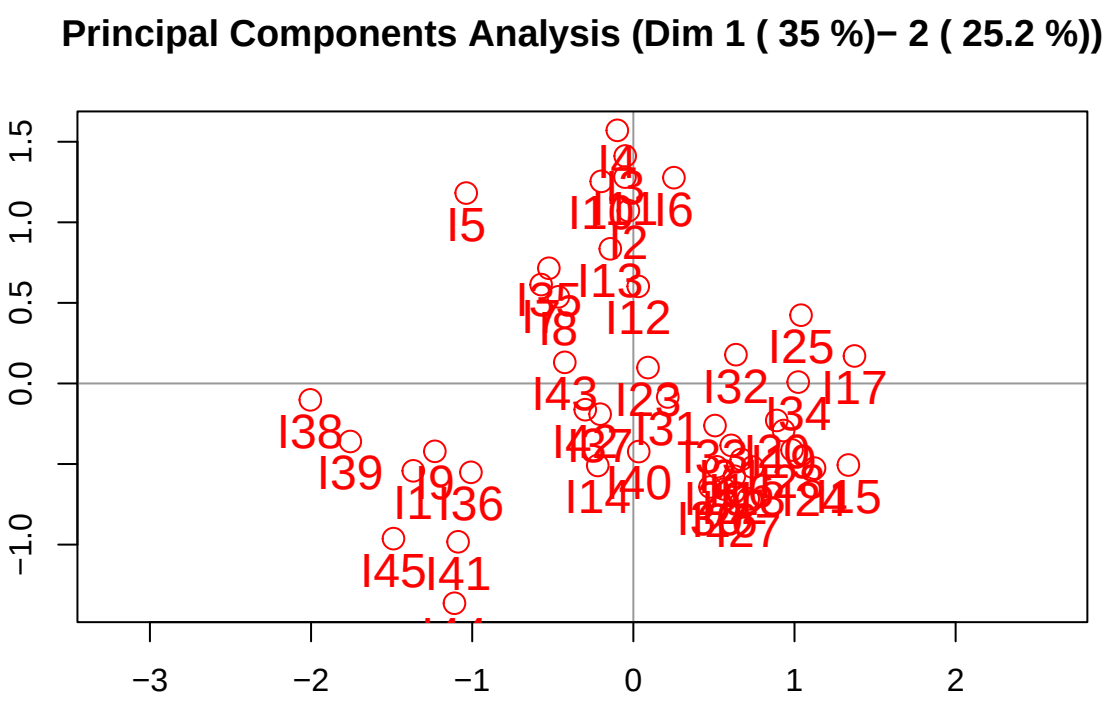
```
acp2<-princomp(X, cor=TRUE, score=TRUE)
plot(acp2)
```

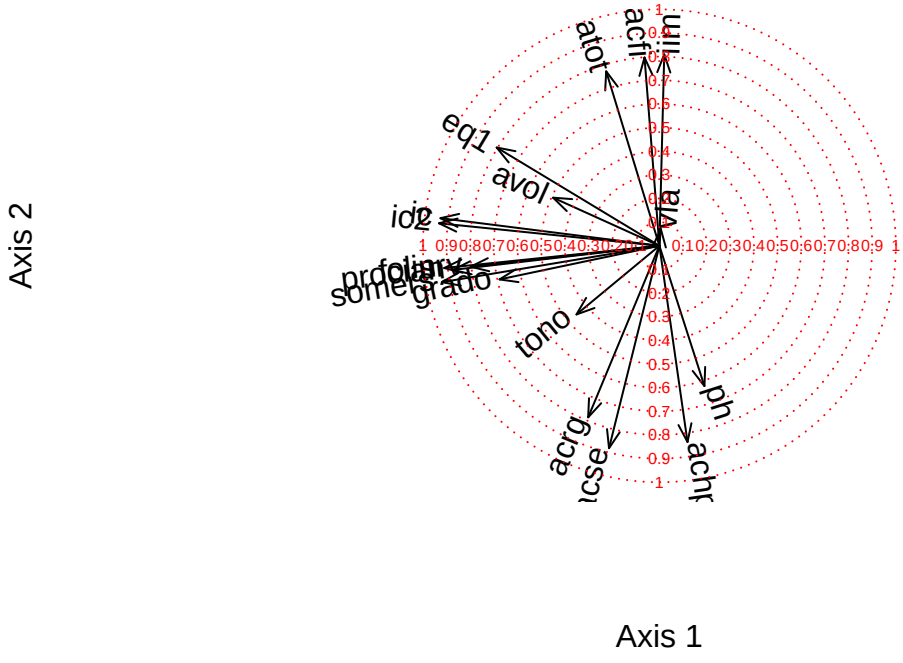


Obserbamos la ditribución de los componentes principales.

