2ª Práctica de Análisis BIPLOT

Consideremos los datos de la matriz X=

$$X=\begin{pmatrix} & Sp1 & Sp2 & Sp3 & Sp4 \\ S1 & 1 & 0 & 0 & 0 \\ S2 & 0 & 0 & 0 & 0 \\ S3 & 0 & 1 & 0 & 0 \\ S4 & 11 & 4 & 0 & 0 \\ S5 & 11 & 5 & 17 & 7 \\ S6 & 9 & 6 & 0 & 0 \\ S7 & 9 & 7 & 13 & 10 \\ S8 & 7 & 8 & 0 & 0 \\ S9 & 7 & 9 & 10 & 13 \\ S10 & 5 & 10 & 0 & 0 \end{pmatrix}$$

recogida en el siguiente artículo, que representa los datos de campo de una investigación medio ambiental, donde las filas representan 10 parcelas y las columnas 4 tipos de especies u organismos, cada elemento de la tabla representa el nº de especies de cada tipo presentes en cada parcela.

Vamos a calcular las coordenadas de las filas y columnas F y G, en cada uno de los 3 tipos de Biplot (SQ, JK, GH), así como la representación gráfica Biplot en R, por último calcularemos la estimación de la matriz X obtenida a través de esta técnica.

También vamos a instalar esta macro en Excel, para realizar el análisis Biplot en la hoja de cálculo, siguiendo las directrices del artículo anterior.

Pasos a realizar en la práctica:

1. Crea el fichero de texto biplot-2.txt con los datos de la matriz X

```
> X
    Sp1 Sp2 Sp3 Sp4
S1    1   0   0   0
S2    0   0   0   0
S3    0   1   0   0
S4   11   4   0   0
S5   11   5  17   7
S6    9   6   0   0
S7    9   7  13  10
S8    7   8   0   0
S9    7   9  10  13
S10   5  10   0   0
```

2. Vamos a comprobar que la matriz X'X es la matriz de covarianzas si se realiza la transformación $x_{ij}=\frac{r_{ij}-\overline{r}_j}{\sqrt{n-1}}$

```
> X_pca
Standard deviations:
[1] 8.577624 4.181997 2.873201 1.567725

Rotation:
          PC1        PC2        PC3        PC4
Sp1 0.3245088  0.6872307  0.5287212 -0.3779708
Sp2 0.2142146  0.6157649 -0.6068400  0.4546327
Sp3 0.7469145 -0.3193073  0.3098449  0.4941232
Sp4 0.5393745 -0.2158481 -0.5061573 -0.6374084

> summary(X_pca)
Importance of components:
                          PC1    PC2     PC3     PC4
Standard deviation     8.5776 4.1820 2.87320 1.56772
Proportion of Variance 0.7229 0.1718 0.08111 0.02415
Cumulative Proportion  0.7229 0.8947 0.97585 1.00000

> cov(X)
          Sp1       Sp2       Sp3       Sp4
Sp1 18.666667  9.444444 14.888889  8.666667
Sp2  9.444444 13.555556  7.333333  8.000000
Sp3 14.888889  7.333333 44.222222 28.777778
Sp4  8.666667  8.000000 28.777778 25.333333

> #Cálculo de la matriz de covarianzas paso a paso

> center=X_pca$center

> center
Sp1 Sp2 Sp3 Sp4
  6   5   4   3

> A1

> A1
      [,1] [,2] [,3] [,4]
 [1,]    1    1    1    1
 [2,]    1    1    1    1
 [3,]    1    1    1    1
 [4,]    1    1    1    1
 [5,]    1    1    1    1
 [6,]    1    1    1    1
 [7,]    1    1    1    1
 [8,]    1    1    1    1
 [9,]    1    1    1    1
[10,]    1    1    1    1

> A2
      [,1] [,2] [,3] [,4]
 [1,]    6    5    4    3
 [2,]    6    5    4    3
 [3,]    6    5    4    3
 [4,]    6    5    4    3
 [5,]    6    5    4    3
 [6,]    6    5    4    3
```

```
 [7,]    6    5    4    3
 [8,]    6    5    4    3
 [9,]    6    5    4    3
[10,]    6    5    4    3
```

```
> A3
    Sp1 Sp2 Sp3 Sp4
S1   -5  -5  -4  -3
S2   -6  -5  -4  -3
S3   -6  -4  -4  -3
S4    5  -1  -4  -3
S5    5   0  13   4
S6    3   1  -4  -3
S7    3   2   9   7
S8    1   3  -4  -3
S9    1   4   6  10
S10  -1   5  -4  -3
```

```
#todo lo anterior se puede hacer con esta orden
```

```
> A3<-sweep(X,2, apply(X,2,mean), FUN="-")#centramos la matriz
> A3
    Sp1 Sp2 Sp3 Sp4
S1   -5  -5  -4  -3
S2   -6  -5  -4  -3
S3   -6  -4  -4  -3
S4    5  -1  -4  -3
S5    5   0  13   4
S6    3   1  -4  -3
S7    3   2   9   7
S8    1   3  -4  -3
S9    1   4   6  10
S10  -1   5  -4  -3
```

```
n=10
```

```
> A4
           Sp1        Sp2       Sp3       Sp4
S1  -1.6666667 -1.6666667 -1.333333 -1.000000
S2  -2.0000000 -1.6666667 -1.333333 -1.000000
S3  -2.0000000 -1.3333333 -1.333333 -1.000000
S4   1.6666667 -0.3333333 -1.333333 -1.000000
S5   1.6666667  0.0000000  4.333333  1.333333
S6   1.0000000  0.3333333 -1.333333 -1.000000
S7   1.0000000  0.6666667  3.000000  2.333333
S8   0.3333333  1.0000000 -1.333333 -1.000000
S9   0.3333333  1.3333333  2.000000  3.333333
S10 -0.3333333  1.6666667 -1.333333 -1.000000
```

```
> t(A4)%*%A4
          Sp1       Sp2       Sp3       Sp4
Sp1 18.666667  9.444444 14.888889  8.666667
Sp2  9.444444 13.555556  7.333333  8.000000
Sp3 14.888889  7.333333 44.222222 28.777778
Sp4  8.666667  8.000000 28.777778 25.333333
```

```
> cov(X)
          Sp1       Sp2       Sp3       Sp4
Sp1 18.666667  9.444444 14.888889  8.666667
Sp2  9.444444 13.555556  7.333333  8.000000
Sp3 14.888889  7.333333 44.222222 28.777778
Sp4  8.666667  8.000000 28.777778 25.333333
```

