Ejemplo 1

Objetivos:

1. **Ayudar a comprender los conceptos relacionados con un análisis de componentes principales. Interpretación de resultados.**
2. **Aprender a ejecutar con R el ACP. Familiarizarse con los términos y funciones ligadas a un ACP.**

Objetivo del análisis de componentes principales: ACP:
Datos
Exploración previa al análisis ACP

Grafico caja
Correlaciones observadas entre pares de variables:
Determinante de la matriz de correlaciones

ACP
Función prcomp()
Función princomp()

Puntuaciones de las observaciones en las componentes
Gráfico de dispersión de las componentes Y1 e Y2:

Análisis mediante extracción de unas pocas componentes:

Matriz de componentes C
Gráficos de componentes
La matriz de correlaciones reproducidas y la matriz residual
Matriz de comunalidades

Rotaciones de la solución
La función varimax()
Uso de la matriz de componentes de la solución no rotada, C, como input en la función varimax

Matriz de componentes rotados (C^R):
Tabla de coeficientes para el cálculo de las puntuaciones en las componentes:

Ejemplo 1: (archivo eje1.DAT)

Un grupo constituido por 13 empresas se ha clasificado según las puntuaciones obtenidas en 8 indicadores económicos (X1 a X8).

Los datos vienen dados en la tabla siguiente:

X1: Indicador de volumen de facturación de la empresa
X2: Indicador del nivel de nueva contratación.
X3: Indicador del total de clientes
X4: Indicador de beneficios de la empresa
X5: Indicador de nivel de retribución salarial de los empleados
X6: Indicador de nivel de organización empresarial dentro de la empresa
X7: Indicador de nivel de relaciones con otras empresas
X8: Indicador de nivel de equipamiento (ordenadores, maquinaria, etc.)

Datos observados en las 13 empresas.

X1	X2	X3	X4	X5	X6	X7	X8
0,13	0,94	2,17	5,79	5,48	11,1	3,98	5,8
1,55	4,3	4,33	5,4	5,34	11,01	5,58	8,1
1,21	6,9	6,5	4,07	6,89	24,76	7,28	14,4
2,48	8,69	8,67	7,41	4,14	31,59	10,34	17,87
6,01	11,24	10,83	0,03	10,04	10,17	1,97	7,55
6,74	13,13	13	1,55	5,87	6,13	1,06	3,95
7,58	14,85	15,17	11,4	2,71	6,96	1,67	4,75
8,01	18,16	17,33	6,33	2,56	41,52	12,37	26,23
8,12	19,57	19,5	6,51	3,46	31,04	10,16	20,15
11,52	21,73	21,67	5,06	5,92	43,28	12,24	26,57
10,73	24,14	23,83	3,97	5,6	11,37	5,24	7,33
11,99	24,81	26	4,89	4,87	9,71	4,09	6,83
14,36	25,43	28,17	7,76	3,52	34,44	9,8	20,41

Objetivo del análisis:

1. Seleccionar un número pequeño de componentes que resuma las 8 variables observadas en unas pocas dimensiones latentes, procurando que la información perdida no revista mucha importancia.

2. Intentar interpretar las componentes principales derivadas del análisis

Datos

```
>#8 variables x1 a x8 observadas en 13 empresas. No olvide usar separador decimal ,
>a=read.table("eje1.DAT", header=T, sep="\t",dec=",")
>a=round(a,2)
```

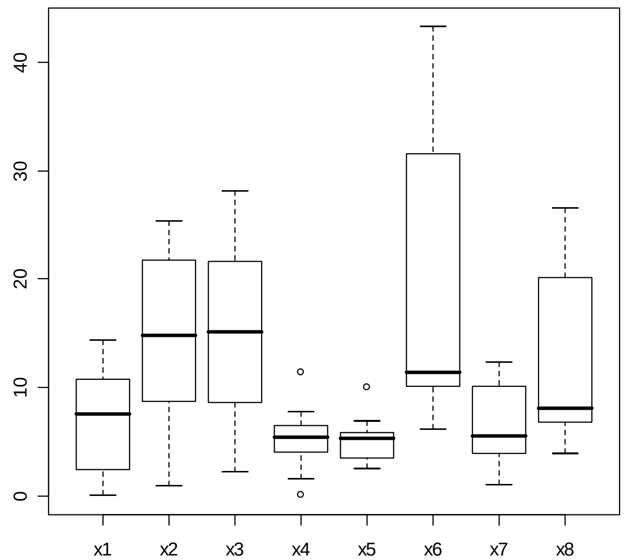
Partimos de los datos, matriz X de 13 filas (empresas) y 8 columnas (indicadores observados).

Exploración previa al análisis ACP

Grafico caja

Dado que se admite que las escalas de los indicadores son más o menos arbitrarias se decide trabajar con la matriz de correlaciones1 [1].

>boxplot(a)



Correlaciones observadas entre pares de variables:

Una exploración previa de las correlaciones nos permite ver si en nuestros datos podrían encontrarse unas componentes capaces de recoger parte de la variabilidad. Puede comprobarse que determinados pares de variables están altamente correlacionadas. En consecuencia, podría tener interés seleccionar, entre unas pocas componentes, la mayor parte de la información original, eliminando las redundancias o variabilidad compartida por grupos de variables.
Veamos primero la matriz implicada en los cálculos: **R** (correlaciones entre las variables originales observadas).
La matriz de correlaciones R viene dada por:

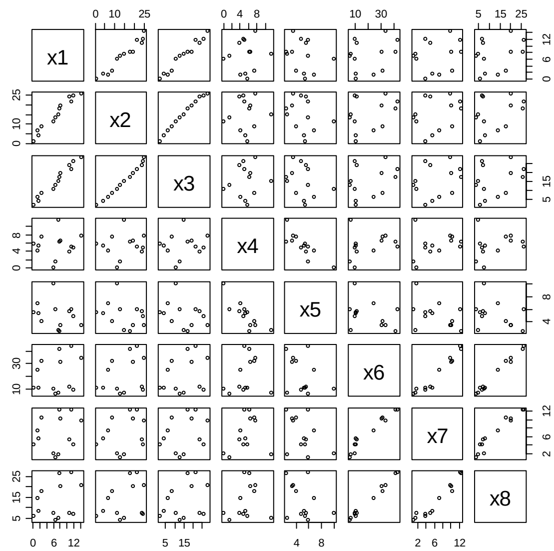
1 Sugerencia: Un modo visual de explorar las diferencias existentes en la variabilidad de las distintas variables es representar gráficamente en un mismo espacio mediante diagramas de caja.