Gráfico circular de correlación

```
acp3<-plot(acpvino, CorrelationCircle=TRUE,
           ShowAxis=TRUE,  CexInd=1.5)
```


Principal Components Analysis (Dim 1 (35 %)− 2 (25.2 %))
1.5
1.0
0.5
0.0
−1.0
−3
−2
−1
0
1
2
I5
I6
I2
I13
I12
I25
I32
I17
I34
I43
I23
I31
I38
I39
I19
I36
I40
I14
I45
I41



Esta es una fora de visualizar el biplot por componentes principales.

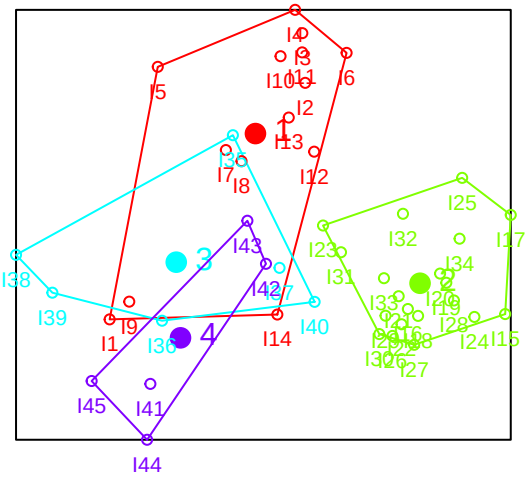
Agregar grupos al biplot definido por usuario

```
acpvino1<-AddCluster2Biplot(acpvino, ClusterType="us",
                            Groups = BD$grupo)
```

Grafico con poligonos

CexInd= tamaño de los argumentos

```
acp4<-plot(acpvino1, PlotClus=TRUE,
           ClustCenters=TRUE, margin=0.05,
           CexInd=0.7, ShowBox=TRUE)
```



Vemos una visulizacion por poligonos

Gráfico con elipses

```
acp5<-plot(acpvino1, PlotClus=TRUE, ClustCenters=TRUE,
     margin=0.05, CexInd=0.7, TypeClus="el",
     ShowBox=F)
```

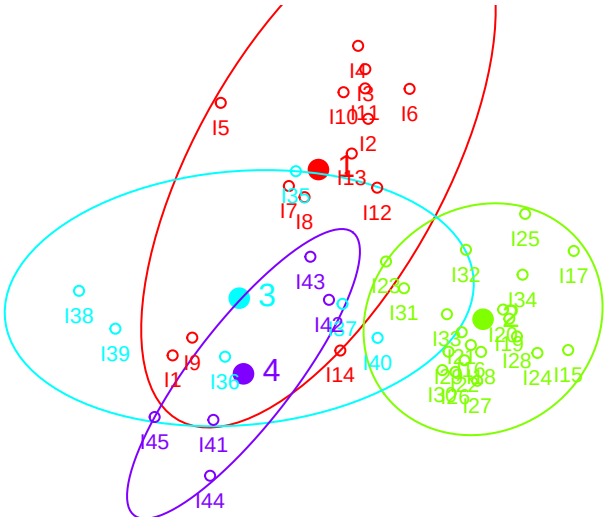
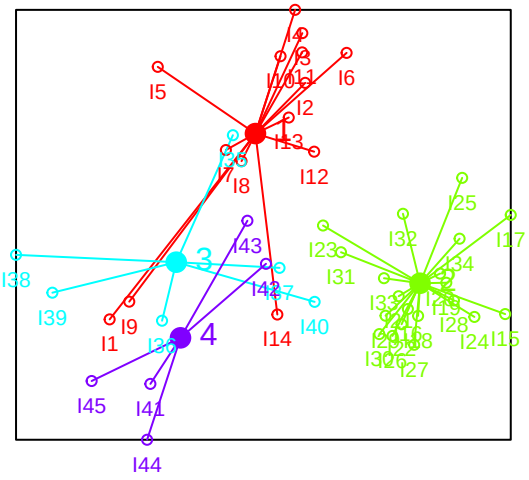


Gráfico con estrellas

```
acp6<-plot(acpvino1, PlotClus=TRUE, ClustCenters=TRUE,
     margin=0.05, CexInd=0.7, TypeClus="st",
     ShowBox=TRUE)
```