Efectivamente la matriz obtenida coincide con la matriz de covarianzas

```
> svd(cov(X))
$d
[1] 73.575628 17.489103  8.255287  2.457760

$u
           [,1]       [,2]       [,3]       [,4]
[1,] -0.3245088 -0.6872307  0.5287212  0.3779708
[2,] -0.2142146 -0.6157649 -0.6068400 -0.4546327
[3,] -0.7469145  0.3193073  0.3098449 -0.4941232
[4,] -0.5393745  0.2158481 -0.5061573  0.6374084
```

```
$v
           [,1]       [,2]       [,3]       [,4]
[1,] -0.3245088 -0.6872307  0.5287212  0.3779708
[2,] -0.2142146 -0.6157649 -0.6068400 -0.4546327
[3,] -0.7469145  0.3193073  0.3098449 -0.4941232
[4,] -0.5393745  0.2158481 -0.5061573  0.6374084
```

Comprobamos que el 1º autovalor 73.575628=8.577624^2 es igual al primer valor singular al cuadrado, además los ejes principales de la función prcomp, coinciden con los obtenidos en la DVS de la matriz de Covarianzas

3. Comprobemos que la matriz X'X es la matriz de correlaciones si se realiza la transformación $x_{ij}=\dfrac{r_{ij}-\overline{r}_j}{s_j\sqrt{n-1}}$

```
> #Cálculo de la matriz de correlaciones paso a paso

> X_pca
Standard deviations:
[1] 1.6011643 0.9486476 0.6686156 0.2988208

Rotation:
          PC1        PC2        PC3        PC4
Sp1 0.4807058  0.4464498  0.7040724 -0.2718208
Sp2 0.4343834  0.6260481 -0.5936736  0.2587014
Sp3 0.5386470 -0.4709652  0.1903786  0.6721661
Sp4 0.5386023 -0.4323634 -0.3399847 -0.6382631

> summary(X_pca)
Importance of components:
                          PC1    PC2    PC3     PC4
Standard deviation     1.6012 0.9486 0.6686 0.29882
Proportion of Variance 0.6409 0.2250 0.1118 0.02232
Cumulative Proportion  0.6409 0.8659 0.9777 1.00000

> cor(X)
          Sp1       Sp2       Sp3       Sp4
Sp1 1.0000000 0.5937237 0.5182134 0.3985406
Sp2 0.5937237 1.0000000 0.2995177 0.4317031
Sp3 0.5182134 0.2995177 1.0000000 0.8597869
Sp4 0.3985406 0.4317031 0.8597869 1.0000000

> center
Sp1 Sp2 Sp3 Sp4
  6   5   4   3

> scale
     Sp1      Sp2      Sp3      Sp4
4.320494 3.681787 6.649979 5.033223

> C1=A4[,1]/scale[1]
> C2
> C3
> C4

> A5
             C1          C2         C3         C4
S1  -0.38575837 -0.45267873 -0.2005019 -0.1986799
S2  -0.46291005 -0.45267873 -0.2005019 -0.1986799
S3  -0.46291005 -0.36214298 -0.2005019 -0.1986799
S4   0.38575837 -0.09053575 -0.2005019 -0.1986799
S5   0.38575837  0.00000000  0.6516311  0.2649065
S6   0.23145502  0.09053575 -0.2005019 -0.1986799
S7   0.23145502  0.18107149  0.4511292  0.4635863
S8   0.07715167  0.27160724 -0.2005019 -0.1986799
S9   0.07715167  0.36214298  0.3007528  0.6622662
S10 -0.07715167  0.45267873 -0.2005019 -0.1986799

> t(A5)%*%A5
         C1        C2        C3        C4
C1 1.0000000 0.5937237 0.5182134 0.3985406
C2 0.5937237 1.0000000 0.2995177 0.4317031
C3 0.5182134 0.2995177 1.0000000 0.8597869
C4 0.3985406 0.4317031 0.8597869 1.0000000
```

```
> cor(X)
          Sp1       Sp2       Sp3       Sp4
Sp1 1.0000000 0.5937237 0.5182134 0.3985406
Sp2 0.5937237 1.0000000 0.2995177 0.4317031
Sp3 0.5182134 0.2995177 1.0000000 0.8597869
Sp4 0.3985406 0.4317031 0.8597869 1.0000000
```

Vemos que también X'X es la matriz de correlaciones en el ACPN

4. En este caso también se cumple la relación entre valores singulares y autovalores 1.6011643^2=2.56372704 y vectores propios y ejes factoriales del ACPN

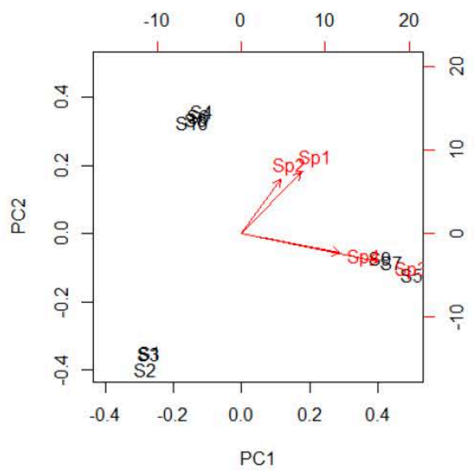
```
> svd(cor(X))
$d
[1] 2.56372704 0.89993231 0.44704679 0.08929386

$u
           [,1]       [,2]       [,3]       [,4]
[1,] -0.4807058  0.4464498 -0.7040724  0.2718208
[2,] -0.4343834  0.6260481  0.5936736 -0.2587014
[3,] -0.5386470 -0.4709652 -0.1903786 -0.6721661
[4,] -0.5386023 -0.4323634  0.3399847  0.6382631

$v
           [,1]       [,2]       [,3]       [,4]
[1,] -0.4807058  0.4464498 -0.7040724  0.2718208
[2,] -0.4343834  0.6260481  0.5936736 -0.2587014
[3,] -0.5386470 -0.4709652 -0.1903786 -0.6721661
[4,] -0.5386023 -0.4323634  0.3399847  0.6382631
```