>round(cor(a),2)

En este ejemplo simple, puede observarse que las variables X1, X2 y X3 están altamente correlacionadas entre sí; por otro lado, lo están las variables X4 y X5, y, por último el grupo formado por las variables X6, X7 y X8.

```
      x1    x2    x3    x4    x5    x6    x7    x8
x1  1.00  0.97  0.98  0.10 -0.26  0.26  0.21  0.30
x2  0.97  1.00  0.99  0.11 -0.31  0.31  0.29  0.35
x3  0.98  0.99  1.00  0.15 -0.33  0.30  0.28  0.33
x4  0.10  0.11  0.15  1.00 -0.84  0.25  0.30  0.23
x5 -0.26 -0.31 -0.33 -0.84  1.00 -0.35 -0.41 -0.34
x6  0.26  0.31  0.30  0.25 -0.35  1.00  0.97  0.99
x7  0.21  0.29  0.28  0.30 -0.41  0.97  1.00  0.96
x8  0.30  0.35  0.33  0.23 -0.34  0.99  0.96  1.00
```

>plot(a) #Visualización de las relaciones entre pares de variables



Si las 8 variables observadas no estuviesen correlacionadas no tendría interés realizar un ACP, pues sería necesario generar tantas componentes como variables existentes para no perder información esencial. Hemos realizado una primera inspección a la matriz de correlaciones y se han observado correlaciones significativas entre las variables.

Determinante de la matriz de correlaciones

Otro indicador adecuado para ver si las variables están correlacionadas es obtener el determinante de la matriz de correlaciones. Si está próximo a cero indica que hay fuerte redundancia en la información registrada en las variables (se dice que están relacionadas).

>det(cor(a))

```
[1] 1.31792e-08
```

Las 3 primeras componentes tienen varianza superior a 1, tal como muestra el resultado siguiente. La cuarta componente presenta un descenso importante en la varianza (con un valor igual a 0.38):

```
> acp=prcomp(a,scale=TRUE)
> acp
```

```
Standard deviations:
[1] 2.04999203 1.45922413 1.20534172 0.38122139 0.21899191 0.11810212 0.08068302
[8] 0.04004820

Rotation:
          PC1        PC2         PC3          PC4        PC5         PC6
x1  0.3551761 -0.4594807  0.03418166 -0.175477344  0.4262555  0.60639707
x2  0.3769254 -0.4286792  0.03801483  0.070933220 -0.2656430 -0.57860281
x3  0.3763377 -0.4334478  0.00620033 -0.005478072 -0.2364483 -0.01686902
x4  0.2140581  0.2014459 -0.67199135 -0.664424455 -0.1253127 -0.05404693
x5 -0.2942276 -0.1413771  0.59850961 -0.699072953 -0.2124654 -0.02748401
x6  0.3894736  0.3446872  0.26238969 -0.118028764  0.2985259 -0.16069794
x7  0.3858970  0.3672919  0.20563146  0.123468724 -0.6503276  0.45248395
x8  0.3964218  0.3208648  0.27645845 -0.069339783  0.3448821 -0.25094786
           PC7          PC8
x1 -0.18198549  0.219628939
x2 -0.26313489  0.439428491
x3  0.45614646 -0.637308817
x4 -0.04376550  0.005167487
x5 -0.02476905 -0.003932809
x6  0.63954673  0.350371723
x7 -0.13749489  0.109903482
x8 -0.50903627 -0.466499638
```

La matriz Rotation muestra los vectores propios asociados a los componentes principales. Son los coeficientes que proporcionan la combinación lineal de las variables Xi que dan lugar a cada una de las componentes. Por ejemplo, la primera componente Y1 se obtiene mediante la combinación lineal:

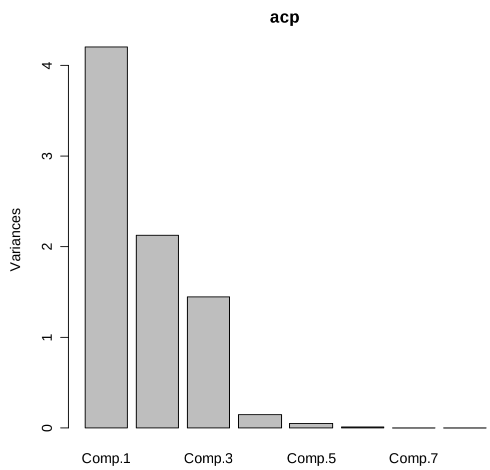
Y1 = `0.3551761 x1 +  0.3769254 x2 + …+ 0.3964218 x8`

La función summary permite realizar un resumen del análisis destacando la importancia de cada componente en el conjunto mediante el porcentaje de variabilidad que captura:

```
> summary(acp)
```

```
Importance of components:
                         PC1   PC2   PC3    PC4     PC5     PC6     PC7    PC8
Standard deviation     2.050 1.459 1.205 0.3812 0.21899 0.11810 0.08068 0.0400
Proportion of Variance 0.525 0.266 0.182 0.0182 0.00599 0.00174 0.00081 0.0002
Cumulative Proportion  0.525 0.791 0.973 0.9912 0.99724 0.99899 0.99980 1.0000
```

El gráfico siguiente muestra visualmente la importancia de cada componente en un gráfico de barras. Las alturas son las varianzas de las correspondientes componentes:



Según la ***Varianza total explicada*** la primera componente Y_1 tiene una varianza $V(Y_1)$ = 4,202 igual al primer autovalor. $V(Y_2)$ = 2,130 y $V(Y_3)$ = 1,453.

```
> predict(acp)  #Proporciona las puntuaciones en las componentes principales Y
             PC1        PC2        PC3          PC4         PC5          PC6
 [1,] -2.6635778  1.3254091 -0.6725289 -0.007877723  0.02462217  0.134964532
 [2,] -2.0471813  1.1095352 -0.4383188  0.138268981 -0.14259286  0.204099471
 [3,] -1.3343457  1.4316100  0.9174669 -0.177057283 -0.13901828 -0.188189303
 [4,]  0.2564168  2.1184313 -0.2819989 -0.026118381 -0.19530523 -0.018938842
 [5,] -2.5867596 -1.1318569  2.1024888 -0.462593645  0.14505663 -0.047087571
 [6,] -1.9901966 -1.3326656  0.2621269  0.664380683  0.37800622 -0.005820114
 [7,] -0.4146824 -0.5833204 -2.9419248 -0.556015653  0.19264898 -0.130032932
 [8,]  2.5417560  1.4095469  0.1648108  0.552251799  0.22774004 -0.066823148
 [9,]  1.7914983  0.4628199 -0.1160498  0.270908453 -0.09869660 -0.114171791
[10,]  2.6333447  0.3594584  1.5592602 -0.458464027  0.08784772  0.070619805
[11,]  0.2487181 -2.0365336  0.1168405  0.183675749 -0.40630641 -0.012524457
[12,]  0.4694430 -2.3576761 -0.4123375  0.161514217 -0.20655122  0.001158728
[13,]  3.0955665 -0.7747582 -0.2598355 -0.282873170  0.13254885  0.172745622
              PC7          PC8
 [1,]  0.08899993 -0.012315001
 [2,] -0.15179307 -0.015020696
 [3,]  0.08668485 -0.021462307
 [4,]  0.07686015  0.044519754
 [5,] -0.06102748 -0.024861565
 [6,]  0.06102395  0.034957575
 [7,] -0.05640002  0.021350829
 [8,] -0.06924813  0.003955654
 [9,] -0.05245509 -0.063484334
[10,] -0.04688297  0.050912887
[11,] -0.01064770  0.064022812
[12,]  0.02246169 -0.045622684
[13,]  0.11242390 -0.036952925
```

Observe que las predicciones se obtienen igualmente mediante la operación matricial:

```
> ZX=scale(a[,1:8]) #variables x1 a x8 estandarizadas
> ZX%*%acp1$rotation
```

Dado que son las combinaciones lineales de las variables X (estandarizadas) descritas por los vectores propios.
Estas puntuaciones variarán ligeramente de las obtenidas con princomp, dado que en la matriz de covarianzas los denominadores N se sustituyen por N-1. Cuando el tamaño muestral es grande los resultados son prácticamente iguales.

Estas puntuaciones, que definen las nuevas variables Y, tienen interés, especialmente las de mayor varianza, tanto para comparar casos como para utilizarlas en otros análisis, sustituyendo las variables X originales.

Lógicamente están incorreladas, la matriz de covarianza es:

> round(cov(predict(acp)),3)

```
      PC1  PC2   PC3   PC4   PC5   PC6   PC7   PC8
PC1 4.202 0.00 0.000 0.000 0.000 0.000 0.000 0.000
PC2 0.000 2.13 0.000 0.000 0.000 0.000 0.000 0.000
PC3 0.000 0.00 1.453 0.000 0.000 0.000 0.000 0.000
PC4 0.000 0.00 0.000 0.145 0.000 0.000 0.000 0.000
PC5 0.000 0.00 0.000 0.000 0.048 0.000 0.000 0.000
PC6 0.000 0.00 0.000 0.000 0.000 0.014 0.000 0.000
PC7 0.000 0.00 0.000 0.000 0.000 0.000 0.007 0.000
PC8 0.000 0.00 0.000 0.000 0.000 0.000 0.000 0.002
```

Función princomp()

La función princomp() de R permite también realizar el mismo análisis:

>acp=princomp(a,cor=TRUE)#equivale a estandarizar#acp2=prcomp(a, scale=T)
>summary(acp)

```
Importance of components:
                          Comp.1    Comp.2    Comp.3     Comp.4      Comp.5
Standard deviation     2.0499036 1.4594230 1.2052970 0.38124395 0.218726278
Proportion of Variance 0.5252631 0.2662394 0.1815926 0.01816837 0.005980148
Cumulative Proportion  0.5252631 0.7915026 0.9730952 0.99126353 0.997243674
                            Comp.6       Comp.7       Comp.8
Standard deviation     0.118072199 0.0806303271 0.0401038291
Proportion of Variance 0.001742631 0.0008126562 0.0002010396
Cumulative Proportion  0.998986304 0.9997989604 1.0000000000
```

Los vectores propios de la matriz de correlaciones de X con la función **princomp** son las columnas de la matriz Loadings

> loadings(acp)# Vectores propios: coeficientes de combinaciones lineales que proporcionan componentes

```
Loadings:
   Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8
x1 -0.355 -0.460        -0.177 -0.426  0.607  0.179  0.220
x2 -0.377 -0.429                0.265 -0.578  0.265  0.440
x3 -0.376 -0.434                0.237        -0.455 -0.638
x4 -0.214  0.202  0.672 -0.664  0.127
x5  0.294 -0.142 -0.598 -0.698  0.214
x6 -0.390  0.344 -0.263 -0.119 -0.298 -0.163 -0.640  0.349
x7 -0.386  0.367 -0.206  0.125  0.650  0.453  0.136  0.111
x8 -0.397  0.320 -0.277        -0.345 -0.249  0.511 -0.466

               Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8
SS loadings     1.000  1.000  1.000  1.000  1.000  1.000  1.000  1.000
Proportion Var  0.125  0.125  0.125  0.125  0.125  0.125  0.125  0.125
Cumulative Var  0.125  0.250  0.375  0.500  0.625  0.750  0.875  1.000
```

Puntuaciones en las componentes de las 13 observaciones

Las puntuaciones de las observaciones o casos en las componentes se obtienen mediante la función predict(). También se pueden obtener predicciones para nuevos datos.