Aplicacion del Biplot

Predeterminado JK

```
bipvino<-PCA.Biplot(X, Scaling = 5)
summary(bipvino)
```

```
##  ###### Biplot for Principal Components Analysis #######
##
## Call
## PCA.Biplot(X = X, Scaling = 5)
## Type of coordinates:
## Transformation of the raw data:
## [1] "Standardize columns"
## Type of Biplot
## [1] "PCA"
##
##  Eigenvalues & Explained Variance (Inertia)
##      Eigenvalue Exp. Var Cummulative
## [1,]  277.12688   34.991      34.991
## [2,]  199.36534   25.172      60.163
## [3,]   85.42317   10.786      70.949
##
##
```

```
##  RELATIVE CONTRIBUTIONS OF THE FACTOR TO THE ELEMENT
##
##  Row Contributions
##     Dim 1 Dim 2 Dim 3
## I1  42.04  6.63 44.06
## I2   0.03 32.75  2.51
## I3   0.07 57.27 18.53
## I4   0.28 70.89  8.84
## I5  29.69 38.53 12.87
## I6   2.64 67.95  1.86
## I7  23.63 27.17  9.95
## I8  15.06 20.09  8.55
## I9  60.32  7.05 10.79
## I10  1.55 61.24 11.60
## I11  0.12 74.78 12.48
## I12  0.09 35.48 28.14
## I13  1.68 57.21 29.79
## I14  2.95 15.65 39.81
## I15 64.41  9.27  3.81
## I16 35.26 17.46  4.85
## I17 63.73  0.99 14.28
## I18 51.05 25.08  0.12
## I19 75.88  7.45  9.01
## I20 77.10  5.13  0.77
## I21 24.48  9.79 18.31
## I22 30.80 25.95  0.05
## I23  2.08  2.45  0.20
## I24 71.22 15.56  0.82
## I25 72.83 12.08  2.23
## I26 32.34 43.51  1.52
## I27 35.29 35.58 11.66
## I28 63.06 11.10  3.65
## I29 16.99 16.98 27.32
## I30 17.97 32.85  0.25
## I31  9.13  1.41 35.53
## I32 55.95  4.40  3.60
## I33 28.58  7.60 26.59
## I34 67.06  0.00  4.97
## I35  7.12 13.24  1.64
## I36 41.97 12.56 27.66
## I37  4.81  4.11 20.46
## I38 83.21  0.21  0.95
## I39 88.41  3.71  2.38
## I40  0.08 13.31  0.69
## I41 42.39 34.59  1.36
## I42  9.24  2.75 29.92
## I43 23.86  2.26  7.75
## I44 29.74 44.90  6.03
## I45 56.52 23.65  1.05
##
##  Column Contributions
##         Dim 1 Dim 2 Dim 3
## grado   45.71  2.02  3.54
## avol    20.23  4.14 26.96
```

```
## atot     5.06 54.44 27.69
## acfi     0.40 63.45 15.73
## ph       3.63 35.20  3.72
## folin   82.89  0.89  0.52
## somers  84.58  2.36  0.81
## srv     63.74  0.78  7.65
## procian 76.19  1.04  0.13
## acrg     9.08 52.64 19.41
## acse     4.54 73.25 13.87
## achplc   1.41 68.84 12.61
## ic      85.75  1.37  0.54
## ic2     86.89  0.91  0.23
## tono    12.30  8.43 37.44
## iim      0.04 65.55  3.20
## eq1     47.38 17.28  6.51
## vla      0.00  0.51 13.58
##
##
##
##  Qualities of representation of the rows (Cummulative contributions)
##     Dim 1 Dim 2 Dim 3
## I1  42.04 48.67 92.73
## I2   0.03 32.78 35.29
## I3   0.07 57.34 75.87
## I4   0.28 71.17 80.01
## I5  29.69 68.22 81.09
## I6   2.64 70.59 72.45
## I7  23.63 50.80 60.75
## I8  15.06 35.15 43.70
## I9  60.32 67.37 78.16
## I10  1.55 62.79 74.39
## I11  0.12 74.90 87.38
## I12  0.09 35.57 63.71
## I13  1.68 58.89 88.68
## I14  2.95 18.60 58.41
## I15 64.41 73.68 77.49
## I16 35.26 52.72 57.57
## I17 63.73 64.72 79.00
## I18 51.05 76.13 76.25
## I19 75.88 83.33 92.34
## I20 77.10 82.23 83.00
## I21 24.48 34.27 52.58
## I22 30.80 56.75 56.80
## I23  2.08  4.53  4.73
## I24 71.22 86.78 87.60
## I25 72.83 84.91 87.14
## I26 32.34 75.85 77.37
## I27 35.29 70.87 82.53
## I28 63.06 74.16 77.81
## I29 16.99 33.97 61.29
## I30 17.97 50.82 51.07
## I31  9.13 10.54 46.07
## I32 55.95 60.35 63.95
## I33 28.58 36.18 62.77
```

```
## I34 67.06 67.06 72.03
## I35  7.12 20.36 22.00
## I36 41.97 54.53 82.19
## I37  4.81  8.92 29.38
## I38 83.21 83.42 84.37
## I39 88.41 92.12 94.50
## I40  0.08 13.39 14.08
## I41 42.39 76.98 78.34
## I42  9.24 11.99 41.91
## I43 23.86 26.12 33.87
## I44 29.74 74.64 80.67
## I45 56.52 80.17 81.22
##
##
##
##  Qualities of representation of the columns (Cummulative contributions)
##         Dim 1 Dim 2 Dim 3
## grado   45.71 47.73 51.27
## avol    20.23 24.37 51.33
## atot     5.06 59.50 87.19
## acfi     0.40 63.85 79.58
## ph       3.63 38.83 42.55
## folin   82.89 83.78 84.30
## somers  84.58 86.94 87.75
## srv     63.74 64.52 72.17
## procian 76.19 77.23 77.36
## acrg     9.08 61.72 81.13
## acse     4.54 77.79 91.66
## achplc   1.41 70.25 82.86
## ic      85.75 87.12 87.66
## ic2     86.89 87.80 88.03
## tono    12.30 20.73 58.17
## iim      0.04 65.59 68.79
## eq1     47.38 64.66 71.17
## vla      0.00  0.51 14.09
```