5. Realiza el **análisis** de Componentes principales sobre X.

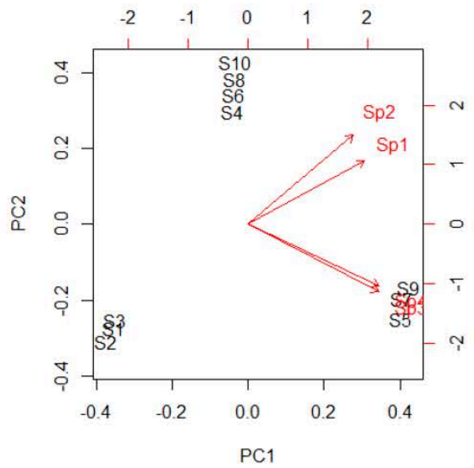
```
> X_pca=prcomp(X, center=FALSE, scale=FALSE)
> X_pca
Standard deviations:b
[1] 12.515847  5.474338  2.873921  1.568157

Rotation:
           PC1        PC2        PC3        PC4
Sp1 -0.5555607 -0.5014692  0.5414099 -0.3830880
Sp2 -0.4409880 -0.4982772 -0.5958536  0.4496753
Sp3 -0.5676054  0.5820419  0.3068285  0.4948814
Sp4 -0.4179786  0.4018403 -0.5076319 -0.6372818
```

6. Realiza la representación gráfica Biplot. Observa la posición de los 4 organismos (columnas), representadas por vectores y de las 10 parcelas (filas) representadas por puntos

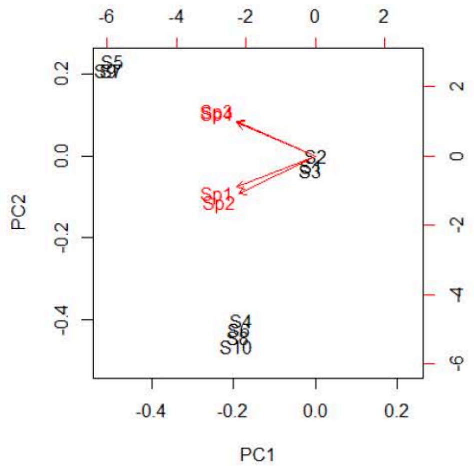
```
> X_pca=prcomp(X, center=TRUE, scale=FALSE)
> biplot(X_pca)
```



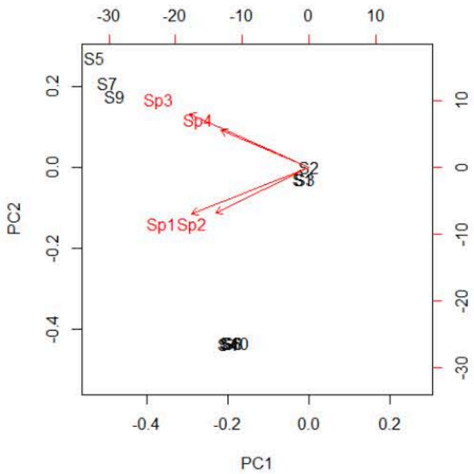

```
> X_pca=prcomp(X, center=TRUE, scale=TRUE)
> biplot(X_pca)
```




```
> X_pca=prcomp(X, center=FALSE, scale=TRUE)
> biplot(X_pca)
```




```
> X_pca=prcomp(X, center=FALSE, scale=FALSE)
> biplot(X_pca)
```



7. Comprobemos si tiene sentido realizar un análisis Biplot sobre la matriz X, en este caso ....

```
> summary(X_pca)
Importance of components:
```

```
                           PC1    PC2     PC3     PC4
Standard deviation     12.5158 5.4743 2.87392 1.56816
Proportion of Variance  0.7938 0.1519 0.04186 0.01246
Cumulative Proportion   0.7938 0.9457 0.98754 1.00000
```

8. Vamos a realizar los cálculos para cada uno de los 3 tipos BIPLOT, comenzamos calculando la (SVD) de la matriz X, (descomposición en valores singulares de X)

```
> svd(X)
$d
[1] 37.547540 16.423015  8.621762  4.704471

$u
               [,1]          [,2]          [,3]          [,4]
 [1,] -1.479619e-02 -3.053454e-02  6.279573e-02 -8.143062e-02
 [2,] -4.175664e-18  8.875375e-18 -3.788861e-17 -9.036090e-17
 [3,] -1.174479e-02 -3.034018e-02 -6.911042e-02  9.558467e-02
 [4,] -2.097373e-01 -4.572406e-01  4.143114e-01 -5.133981e-01
 [5,] -5.563946e-01  2.861867e-01  5.380456e-01  4.222409e-01
 [6,] -2.036345e-01 -4.568519e-01  1.504991e-01 -1.593675e-01
 [7,] -5.232198e-01  2.182173e-01 -4.475132e-02 -5.089302e-02
 [8,] -1.975317e-01 -4.564632e-01 -1.133132e-01  1.946631e-01
 [9,] -5.051620e-01  1.856885e-01 -5.919605e-01 -4.188331e-01
[10,] -1.914289e-01 -4.560745e-01 -3.771255e-01  5.486937e-01

$v
           [,1]       [,2]       [,3]       [,4]
[1,] -0.5555607 -0.5014692  0.5414099 -0.3830880
[2,] -0.4409880 -0.4982772 -0.5958536  0.4496753
[3,] -0.5676054  0.5820419  0.3068285  0.4948814
[4,] -0.4179786  0.4018403 -0.5076319 -0.6372818
```