> predict(acp)#equivalente acp$scores son las puntuaciones de los datos sobre las comp princ

```
          Comp.1     Comp.2     Comp.3       Comp.4      Comp.5       Comp.6
 [1,]  2.7723397  1.3795296  0.6999903 -0.008199394 -0.02562756  0.140475539
 [2,]  2.1307739  1.1548408  0.4562167  0.143914918  0.14841536  0.212433464
 [3,]  1.3888311  1.4900669 -0.9549298 -0.184287063  0.14469482 -0.195873636
 [4,] -0.2668871  2.2049332  0.2935138 -0.027184873  0.20328013 -0.019712172
 [5,]  2.6923847 -1.1780740 -2.1883397 -0.481482732 -0.15097973 -0.049010298
 [6,]  2.0714623 -1.3870823 -0.2728304  0.691509340 -0.39344135 -0.006057766
 [7,]  0.4316151 -0.6071391  3.0620524 -0.578719440 -0.20051542 -0.135342567
 [8,] -2.6455435  1.4671029 -0.1715405  0.574801897 -0.23703934 -0.069551738
 [9,] -1.8646505  0.4817182  0.1207885  0.281970458  0.10272668 -0.118833768
[10,] -2.7408720  0.3741362 -1.6229295 -0.477184488 -0.09143481  0.073503424
[11,] -0.2588740 -2.1196914 -0.1216115  0.191175781  0.42289711 -0.013035868
[12,] -0.4886117 -2.4539470  0.4291744  0.168109327  0.21498533  0.001206043
[13,] -3.2219678 -0.8063939  0.2704453 -0.294423730 -0.13796122  0.179799344
           Comp.7       Comp.8
 [1,] -0.09263407 -0.012817859
 [2,]  0.15799123 -0.015634036
 [3,] -0.09022445 -0.022338677
 [4,] -0.07999858  0.046337629
 [5,]  0.06351942 -0.025876737
 [6,] -0.06351574  0.036384997
 [7,]  0.05870301  0.022222647
 [8,]  0.07207574  0.004117176
 [9,]  0.05459699 -0.066076590
[10,]  0.04879735  0.052991813
[11,]  0.01108247  0.066637055
[12,] -0.02337887 -0.047485595
[13,] -0.11701450 -0.038461824
```

De modo equivalente se obtienen las puntuaciones a través de la matriz de vectores propios mediante la operación matricial:

>ZX=scale(a[,1:8]) #variables x1 a x8 estandarizadas

>A=acp$loadings[,]#sus columnas son los vectores propios de matriz de correlaciones de X

> Y=ZX%*%A #puntuaciones de los casos en las componentes. Y=ZX*A

NOTA: Las puntuaciones Y difieren de las predichas por R en la cte sqrt(n/n-1)
Es decir, Y*sqrt(13/12)=predict(acp)

Las variables nuevas, Y, están incorreladas y sin estandarizar:

> round(cov(predict(acp)),3)

```
       Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8
Comp.1  4.552  0.000  0.000  0.000  0.000  0.000  0.000  0.000
Comp.2  0.000  2.307  0.000  0.000  0.000  0.000  0.000  0.000
Comp.3  0.000  0.000  1.574  0.000  0.000  0.000  0.000  0.000
Comp.4  0.000  0.000  0.000  0.157  0.000  0.000  0.000  0.000
Comp.5  0.000  0.000  0.000  0.000  0.052  0.000  0.000  0.000
Comp.6  0.000  0.000  0.000  0.000  0.000  0.015  0.000  0.000
Comp.7  0.000  0.000  0.000  0.000  0.000  0.000  0.007  0.000
Comp.8  0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.002
```

Las puntuaciones estandarizadas en las 3 primeras componentes se pueden obtener mediante la función scale()

> scale(predict(acp)[,1:3])

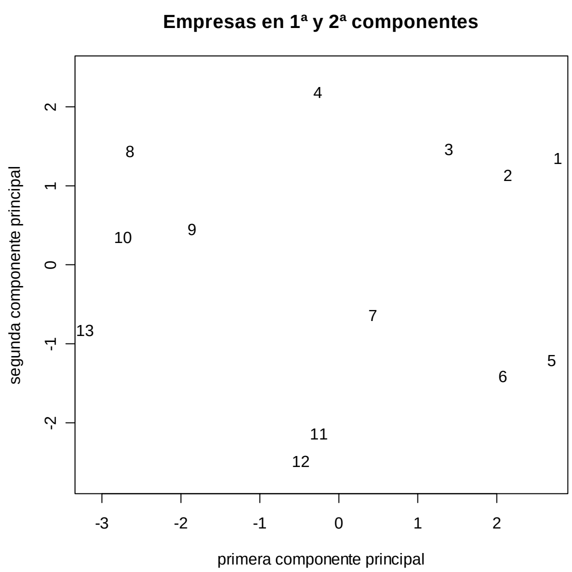
```
          Comp.1     Comp.2      Comp.3
 [1,]  1.2993673  0.9081734  0.55797774
 [2,]  0.9986720  0.7602560  0.36366044
 [3,]  0.6509310  0.9809424 -0.76119572
 [4,] -0.1250873  1.4515540  0.23396633
 [5,]  1.2618933 -0.7755509 -1.74437406
 [6,]  0.9708732 -0.9131455 -0.21747914
 [7,]  0.2022936 -0.3996925  2.44082983
 [8,] -1.2399393  0.9658247 -0.13673877
 [9,] -0.8739427  0.3171252  0.09628317
[10,] -1.2846188  0.2463017 -1.29367307
[11,] -0.1213316 -1.3954375 -0.09693921
[12,] -0.2290073 -1.6154850  0.34210446
```

```
[13,] -1.5101034 -0.5308661  0.21557797
attr(,"scaled:center")
      Comp.1       Comp.2       Comp.3
1.750736e-16 1.281027e-16 2.070993e-16
attr(,"scaled:scale")
  Comp.1   Comp.2   Comp.3
2.133607 1.519016 1.254513
```

Gráfico de dispersión de las componentes Y1 e Y2:

Pueden realizarse gráficos con las puntuaciones en las componentes principales y visualizar qué casos destacan en determinadas componentes:

```
require(MASS)
eqscplot(predict(acp)[,1:2],type="n",xlab="primera componente principal"
,ylab="segunda componente principal")#Gráfico que usa la misma escala en los ejes
text(predict(acp) [,1:2],labels=as.character(row.names(a)))
title(main="Empresas en 1ª y 2ª componentes")
```



Las puntuaciones de las variables en 2 primeras componentes recogen el 79,1% de variabilidad de los datos. Estas nuevas variables están incorrelacionadas entre sí.
Las empresas número 4 y 12 son las más diferenciadas en la segunda componente.
Las empresas 13,10 y 8 son las más alejadas de las empresas 1 y 5 en la componente primera.
A veces es útil guardar los valores que presentan los casos en las componentes principales, como un nuevo conjunto de variables incorrelacionadas que puede ser usado en análisis posteriores. Otras veces puede interesar conocer el valor concreto alcanzado por determinados casos en una o más componentes. Así por ejemplo, si la componente considerada es la Y1 (vea tabla de puntuaciones estandarizadas de los casos en las componentes), el último caso presenta la puntuación más baja (-1,5101) y el primer caso la más alta (1,29937). Es importante que tenga presente que el signo en la escala podría haber resultado al contrario y por tanto la puntuación tomaría los signos cambiados. Lo único que debe tenerse en cuenta es la contraposición (en extremos de la escala) pero no el signo en sí mismo.

Análisis mediante extracción de unas pocas componentes:

De modo similar a como se procede en un análisis factorial, extraeremos del conjunto la máxima variabilidad posible con el menor número posible de componentes. Aquí puede verse que el 97,3 % de la variabilidad de los datos se encuentra representado por las 3 primeras componentes.

Matriz de componentes C

```
>ZX=scale(a[,1:8]) #variables x1 a x8 estandarizadas
>A=acp$loadings[,]#sus columnas son los vectores propios de matriz de correlaciones de X
> Y=ZX%*%A  #puntuaciones de los casos en las componentes.  Y=ZX*A
```

La **matriz de componentes** contiene las **correlaciones** entre las **variables originales (ZX)** y las **componentes** (Y). Se puede obtener, por tanto con la función cor, cor(X,Y)

Tomamos las 3 primeras componentes (97,3% de variabilidad)

```
> C=cor(a,predict(acp)[,1:3]) #Matriz de componentes
> C

       Comp.1     Comp.2       Comp.3
x1 -0.7280141 -0.6706139 -0.040499311
x2 -0.7725832 -0.6257327 -0.045165677
x3 -0.7712993 -0.6327374 -0.006976095
x4 -0.4386716  0.2949344  0.809722115
x5  0.6031344 -0.2073072 -0.721199455
x6 -0.7984840  0.5024789 -0.316867759
x7 -0.7910934  0.5358628 -0.248069440
x8 -0.8128367  0.4676518 -0.333601277
```

La ***matriz de componentes*** muestra las correlaciones entre cada variable y cada una de las 3 componentes.
Por ejemplo, la variable X1 presenta coeficientes de correlación relativamente altos con las componentes 1 y 2 (-0,728 y -0,671, respectivamente) y correlación insignificante con la componente 3 (igual a -0,0405).
Puede observarse que no hay una estructura clara. Por ejemplo, las variables X1, X2 y X3 presentan correlaciones altas en las componentes 1 y 2.

Prácticamente todas las variables están correlacionadas con la primera componente. En menor grado lo está X4. La componente 2 contrasta el grupo de variables X1, X2 y X3 con el constituido por las variables X6, X7 y X8. Pero la existencia de variables altamente correlacionadas simultáneamente con dos o más componentes supone un inconveniente para tratar de explicar su significado.

Normalmente, es preciso efectuar una rotación de los ejes en la solución inicial para intentar mejorar la interpretación de las componentes.

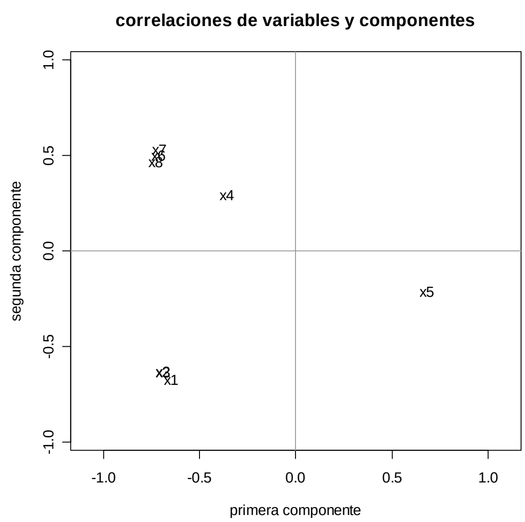
Los **gráficos de componentes** visualizan la tabla anterior. Los presentamos en planos, de modo que puedan interpretarse más claramente. Por ejemplo, puede observarse que las variables X1, X2 y X3 están situadas "cerca" de la bisectriz formada por el plano constituido por las componentes 1 y 2 y próximas al origen en la componente 3.