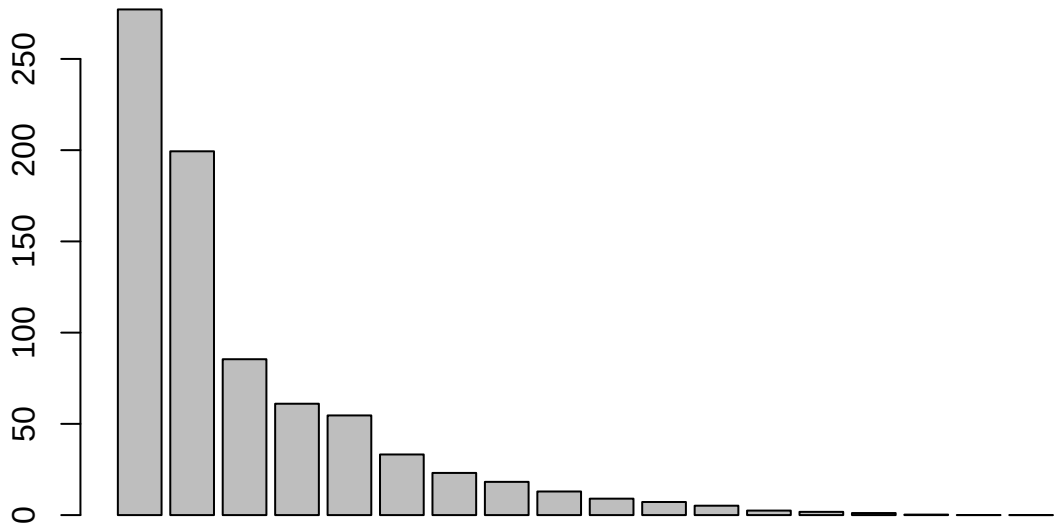
Valores propios

```
bipvino$EigenValues
```

```
##  [1] 277.12687550 199.36534193  85.42316719  61.02361652  54.61472549
##  [6]  33.21950770  23.10087611  18.20271969  12.93567822   8.99721387
## [11]   7.17039349   5.14634483   2.46693118   1.76863760   1.12884586
## [16]   0.26153511   0.02966717   0.01792254
```