9. Vamos a comparar estos resultados de la DVS de X con los de la DVS de X'X, (XX1), y XX', (XX2)

```
> XX1
    Sp1 Sp2 Sp3 Sp4
Sp1 528 385 374 258
Sp2 385 372 266 222
Sp3 374 266 558 379
Sp4 258 222 379 318

> XX2
    S1 S2 S3  S4  S5  S6  S7  S8  S9 S10
S1   1  0  0  11  11   9   9   7   7   5
S2   0  0  0   0   0   0   0   0   0   0
S3   0  0  1   4   5   6   7   8   9  10
S4  11  0  4 137 141 123 127 109 113  95
S5  11  0  5 141 484 129 425 117 383 105
S6   9  0  6 123 129 117 123 111 117 105
S7   9  0  7 127 425 123 399 119 386 115
S8   7  0  8 109 117 111 119 113 121 115
S9   7  0  9 113 383 117 386 121 399 125
S10  5  0 10  95 105 105 115 115 125 125

> svd(XX1)
$d
[1] 1409.81776  269.71541   74.33479   22.13205

$u
           [,1]       [,2]       [,3]       [,4]
[1,] -0.5555607 -0.5014692  0.5414099  0.3830880
[2,] -0.4409880 -0.4982772 -0.5958536 -0.4496753
[3,] -0.5676054  0.5820419  0.3068285 -0.4948814
[4,] -0.4179786  0.4018403 -0.5076319  0.6372818

$v
           [,1]       [,2]       [,3]       [,4]
[1,] -0.5555607 -0.5014692  0.5414099  0.3830880
[2,] -0.4409880 -0.4982772 -0.5958536 -0.4496753
[3,] -0.5676054  0.5820419  0.3068285 -0.4948814
[4,] -0.4179786  0.4018403 -0.5076319  0.6372818

> svd(XX2)
$d
```

```
 [1] 1.409818e+03 2.697154e+02 7.433479e+01 2.213205e+01 1.775996e-14
 [6] 1.368788e-14 9.490847e-15 5.201203e-15 3.582003e-15 2.531395e-16

$u
             [,1]        [,2]          [,3]          [,4]         [,5]
 [1,] -0.01479619  0.03053454  6.279573e-02 -8.143062e-02  0.885503667
 [2,]  0.00000000  0.00000000  1.110223e-16  3.330669e-16  0.143578620
 [3,] -0.01174479  0.03034018 -6.911042e-02  9.558467e-02 -0.062926944
 [4,] -0.20973728  0.45724064  4.143114e-01 -5.133981e-01  0.079156569
 [5,] -0.55639461 -0.28618670  5.380456e-01  4.222409e-01 -0.008695919
 [6,] -0.20363448  0.45685193  1.504991e-01 -1.593675e-01 -0.124869497
 [7,] -0.52321983 -0.21821727 -4.475132e-02 -5.089302e-02  0.019030200
 [8,] -0.19753167  0.45646321 -1.133132e-01  1.946631e-01 -0.295339858
 [9,] -0.50516203 -0.18568848 -5.919605e-01 -4.188331e-01 -0.009956197
[10,] -0.19142886  0.45607450 -3.771255e-01  5.486937e-01  0.285811049
            [,6]         [,7]        [,8]       [,9]       [,10]
 [1,] -0.2807218  0.192650300  0.01376733  0.2961327  0.01883516
 [2,]  0.6149250  0.434333150 -0.56347743 -0.1208159  0.28373205
 [3,]  0.3789240 -0.167436140 -0.14196695  0.7321304 -0.50327056
 [4,]  0.3221626  0.001821785  0.27809006 -0.2142584 -0.28001997
 [5,]  0.1301137 -0.010039577  0.22253907  0.1341877  0.23686676
 [6,] -0.2590610 -0.400658833 -0.46114695  0.2496292  0.42844787
 [7,] -0.2847415  0.021970669 -0.48700580 -0.2936572 -0.51836060
 [8,] -0.2624094  0.701474255  0.10976410  0.1969713 -0.05378345
 [9,]  0.1489707 -0.011494588  0.25479111  0.1536352  0.27119528
[10,]  0.1987953 -0.304783754  0.09216008 -0.2946698 -0.05136363

$v
             [,1]          [,2]          [,3]          [,4]         [,5]
 [1,] -0.01479619  3.053454e-02  6.279573e-02 -8.143062e-02 -0.306314174
 [2,]  0.00000000  1.249001e-16  7.441964e-16  1.020763e-15  0.077140946
 [3,] -0.01174479  3.034018e-02 -6.911042e-02  9.558467e-02 -0.026503958
 [4,] -0.20973728  4.572406e-01  4.143114e-01 -5.133981e-01 -0.002799065
 [5,] -0.55639461 -2.861867e-01  5.380456e-01  4.222409e-01  0.352863922
 [6,] -0.20363448  4.568519e-01  1.504991e-01 -1.593675e-01  0.058296097
 [7,] -0.52321983 -2.182173e-01 -4.475132e-02 -5.089302e-02 -0.772209452
 [8,] -0.19753167  4.564632e-01 -1.133132e-01  1.946631e-01  0.068758932
 [9,] -0.50516203 -1.856885e-01 -5.919605e-01 -4.188331e-01  0.404003621
[10,] -0.19142886  4.560745e-01 -3.771255e-01  5.486937e-01 -0.085703385
             [,6]        [,7]        [,8]        [,9]       [,10]
 [1,]  0.14985361 -0.09327670  0.01782062 -0.68224191 -0.63049016
 [2,] -0.81187493  0.35155988  0.02736274 -0.42352602  0.17660978
 [3,] -0.11340338  0.53681602  0.39131381  0.44222975 -0.57844704
 [4,] -0.29730595 -0.36966019  0.15156550  0.22339030 -0.11741861
 [5,]  0.01924221 -0.04269023 -0.01093514 -0.05275758 -0.10551404
 [6,]  0.27871286  0.57604928 -0.53130422 -0.05495498  0.05962428
 [7,] -0.04210976  0.09342355  0.02393052  0.11545500  0.23090754
 [8,]  0.26996289  0.06785373  0.66473428 -0.27099337  0.32306040
 [9,]  0.02203093 -0.04887722 -0.01251994 -0.06040361 -0.12080593
[10,] -0.25290742 -0.30579195 -0.31277831  0.11611212 -0.18956366
```

Se observa la relación existente entre las DVS de X'X y XX', (que coinciden con las autodescomposiciones de X'X y XX') y la DVS de X, en el 1º caso los vectores singulares a izquierda y derecha son iguales, vectores U, para X'X y vectores V para XX', (aquí los nombres u y v están intercambiados en relación a la teoría), y también se observa que los valores singulares son las raíces cuadradas de los autovalores, ya que en la DVS, se obtienen valores singulares y en las autodescomposiciones, autovalores.
También se puede comprobar que las componentes principales del análisis realizado en el apartado 2, coinciden con los vectores U de X'X.