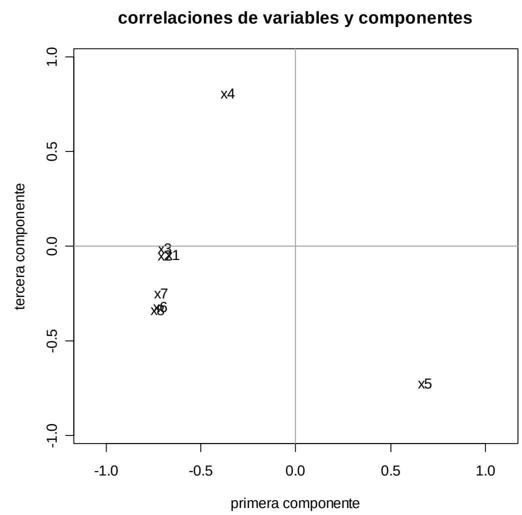
Gráficos de componentes

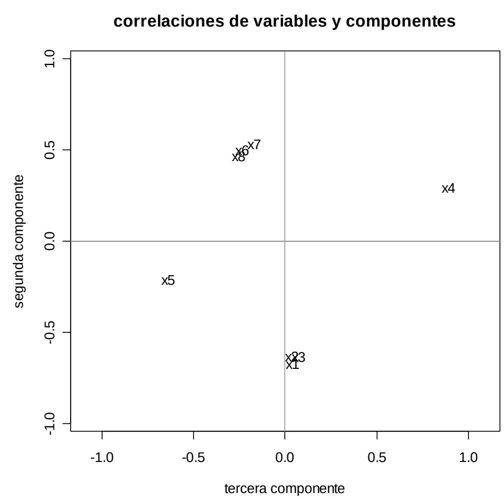
```
#Gráficos de componentes
eqscplot(C[,1],C[,2],type="n",xlim=c(-1,1),ylim=c(-1,1),
 xlab="primera componente",ylab="segunda componente")
text(x=C[,1],y=C[,2],labels=row.names(C),pos=4)
title(main="correlaciones de variables y componentes")
abline(h=0, v=0, col = "gray60")


eqscplot(C[,1],C[,3],type="n",xlim=c(-1,1),ylim=c(-1,1),
 xlab="primera componente",ylab="tercera componente")
text(x=C[,1],y=C[,3],labels=row.names(C),pos=4)
title(main="correlaciones de variables y componentes")
abline(h=0, v=0, col = "gray60")

eqscplot(C[,3],C[,2],type="n",xlim=c(-1,1),ylim=c(-1,1),
 xlab="tercera componente",ylab="segunda componente")
text(x=C[,3],y=C[,2],labels=row.names(C),pos=4)
title(main="correlaciones de variables y componentes")
abline(h=0, v=0, col = "gray60")
```







La matriz de componentes puede ser muy útil para interpretar los componentes, así como su representación gráfica en el **Gráfico de componentes**.

La matriz de correlaciones reproducidas y la matriz residual

Si comparamos la matriz de correlaciones reproducidas con la matriz de correlaciones observadas (su diferencia es la matriz residual) podemos comprobar que están muy próximas. Esto es un buen indicador de que la variabilidad captada por las componentes permite resumir adecuadamente la información contenida en las variables (X’s) observadas. Sólo un 3% de los residuos supera el valor 0,05 en términos absolutos (hay 1 residuo, no redundante, que los supera en el conjunto).
La matriz reproducida se obtiene mediante el producto de matrices:
R* = CC’
donde C representa la matriz de componentes.

```
> Rrep=C%*%t(C)#Correlaciones reproducidas
> Rrep
            x1         x2         x3          x4         x5         x6
x1  0.98136771  0.9839057  0.9861218  0.08877884 -0.2708592  0.2571713
x2  0.98390568  0.9904661  0.9921324  0.11778858 -0.3036792  0.3167895
x3  0.98612175  0.9921324  0.9953078  0.14608239 -0.3289950  0.3001435
x4  0.08877884  0.1177886  0.1460824  0.93506894 -0.9096911  0.2418957
x5 -0.27085921 -0.3036792 -0.3289950 -0.90969107  0.9268760 -0.3572358
x6  0.25717127  0.3167895  0.3001435  0.24189572 -0.3572358  0.9904670
x7  0.22661680  0.2870829  0.2728400  0.30420724 -0.4093163  0.9795404
x8  0.29165343  0.3504263  0.3333668  0.22437060 -0.3466043  0.9897297
           x7         x8
x1  0.2266168  0.2916534
x2  0.2870829  0.3504263
x3  0.2728400  0.3333668
x4  0.3042072  0.2243706
x5 -0.4093163 -0.3466043
x6  0.9795404  0.9897297
x7  0.9745162  0.9763832
x8  0.9763832  0.9906915
```

```
> Resi=cor(a)-Rrep  #diferencias entre la matriz R de correlaciones de los datos y matriz reproducida
```

```
 > Resi
             x1           x2           x3           x4           x5
x1  0.018632288 -0.011670970 -0.005611114  0.014057426  0.013358621
x2 -0.011670970  0.009533859  0.001868809 -0.004799928 -0.004302133
x3 -0.005611114  0.001868809  0.004692163  0.001786045  0.002859325
x4  0.014057426 -0.004799928  0.001786045  0.064931063  0.068752697
x5  0.013358621 -0.004302133  0.002859325  0.068752697  0.073123983
x6  0.007119457 -0.004559013 -0.001721108  0.009601866  0.008960217
x7 -0.012422185  0.006215123  0.006649191 -0.008429051 -0.006184500
x8  0.007151329 -0.002554137 -0.004825921  0.005021547  0.003782750
             x6           x7           x8
x1  0.007119457 -0.012422185  0.007151329
x2 -0.004559013  0.006215123 -0.002554137
x3 -0.001721108  0.006649191 -0.004825921
x4  0.009601866 -0.008429051  0.005021547
x5  0.008960217 -0.006184500  0.003782750
x6  0.009533026 -0.012965066  0.004310092
x7 -0.012965066  0.025483813 -0.013196832
x8  0.004310092 -0.013196832  0.009308547
```

Solo un residuo no redundante es mayor en valor absoluto a 0.05.