Screeplot

```
SC<-barplot(bipvino$EigenValues)
```



Vectores propios

```
bipvino$EV
```

```
##               [,1]        [,2]        [,3]
##  [1,] -0.269400471 -0.06678758  0.13502664
##  [2,] -0.179235894  0.09563188 -0.37266607
##  [3,] -0.089642289  0.34663991 -0.37767939
##  [4,] -0.025075364  0.37420670 -0.28461188
##  [5,]  0.075921760 -0.27872944 -0.13842752
##  [6,] -0.362771201 -0.04421297 -0.05176113
##  [7,] -0.366464498 -0.07220257 -0.06472232
##  [8,] -0.318130606 -0.04157401  0.19854164
##  [9,] -0.347804576 -0.04785685  0.02584725
## [10,] -0.120049408 -0.34086254 -0.31617278
## [11,] -0.084888000 -0.40207820 -0.26728099
## [12,]  0.047378644 -0.38977456 -0.25488092
## [13,] -0.368971746  0.05491570 -0.05287232
## [14,] -0.371435455  0.04476039 -0.03421019
## [15,] -0.139772430 -0.13640832  0.43913353
## [16,]  0.008178563  0.38035721 -0.12838425
## [17,] -0.274261123  0.19527349  0.18313281
## [18,]  0.002361018  0.03345360  0.26444673
```

Tabla de inercias

```
Inercias<-data.frame(paste("Eje",1:length(bipvino$EigenValues)),
                     bipvino$EigenValues, bipvino$Inertia,
                     bipvino$CumInertia)
colnames(Inercias)<-c("Eje", "Valor Propio",
                     "Inercia", "Inercia acumulada")
kable(Inercias)
```

Eje	Valor Propio	Inercia	Inercia acumulada
Eje 1	277.1268755	34.991	34.991
Eje 2	199.3653419	25.172	60.163
Eje 3	85.4231672	10.786	70.949
Eje 4	61.0236165	7.705	78.654
Eje 5	54.6147255	6.896	85.550
Eje 6	33.2195077	4.194	89.744
Eje 7	23.1008761	2.917	92.661
Eje 8	18.2027197	2.298	94.959
Eje 9	12.9356782	1.633	96.592
Eje 10	8.9972139	1.136	97.728
Eje 11	7.1703935	0.905	98.633
Eje 12	5.1463448	0.650	99.283
Eje 13	2.4669312	0.311	99.594
Eje 14	1.7686376	0.223	99.817
Eje 15	1.1288459	0.143	99.960
Eje 16	0.2615351	0.033	99.993
Eje 17	0.0296672	0.004	99.997
Eje 18	0.0179225	0.002	99.999

Tabla contribución de columnas

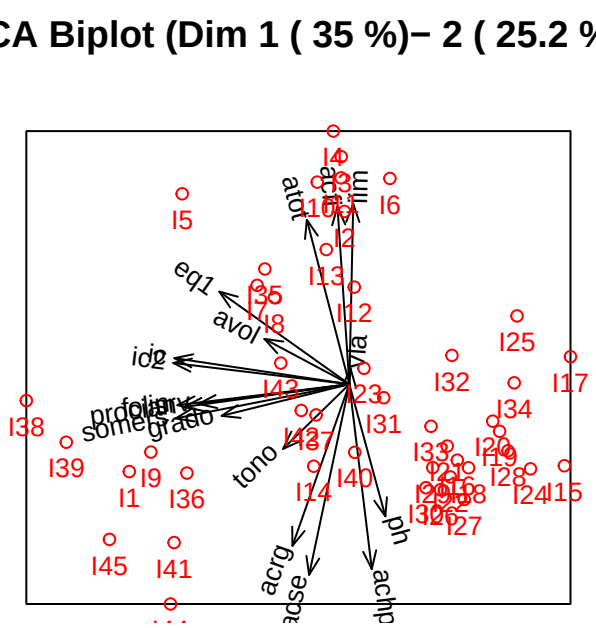
```
kable(bipvino$ColContributions)
```

	Dim 1	Dim 2	Dim 3
grado	45.71	2.02	3.54
avol	20.23	4.14	26.96
atot	5.06	54.44	27.69
acfi	0.40	63.45	15.73
ph	3.63	35.20	3.72
folin	82.89	0.89	0.52
somers	84.58	2.36	0.81
srv	63.74	0.78	7.65
procian	76.19	1.04	0.13
acrg	9.08	52.64	19.41
acse	4.54	73.25	13.87
achplc	1.41	68.84	12.61
ic	85.75	1.37	0.54
ic2	86.89	0.91	0.23

	Dim 1	Dim 2	Dim 3
tono	12.30	8.43	37.44
iim	0.04	65.55	3.20
eq1	47.38	17.28	6.51
vla	0.00	0.51	13.58