Al no haber realizado el ACP con el centrado de la tabla, las desviaciones típicas del ACP, no van a coincidir con las raíces cuadradas de los autovalores de la DVS de la matriz de covarianzas, y al no haber utilizado la escala, tampoco coincidirán con las raíces cuadradas de los autovalores de la DVS de la matriz de correlaciones.
Solo en el caso en que se centren los datos (ACP) se dará la relación entre las desviaciones típicas de la función prcom, y los autovalores de la DVS de la matriz de covarianzas.
Solo en el caso en que se centren los datos y se normalicen (ACPN) se dará la relación entre las desviaciones típicas de la función prcom, y los autovalores de la DVS de la matriz de correlaciones.
En cualquier otro caso, no se cumple, por ejemplo datos no centrados y tipificados.

10. Vamos a ver si se cumplen estas coincidencias en el caso de centrar y tipificar los datos:
    a. Calculamos el ACP con los datos centrados sin tipificar

```
> X_pca
Standard deviations:
[1] 8.577624 4.181997 2.873201 1.567725

Rotation:
          PC1        PC2        PC3        PC4
Sp1 0.3245088  0.6872307  0.5287212 -0.3779708
Sp2 0.2142146  0.6157649 -0.6068400  0.4546327
Sp3 0.7469145 -0.3193073  0.3098449  0.4941232
Sp4 0.5393745 -0.2158481 -0.5061573 -0.6374084
> summary(X_pca)
Importance of components:
                          PC1    PC2     PC3     PC4
Standard deviation     8.5776 4.1820 2.87320 1.56772
Proportion of Variance 0.7229 0.1718 0.08111 0.02415
Cumulative Proportion  0.7229 0.8947 0.97585 1.00000

> center
Sp1 Sp2 Sp3 Sp4
  6   5   4   3

> scale
[1] FALSE

> svd(cov(X))
$d
[1] 73.575628 17.489103  8.255287  2.457760

$u
           [,1]       [,2]       [,3]       [,4]
[1,] -0.3245088 -0.6872307  0.5287212  0.3779708
[2,] -0.2142146 -0.6157649 -0.6068400 -0.4546327
[3,] -0.7469145  0.3193073  0.3098449 -0.4941232
[4,] -0.5393745  0.2158481 -0.5061573  0.6374084

$v
           [,1]       [,2]       [,3]       [,4]
[1,] -0.3245088 -0.6872307  0.5287212  0.3779708
[2,] -0.2142146 -0.6157649 -0.6068400 -0.4546327
[3,] -0.7469145  0.3193073  0.3098449 -0.4941232
[4,] -0.5393745  0.2158481 -0.5061573  0.6374084
```

Efectivamente se cumple la relación $8.577624^2 = 73.575628$

    b. Calculamos ahora el ACPN con los datos centrados y tipificados

```
Standard deviations:
[1] 1.6011643 0.9486476 0.6686156 0.2988208

Rotation:
          PC1        PC2        PC3        PC4
Sp1 0.4807058  0.4464498  0.7040724 -0.2718208
Sp2 0.4343834  0.6260481 -0.5936736  0.2587014
Sp3 0.5386470 -0.4709652  0.1903786  0.6721661
Sp4 0.5386023 -0.4323634 -0.3399847 -0.6382631

> summary(X_pca)
Importance of components:
                          PC1    PC2    PC3     PC4
Standard deviation     1.6012 0.9486 0.6686 0.29882
Proportion of Variance 0.6409 0.2250 0.1118 0.02232
Cumulative Proportion  0.6409 0.8659 0.9777 1.00000

> center
Sp1 Sp2 Sp3 Sp4
  6   5   4   3

> scale
     Sp1      Sp2      Sp3      Sp4
4.320494 3.681787 6.649979 5.033223
```

```
> svd(cor(X))
$d
[1] 2.56372704 0.89993231 0.44704679 0.08929386

$u
           [,1]       [,2]       [,3]       [,4]
[1,] -0.4807058  0.4464498 -0.7040724  0.2718208
[2,] -0.4343834  0.6260481  0.5936736 -0.2587014
[3,] -0.5386470 -0.4709652 -0.1903786 -0.6721661
[4,] -0.5386023 -0.4323634  0.3399847  0.6382631

$v
           [,1]       [,2]       [,3]       [,4]
[1,] -0.4807058  0.4464498 -0.7040724  0.2718208
[2,] -0.4343834  0.6260481  0.5936736 -0.2587014
[3,] -0.5386470 -0.4709652 -0.1903786 -0.6721661
[4,] -0.5386023 -0.4323634  0.3399847  0.6382631
```