Matriz de comunalidades:

El valor de la **comunalidad** de cada variable de la matriz de datos X, lo podemos interpretar como el coeficiente de correlación múltiple entre cada variable observada (X_i) y todas las componentes principales. Su valor es 1, si tomamos todas las componentes principales, dado que toda variable X_i puede expresarse de modo exacto como combinación lineal de las componentes.
Su valor cuando se realiza la **Extracción** de sólo unas pocas componentes refleja también el coeficiente R^2, que tiene cada variable X_i, pero ahora en función de sólo las componentes extraídas (las 3 primeras en el ejemplo). Se interpreta como la proporción de varianza explicada por las 3 componentes[2]. Se obtiene a partir de la **matriz de componentes** (C). Por ejemplo, a la variable X_1 le corresponde una **comunalidad** igual a:

---

[2] Dado que las componentes principales, que asumen el papel de variables independientes en su relación con X_1, son ortogonales y que los coeficientes asociados a ellas juegan el papel de pendientes estandarizadas, la suma de sus valores al cuadrado coincide con R^2

$0{,}981 = 0{,}728^2 + 0{,}671^2 + 0{,}0405^2$

#Comunalidades correspondientes a las 8 variables
o equivalente suma de cuadrados de cada fila de C

```
>apply(C*C,1,sum)

       x1        x2        x3        x4        x5        x6        x7        x8
0.9813677 0.9904661 0.9953078 0.9350689 0.9268760 0.9904670 0.9745162 0.9906915
```

(Vea la fila de la matriz de componentes asociada a X_1 (la primera)).

Puede observarse que todas las variables presentan valores altos en la comunalidad (con extracción de 3 componentes). Por tanto, todas están bien representadas por las 3 componentes extraídas. La proporción de variabilidad de cada variable que es explicada por las componentes supera en todas ellas el 90%. En particular X_3 presenta una comunalidad de 0,995.

En resumen, las comunalidades de todas las variables presentan valores altos. Esto es un buen resultado del análisis porque indica que la variabilidad presente en cada variable observada está compartida casi en su totalidad por las 3 componentes extraídas. Por tanto, si uno de nuestros objetivos fuese sustituir las 8 variables por sólo las 3 componentes, puede admitirse que se perdería poca información.

Rotaciones de la solución

R permite rotar la solución de modo que sea más interpretable.

La matriz de comunalidades lógicamente es la misma. La solución se busca en un espacio de la misma dimensión (3).

La función varimax()

La función **varimax** proporciona los **loadings** rotados[3] y la **matriz de rotación**.

El input lo constituyen los loadings o la matriz de componentes de la solución no rotada.

Uso de la matriz de componentes de la solución no rotada, C, como input en la función varimax

>rota=varimax(C, normalize = T) # C es la matriz de componentes (correlaciones entre X e Y) en solución no rotada, ya obtenida en párrafos anteriores.

[3] Obtenidos como el producto de los loadings no rotados por la matriz de rotación

```
> rota$loadings
Loadings:
   Comp.1 Comp.2 Comp.3
x1 -0.110 -0.983
x2 -0.169 -0.977
x3 -0.147 -0.980  0.117
x4 -0.118         0.960
x5  0.215  0.194 -0.918
x6 -0.976 -0.145  0.130
x7 -0.961 -0.111  0.197
x8 -0.972 -0.181  0.112

               Comp.1 Comp.2 Comp.3
SS loadings     2.943  2.984  1.857
Proportion Var  0.368  0.373  0.232
Cumulative Var  0.368  0.741  0.973
> rota$rotmat
           [,1]       [,2]       [,3]
[1,]  0.6745165 0.64268537 -0.3632947
[2,] -0.5950056 0.76456137  0.2478190
[3,]  0.4370307 0.04900444  0.8981106
```

Matriz de componentes rotados (C^R):

Para encontrar una solución más interpretable se ha efectuado la rotación o giro de los ejes coordenados (que representan a las componentes) tal que las distintas variables, representadas por puntos cuyas coordenadas constituyen los pesos o elementos de la matriz de componentes, "caigan" o se sitúen de forma que se organicen subgrupos claramente definidos y próximos a diferentes ejes.

Esta matriz se obtiene a partir de la matriz de componentes sin rotar C mediante el producto:

$C^R = C\,T$

Donde la matriz T es la constituida por los coeficientes que definen la rotación ejercida (rota$rotmat).

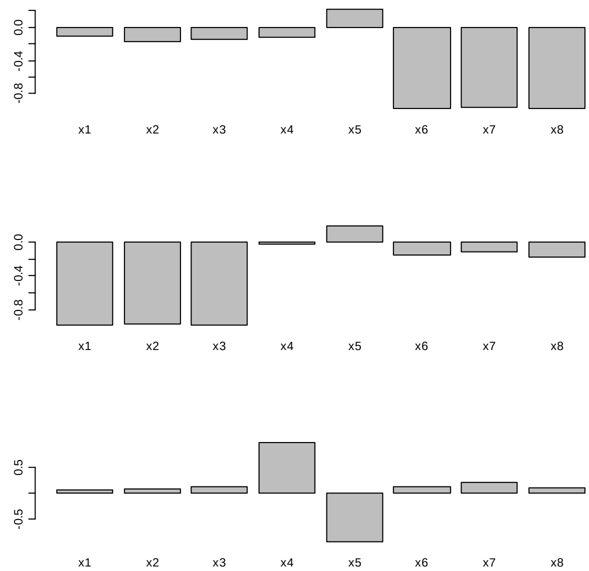
Varimax proporciona directamente la **matriz de componentes rotadas** (rota$loadings).

```
> rota$loadings

Loadings:
   Comp.1 Comp.2 Comp.3
x1 -0.110 -0.983
x2 -0.169 -0.977
x3 -0.147 -0.980  0.117
x4 -0.118         0.960
x5  0.215  0.194 -0.918
x6 -0.976 -0.145  0.130
x7 -0.961 -0.111  0.197
x8 -0.972 -0.181  0.112
```

Tal como se aprecia en la tabla anterior encontramos una estructura simple: Las variables X1 X2 y X3 están altamente correlacionadas con la primera componente. Las variables X6, X7 y X8, lo están con la componente segunda, las variables X4 y X5, con la tercera.
Nota: observe que la componente más importante está en la tabla anterior en la segunda columna.