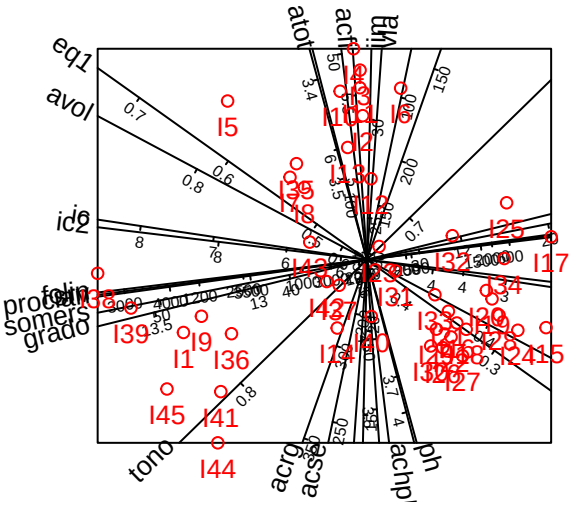
Gráficos Biplot

```
plot(bipvino, ShowBox=TRUE)
```



- Prolongación de vectores linea recta

```
BP1<-plot(bipvino, mode="s",
     margin=0.1, ShowBox=TRUE)
```



- Prolongación de vectores con flechas y linea punteada

```
BP2<-plot(bipvino, mode="ah", margin=0.05,
     ShowBox=TRUE)
```