Podemos comprobar que también se cumple la relación $1.6011643^2=2.56372704$

11. Vamos a calcular las coordenadas de las filas y columnas F y G de los 3 tipos de Biplot

    a. SQ Biplot (c=0.5) $F=VD_\lambda^{1/4}$ $G= UD_\lambda^{1/4}$

```
> F1
[1] -9.066522e-02 -2.558682e-17 -7.196743e-02 -1.285187e+00 -3.409366e+00
[6] -1.247792e+00 -3.206084e+00 -1.210396e+00 -3.095433e+00 -1.173000e+00

> F2
[1] -1.237422e-01  3.596774e-17 -1.229545e-01 -1.852982e+00  1.159781e+00
[6] -1.851407e+00  8.843325e-01 -1.849832e+00  7.525085e-01 -1.848256e+00

> F
                F1            F2
 [1,] -9.066522e-02 -1.237422e-01
 [2,] -2.558682e-17  3.596774e-17
 [3,] -7.196743e-02 -1.229545e-01
 [4,] -1.285187e+00 -1.852982e+00
 [5,] -3.409366e+00  1.159781e+00
 [6,] -1.247792e+00 -1.851407e+00
 [7,] -3.206084e+00  8.843325e-01
 [8,] -1.210396e+00 -1.849832e+00
 [9,] -3.095433e+00  7.525085e-01
[10,] -1.173000e+00 -1.848256e+00

> G1
[1] -3.404256 -2.702200 -3.478062 -2.561207

> G2
[1] -2.032220 -2.019284  2.358743  1.628471

> G
            G1        G2
[1,] -3.404256 -2.032220
[2,] -2.702200 -2.019284
[3,] -3.478062  2.358743
[4,] -2.561207  1.628471

F%*%t(G)
              [,1]          [,2]          [,3]          [,4]
 [1,] 5.601190e-01  4.948662e-01  2.346315e-02  3.070191e-02
 [2,] 1.400974e-17 -3.488392e-18  1.738312e-16  1.241056e-16
 [3,] 4.948662e-01  4.427506e-01 -3.971110e-02 -1.590438e-02
 [4,] 8.140774e+00  7.214531e+00  9.925023e-02  2.741035e-01
 [5,] 9.249427e+00  6.870862e+00  1.459361e+01  1.062076e+01
 [6,] 8.010268e+00  7.110300e+00 -2.709826e-02  1.808909e-01
 [7,] 9.117174e+00  6.877762e+00  1.323687e+01  9.651557e+00
 [8,] 7.879763e+00  7.006068e+00 -1.534468e-01  8.767832e-02
 [9,] 9.008385e+00  6.844951e+00  1.254108e+01  9.153485e+00
[10,] 7.749257e+00  6.901837e+00 -2.797952e-01 -5.534254e-03

> (round(F%*%t(G),2))
      [,1] [,2]  [,3]  [,4]
 [1,] 0.56 0.49  0.02  0.03
 [2,] 0.00 0.00  0.00  0.00
 [3,] 0.49 0.44 -0.04 -0.02
```

```
 [4,] 8.14 7.21  0.10  0.27
 [5,] 9.25 6.87 14.59 10.62
 [6,] 8.01 7.11 -0.03  0.18
 [7,] 9.12 6.88 13.24  9.65
 [8,] 7.88 7.01 -0.15  0.09
 [9,] 9.01 6.84 12.54  9.15
[10,] 7.75 6.90 -0.28 -0.01
```

b. JK Biplot (c=1) $F=VD_\lambda^{1/2}$ G= U

```
> F1
[1] -5.555607e-01 -1.567859e-16 -4.409880e-01 -7.875119e+00 -2.089125e+01
[6] -7.645974e+00 -1.964562e+01 -7.416828e+00 -1.896759e+01 -7.187683e+00

> F2
[1] -5.014692e-01  1.457604e-16 -4.982772e-01 -7.509270e+00  4.700048e+00
[6] -7.502886e+00  3.583785e+00 -7.496502e+00  3.049565e+00 -7.490118e+00

> F
                 F1            F2
 [1,] -5.555607e-01 -5.014692e-01
 [2,] -1.567859e-16  1.457604e-16
 [3,] -4.409880e-01 -4.982772e-01
 [4,] -7.875119e+00 -7.509270e+00
 [5,] -2.089125e+01  4.700048e+00
 [6,] -7.645974e+00 -7.502886e+00
 [7,] -1.964562e+01  3.583785e+00
 [8,] -7.416828e+00 -7.496502e+00
 [9,] -1.896759e+01  3.049565e+00
[10,] -7.187683e+00 -7.490118e+00

> G1
[1] -0.5555607 -0.4409880 -0.5676054 -0.4179786

> G2
[1] -0.5014692 -0.4982772  0.5820419  0.4018403

> G
             G1         G2
[1,] -0.5555607 -0.5014692
[2,] -0.4409880 -0.4982772
[3,] -0.5676054  0.5820419
[4,] -0.4179786  0.4018403

round(F%*%t(G),2)
      [,1] [,2]  [,3]  [,4]
 [1,] 0.56 0.49  0.02  0.03
 [2,] 0.00 0.00  0.00  0.00
 [3,] 0.49 0.44 -0.04 -0.02
 [4,] 8.14 7.21  0.10  0.27
 [5,] 9.25 6.87 14.59 10.62
 [6,] 8.01 7.11 -0.03  0.18
 [7,] 9.12 6.88 13.24  9.65
 [8,] 7.88 7.01 -0.15  0.09
 [9,] 9.01 6.84 12.54  9.15
[10,] 7.75 6.90 -0.28 -0.01
```

c. GH Biplot (c=0) F=V G= $UD_\lambda^{1/2}$

```
> F1
[1] -1.479619e-02 -4.175664e-18 -1.174479e-02 -2.097373e-01 -5.563946e-01
[6] -2.036345e-01 -5.232198e-01 -1.975317e-01 -5.051620e-01 -1.914289e-01

> F2
[1] -3.053454e-02  8.875375e-18 -3.034018e-02 -4.572406e-01  2.861867e-01
[6] -4.568519e-01  2.182173e-01 -4.564632e-01  1.856885e-01 -4.560745e-01

> F
                 F1            F2
 [1,] -1.479619e-02 -3.053454e-02
 [2,] -4.175664e-18  8.875375e-18
 [3,] -1.174479e-02 -3.034018e-02
 [4,] -2.097373e-01 -4.572406e-01
```

```
 [5,] -5.563946e-01  2.861867e-01
 [6,] -2.036345e-01 -4.568519e-01
 [7,] -5.232198e-01  2.182173e-01
 [8,] -1.975317e-01 -4.564632e-01
 [9,] -5.051620e-01  1.856885e-01
[10,] -1.914289e-01 -4.560745e-01
```

```
> G1
[1] -20.85994 -16.55801 -21.31219 -15.69407
```

```
> G2
[1] -8.235635 -8.183214  9.558883  6.599430
```

```
> G
            G1        G2
[1,] -20.85994 -8.235635
[2,] -16.55801 -8.183214
[3,] -21.31219  9.558883
[4,] -15.69407  6.599430
```

```
> round(F%*%t(G),2)
      [,1] [,2]  [,3]  [,4]
 [1,] 0.56 0.49  0.02  0.03
 [2,] 0.00 0.00  0.00  0.00
 [3,] 0.49 0.44 -0.04 -0.02
 [4,] 8.14 7.21  0.10  0.27
 [5,] 9.25 6.87 14.59 10.62
 [6,] 8.01 7.11 -0.03  0.18
 [7,] 9.12 6.88 13.24  9.65
 [8,] 7.88 7.01 -0.15  0.09
 [9,] 9.01 6.84 12.54  9.15
[10,] 7.75 6.90 -0.28 -0.01
```