Representación gráfica de las columnas de la matriz de componentes rotados:



Las **saturaciones al cuadrado** correspondientes a las componentes. (Observe que la componente renombrada como 2 acumula mayor varianza)

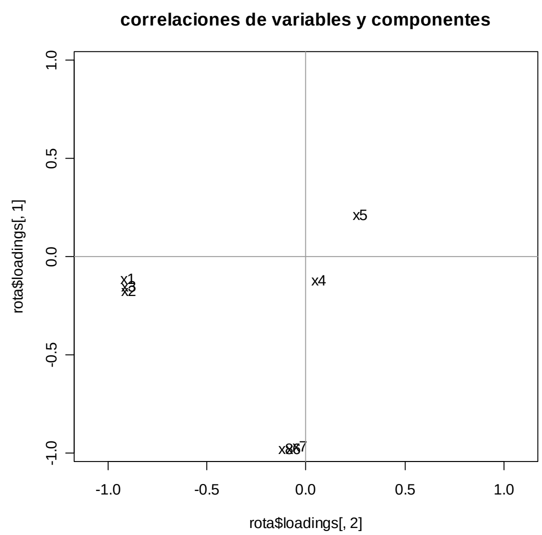
```
               Comp.1 Comp.2 Comp.3
SS loadings     2.943  2.984  1.857
Proportion Var  0.368  0.373  0.232
Cumulative Var  0.368  0.741  0.973
```

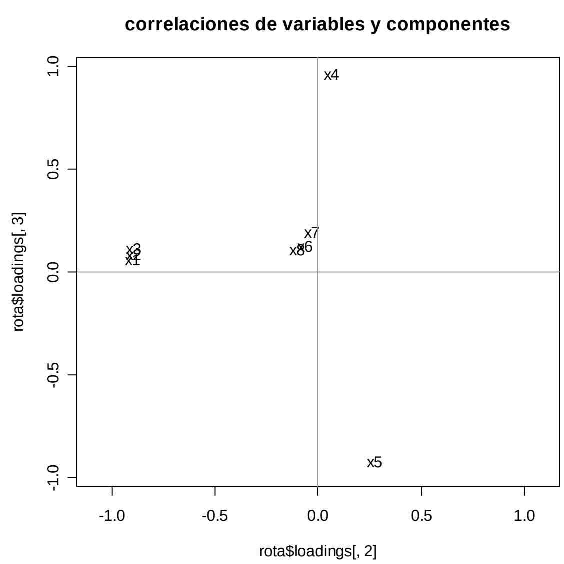
La suma de las saturaciones al cuadrado de las 3 componentes es la misma, pero la variabilidad capturada por cada componente difiere:
En particular, la primera componente pasa de un valor igual a 4,202(sin rotación) a otro igual a 2,984, disminuyendo la variabilidad capturada (en porcentajes: 52,526 frente a 37,3); por el contrario, la componente segunda pasa de describir una variabilidad igual a 2,130 (sin rotar) a 2,943, aumentando su capacidad explicativa (en porcentajes: 26,624 frente a 36,8). Lógicamente el porcentaje acumulado (explicado por las 3 componentes) es el mismo en ambas soluciones. La variabilidad de los puntos en el espacio tridimensional es la misma. Al cambiar los ejes, cambia la variabilidad relativa a cada eje, pero la total es la misma. Geométricamente nos imaginamos las variables ZX_i como puntos en un espacio tridimensional cuyos ejes son las componentes principales. La distancia al cuadrado de cada punto al origen es la comunalidad. Ésta permanece invariante al girar los ejes; lo que sí cambia es la proyección del punto sobre cada nuevo eje rotado.

A diferencia de la matriz de componentes sin rotar, no hay ninguna variable que esté altamente correlacionada con más de una componente. Esto facilita la interpretación de las componentes que habrá de efectuarse teniendo en cuenta la naturaleza de cada subgrupo de variables relacionadas con ellas.

Puede observarse, por ejemplo, que las variables X4 y X5 vienen explicadas fundamentalmente por la componente 3. Aparecen ubicadas a izquierda y derecha distanciadas del origen) y próximas al eje descrito por la componente 3. Los otros dos subgrupos también aparecen distanciados del origen a lo largo de una sola componente.

Los gráficos siguientes muestran los puntos constituidos por las variables en el espacio de las componentes. Las coordenadas de cada variable vienen dadas en las filas de la tabla **Matriz de componentes rotada** (correlaciones de cada variable con cada componente).





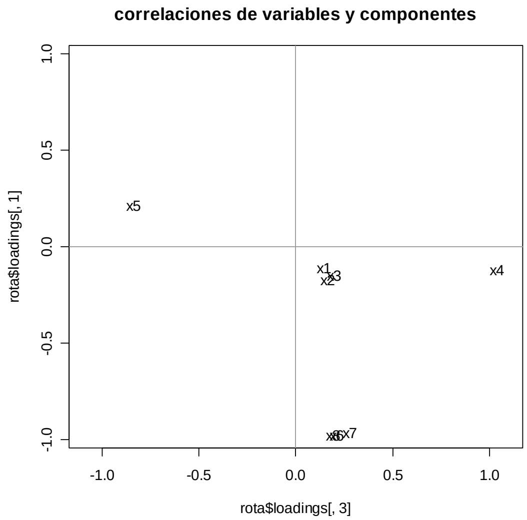


Tabla de coeficientes para el cálculo de las puntuaciones en las componentes:

El resultado es la matriz de puntuaciones estandarizadas en componentes rotadas, dada en la tabla siguiente:

```
> scale(predict(acp)[,1:3])%*% rota$rotmat
            [,1]        [,2]         [,3]
 [1,]  0.5799298  1.55678203  0.254135012
 [2,]  0.3801949  1.24091523  0.152200904
 [3,] -0.4772685  1.13103251 -0.677021626
 [4,] -0.8458057  1.04087572  0.615293779
 [5,]  0.5502799  0.13256201 -2.217276267
 [6,]  1.1031516 -0.08484720 -0.774328216
 [7,]  1.4409873 -0.05596679  2.019591662
 [8,] -1.4707896 -0.06515935  0.567006533
 [9,] -0.7361013 -0.31449016  0.482561350
[10,] -1.5784223 -0.70068861 -0.634128075
[11,]  0.7060876 -1.14962612 -0.388798860
[12,]  0.9562636 -1.36555246 -0.009903002
[13,] -0.6085071 -1.36583681  0.610666807
```

Se obtienen multiplicando la matriz Y_z de componentes sin rotar por T

$$Y_z^R = Y_z T$$