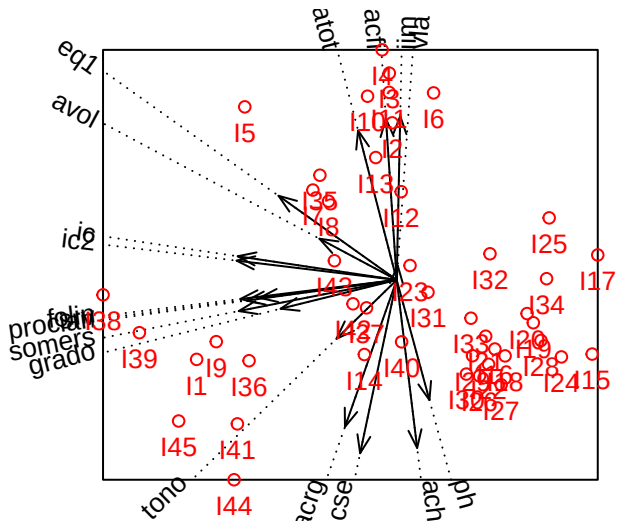


Gráfico circular correlaciones

```
GC<-CorrelationCircle(bipvino)
```

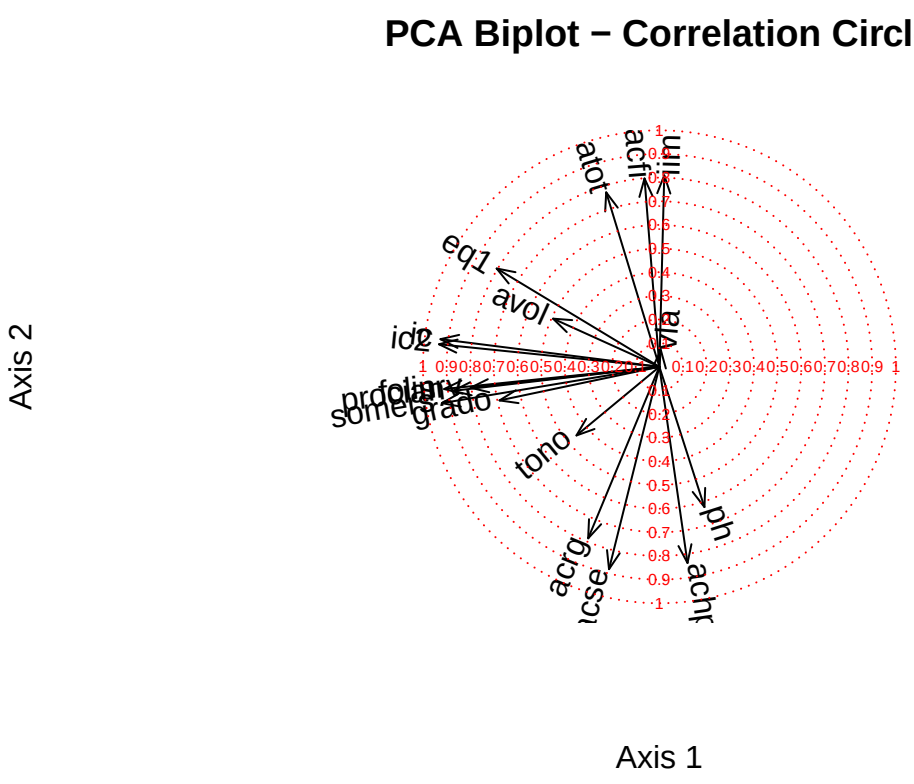
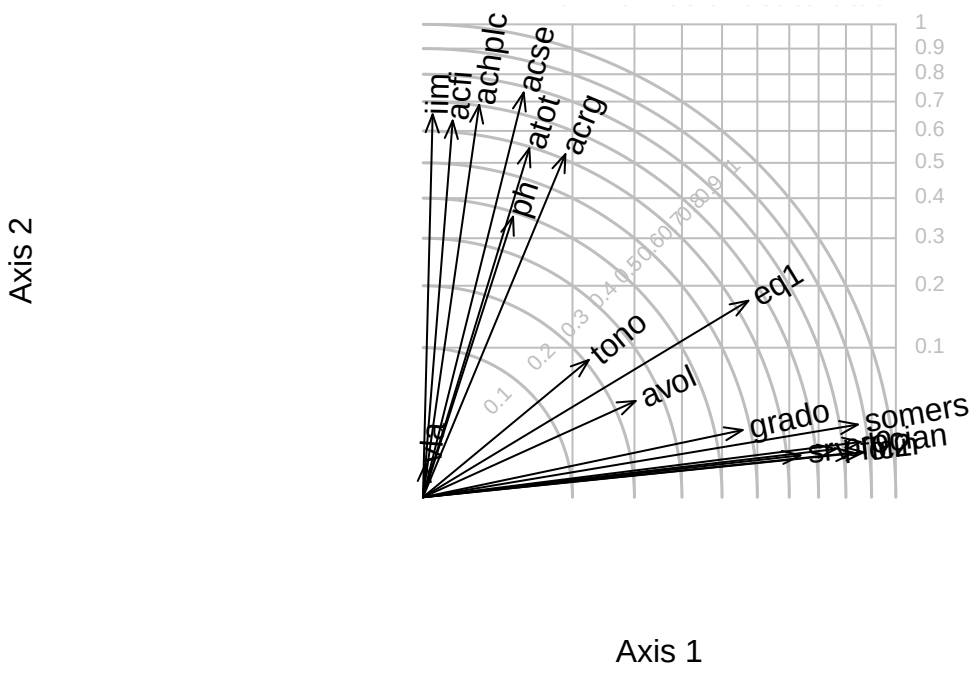


Gráfico contribuciones de los vectores

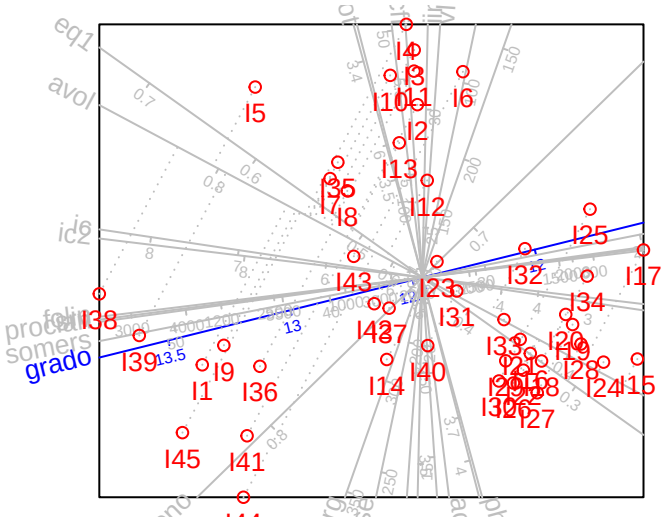
- Calidad de representacion eje 1, 2 y 1+2

```
ColContributionPlot(bipvino, AddSigns2Labs = FALSE)
```