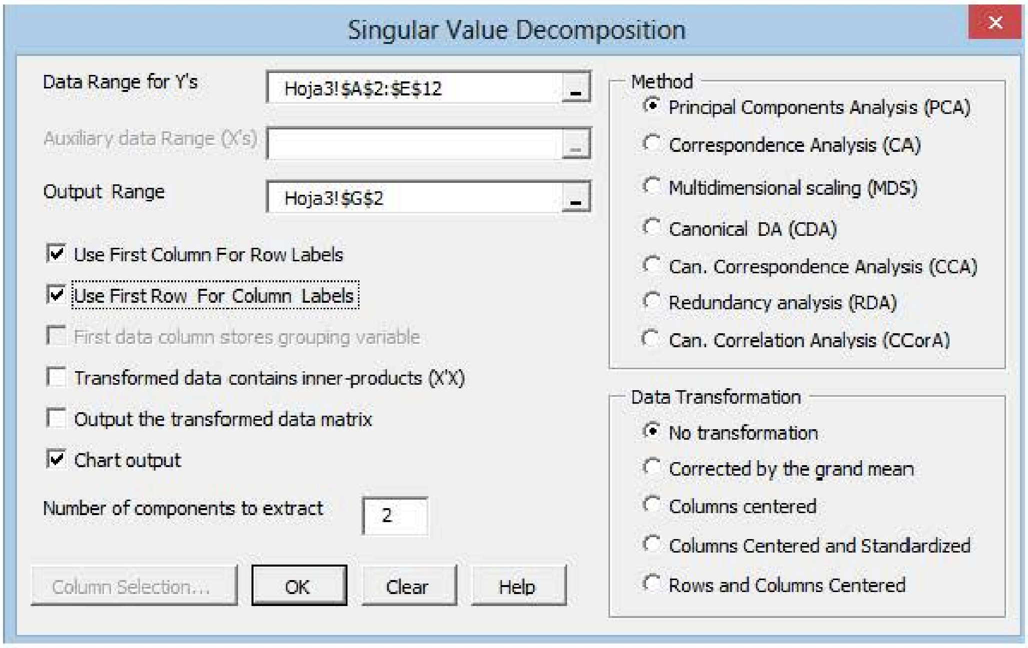
12. Vamos a obtener la matriz estimada de X por el Análisis Biplot, no por el producto de FxG, sino como la reconstrucción de 2º grado de la matriz X

```
>round(((svd(X)$d[1]))*(svd(X)$u[,1])%*%(t(svd(X)$v[,1]))+((svd(X)$d[2]))*(svd(X)$u
[,2])%*%(t(svd(X)$v[,2])),2)
      [,1] [,2]  [,3]  [,4]
 [1,] 0.56 0.49  0.02  0.03
 [2,] 0.00 0.00  0.00  0.00
 [3,] 0.49 0.44 -0.04 -0.02
 [4,] 8.14 7.21  0.10  0.27
 [5,] 9.25 6.87 14.59 10.62
 [6,] 8.01 7.11 -0.03  0.18
 [7,] 9.12 6.88 13.24  9.65
 [8,] 7.88 7.01 -0.15  0.09
 [9,] 9.01 6.84 12.54  9.15
[10,] 7.75 6.90 -0.28 -0.01
```

Se puede comprobar que los 3 tipos de Biplot dan la misma matriz estimada, que coincide con la reconstrucción de la matriz de 2º nivel.

Vamos ahora a utilizar la macro para EXCEL, que permite realizar el análisis Biplot

1. En 'Datos', elegimos la opción 'desde texto', se elige el archivo de texto con los datos de la tabla, pulsamos en 'Importar', marcamos la opción 'de ancho fijo', pulsamos en 'siguiente' y en 'finalizar' y ya tenemos los datos incorporados en Excel
2. Instalamos la macro en y en Complementos, elegimos Biplot
3. Marcamos la opción: Descomposición en valores singulares
   a. en Data Range for Y's: se marcan todas las casillas de la tabla, incluidas las etiquetas
   b. en Output Range se marca la casilla a partir de la cual se obtienen los resultados de la DVS
   c. marcamos las 2 casillas de usar las etiquetas de las filas y columnas
   d. las demás casillas de la izquierda las dejamos como están
   e. en la parte de la derecha, en 'Method' marcamos ACP y en 'Data transformation' ninguna

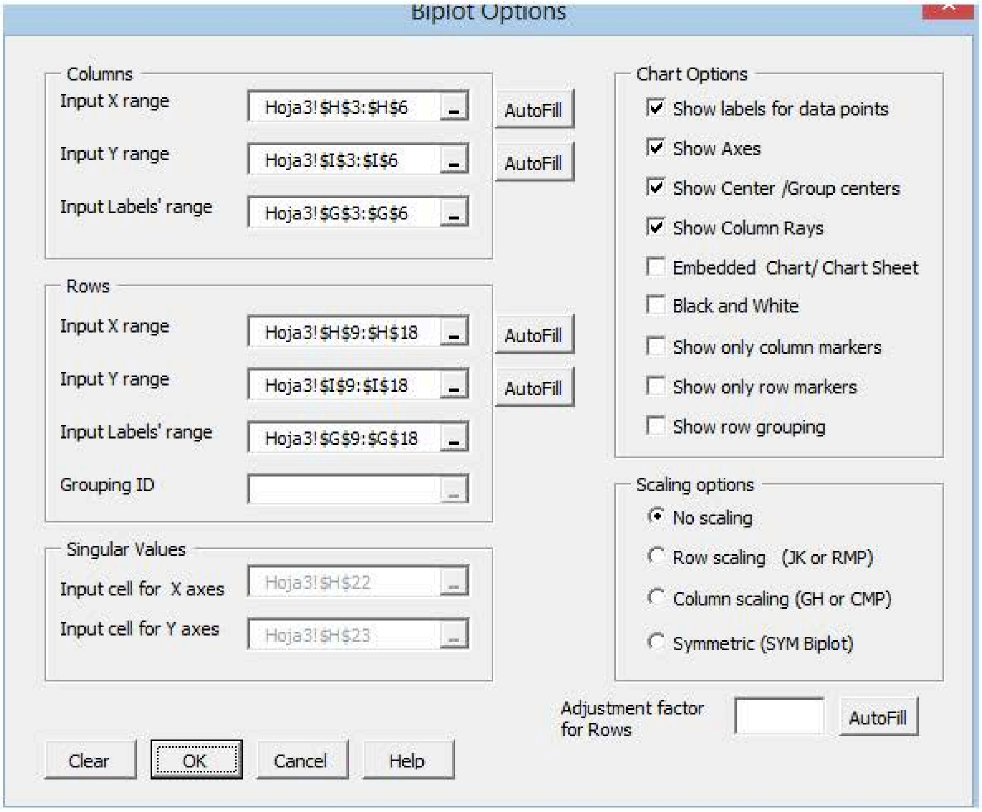


f. Los resultados obtenidos deben parecerse a éstos:

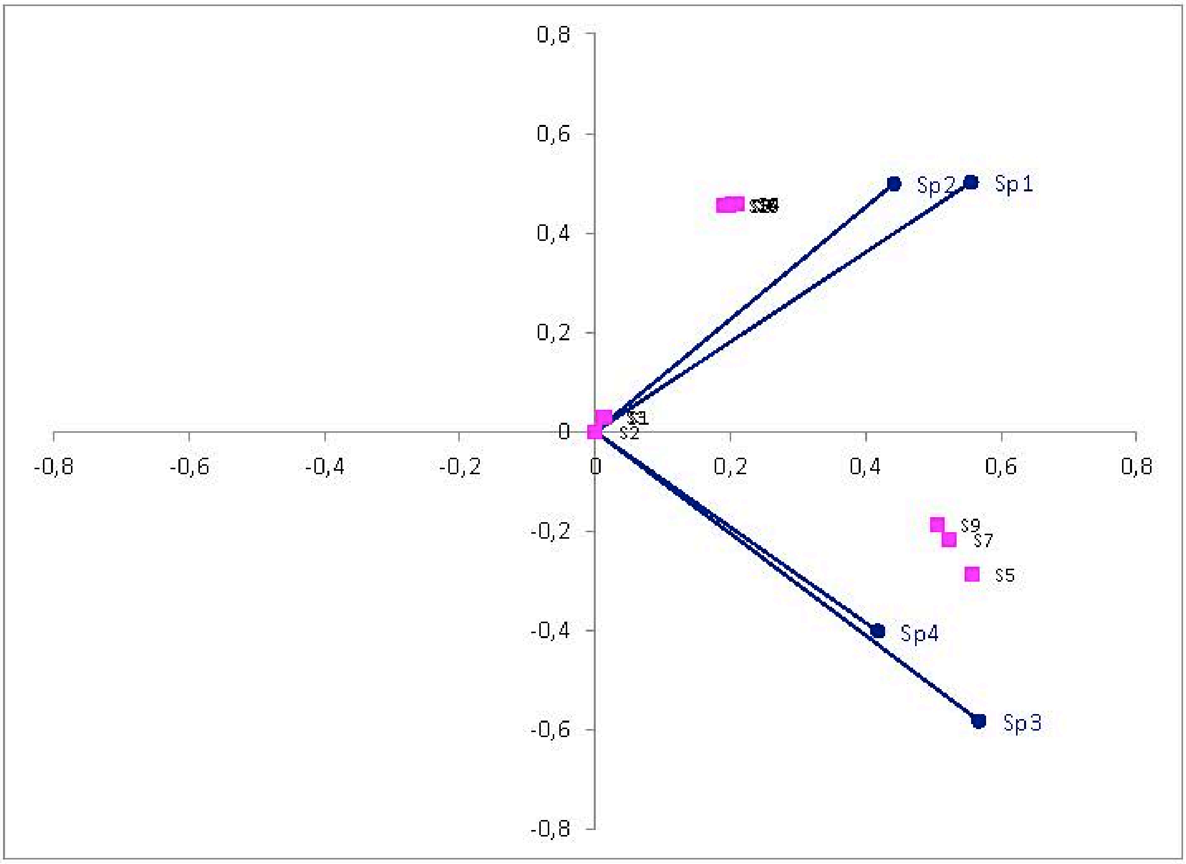
	A	B	C	D	E	F	G	H	I	J	K
1											
2		Sp1	Sp2	Sp3	Sp4		**V matrix of the U LAMBDA V' decomposition**				
3	S1	1	0	0	0		Sp1	0,55556066	0,50146916		
4	S2	0	0	0	0		Sp2	0,44098796	0,49827723		
5	S3	0	1	0	0		Sp3	0,5676054	-0,58204194		
6	S4	11	4	0	0		Sp4	0,41797857	-0,40184034		
7	S5	11	5	17	7						
8	S6	9	6	0	0		*U matrix of the U LAMBDA V' decomposition*				
9	S7	9	7	13	10		S1	0,01479619	0,03053454		
10	S8	7	8	0	0		S2	0	0		
11	S9	7	9	10	13		S3	0,01174479	0,03034018		
12	S10	5	10	0	0		S4	0,20973728	0,45724064		
13							S5	0,55639461	0,2861867		
14							S6	0,20363448	0,45685193		
15							S7	0,52321983	-0,21821727		
16							S8	0,19753167	0,45646321		
17							S9	0,50516203	-0,18568848		
18							S10	0,19142886	0,4560745		
19											
20							**Singular and eigenvalues for the SVD (U LAMBDA V')**				
21								Singular valu	Eigen values	Cumulative % of Eigenvalues	
22								37,54754	1409,81776	0,79381631	
23								16,4230146	269,715409	0,94568309	
24											
25							*Sum of eigenvalues*		1776		
26											

g. Se puede observar que las coordenadas de los vectores u y v de la descomposición en valores singulares de la tabla de partida en Excel, salen con valores opuestos a los obtenidos en R

h. Aparece a continuación la ventana de la representación Biplot, la parte de la izquierda se queda como está, y en la derecha en 'Chart Options', marcamos las 4 primeras, en 'Scaling options' marcamos ninguna,



i. El gráfico obtenido debe ser el siguiente:



j. Efectivamente el gráfico sale simétrico respecto al 2º eje del obtenido en R, por los valores opuestos de los vectores de la DVS, se puede observar una alta correlación entre los organismos Sp1 y Sp2 y entre Sp3 y Sp4, además de comportamientos similares en las parcelas (1,2,3), (5,7,9) y (4,6,8,10)