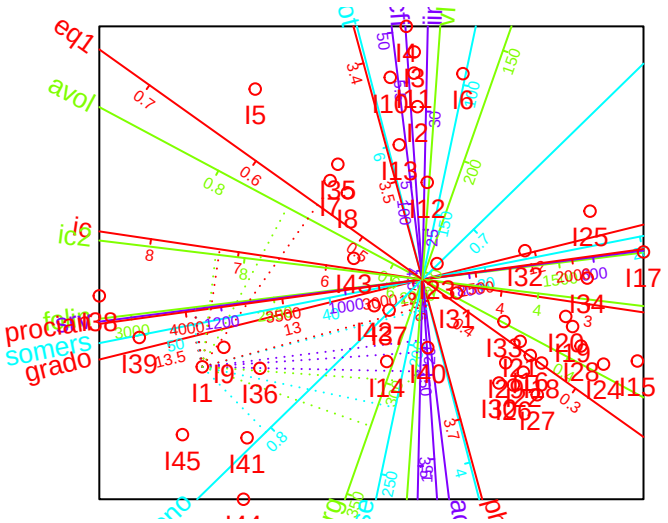
- Proyección individuos sobre una variable donde dp= selecciona la variable

```
BP3<-plot(bipvino, dp=2, mode="s",
     ColorVar=c("blue", rep("grey",17)),
     ShowBox=TRUE)
```



- Proyección de ind sobre todas las variables con *PredPoints= individuo*

```
BP4<-plot(bipvino, PredPoints=1, mode="s",
     ColorVar=1:18, ShowBox=TRUE)
```

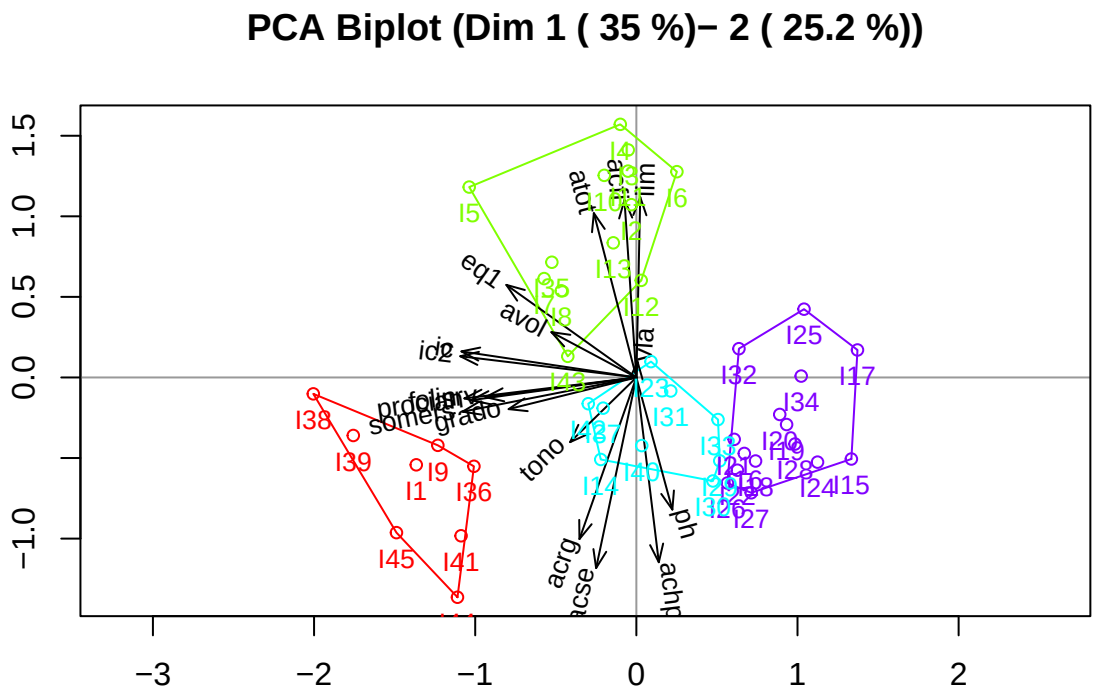


- Cluster Jerárquico con datos originales con el metodo *ward.D*

```
bipvino=AddCluster2Biplot(bipvino, NGroups=4,
                          ClusterType="hi",
                          method="ward.D",
                          Original=TRUE)
```

- Cluster aplicado al biplot

```
clusBP<-plot(bipvino, PlotClus=TRUE,ShowAxis=TRUE)
```


PCA Biplot (Dim 1 (35 %)− 2 (25.2 %))
eq1
avol
tono
acrg
acse
ph
achp
I38
I39
I9
I1
I36
I45
I41
I5
I6
I25
I32
I17
I34
I24
I15
I27
−3
−2
−1
0
1
2
−1.0
0.0
0.5
1.0
1.5

```
clusBP
```

```
## NULL
```