Análisis de Componentes Principales (Segunda Parte)

Mathias Bourel

IMERL - Facultad de Ingeniería, Universidad de la República, Uruguay

May 7, 2019

Introducción

La idea del *Análisis de Componentes Principales*, y del *Análisis Factorial* en general, consiste en considerar las filas o las columnas de la matriz de datos como nube de puntos de manera de simplificar el análisis de la misma sin modificar la estructura general queriendo ganar en simplicidad.

Usamos la palabra *factorial* en el sentido que la descomposición que se pretende obtener de la matriz de datos involucra combinaciones lineales de las variables (los factores).

Consideramos distintos tipos de Análisis Factorial:

- El *Análisis en Componentes Principales* (ACP) que trabaja con matrices de individuos con todas las variables continuas.
- El *Análisis de Correspondencias Simples* (ACS) que trabaja con tablas de contingencias entre dos variables.
- El *Análisis de Correspondencias Múltiples* (ACM), que generaliza lo anterior a más de dos variables.
- El *Análisis Discriminante* (AD).

Veremos que estos métodos se basan todos en el mismo principio: se buscará proyectar las nubes de puntos de filas y las nubes de puntos de columnas en nuevos espacios euclideos para simplificar y ganar en interpretabilidad y eventualmente tomar decisiones.

Introducción

A partir de la matriz de datos

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix}$$

- Los n puntos $\begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1p} \end{pmatrix}, \begin{pmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2p} \end{pmatrix}, \dots, \begin{pmatrix} x_{n1} \\ x_{n2} \\ \vdots \\ x_{np} \end{pmatrix}$ representan la nube de puntos de las filas en $\mathbb{R}^p$.

  Podemos querer ponderar los datos acordandoles un peso p_i y considerar un vector de probabilidad $\mathbf{p} = (p_1, \dots, p_n)$.

- Los p puntos $\begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix}, \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix}, \dots, \begin{pmatrix} x_{1p} \\ x_{2p} \\ \vdots \\ x_{np} \end{pmatrix}$ representan la nube de puntos de las columnas en $\mathbb{R}^n$.

Se trabajará por separado con estas nubes pero también se buscará relacionarlas.

Introducción

Recordamos esta transparencia importante:

- Si suponemos que $\overline{x_j} = \overline{x_k} = 0$ y que $s_j = s_k = 1$ (matriz centrada y reducida) entonces:

$$cos(x_j, x_k) = x_j' x_k = \sum_{i=1}^{n} (x_{ij} - \overline{x_j})(x_{ik} - \overline{x_k}) = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \overline{x_j})(x_{ik} - \overline{x_k})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \overline{x_j})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ik} - \overline{x_k})^2}} = r_{jk}$$

- Si suponemos que $\overline{x_j} = \overline{x_k} = 0$ se tiene que:

$$cos(x_j, x_k) = \frac{x_j' x_k}{||x_j'|| ||x_k||} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \overline{x_j})(x_{ik} - \overline{x_k})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \overline{x_j})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ik} - \overline{x_k})^2}} = r_{jk}$$

Le podemos dar pesos a las observaciones. Si $\mathbf{p}$ es el vector de probabilidades de la transparencia anterior, considero $P = diag(\mathbf{p})$. Si todos tienen el mismo peso $P = \frac{1}{n} I_n$ donde I_n es la identidad. Podemos definir las mismas nociones:

- El producto escalar entre dos variables centradas es su covarianza: $\langle x_j, x_k \rangle_P = x_j' P x_k = Cov(x_j, x_k)$ $\quad ||x_j||_P^2 = x_j' P x_j = Var(x_j)$
- El producto escalar entre dos variables centradas y reducidas es el coeficiente de correlación $\langle x_j, x_k \rangle_P = x_j' P x_k = r_{jk}$

Componentes principales

Volvamos a recordar el método de las componentes principales. Recordamos que:

- Cada eje factorial, de dirección a_k, es una combinación lineal de las variables originales. Inducen una nueva variable z_k que se llama *componente principal*. El valor en esta componente z_k del i-esimo individuo es

$$c_{ik} = \langle \mathbf{x_i}, a_k \rangle = \mathbf{x_i}' a_k = x_{i1} a_{k1} + x_{i2} a_{k2} + \cdots + x_{ip} a_{kp}$$

- Se construyen las componentes principales de manera a conservar la mayor parte de la información de la matriz de datos, deformando lo menos posible la información contenida en ella. Para eso, se trata de elegir los ejes que maximizan la proyección de la "inercia" de la nube de puntos.
- Las componentes principales son no correladas, es decir que los ejes son ortogonales.
- Trabajamos siempre con una matriz de datos centrada y se trata también siempre de reducir la matriz para darle la misma importancia a todas las variables.

Componentes principales

Se quiere tratar de condensar lo más posible la información de manera a retener variables que son realmente importantes y características. Para eso vamos a querer determinar un subespacio, llamado *subespacio factorial de la nube*, de dimensión q menor a p, y para ello necesitaremos de q nuevos ejes donde vamos a proyectar las nubes de puntos para que la información sea a la vez la más "visual" y fiel posible.
Tenemos entonces dos grandes etapas:

- Hacer un cambio de referencial entre el referencial definida por las variables viejas $x_1, \ldots, x_p$ y un referencial definida por p nuevos ejes, que pasan por el centro de gravedad de la nube de puntos, y que llamaremos *ejes factoriales*.
- nos quedamos con los q primeros ejes del nuevo sistema de coordenadas tratando de recuperar las relaciones más significativas de la matriz de datos.

Los ejes factoriales se determinan de la siguientes manera:

- El primer eje es el eje en el cual la nube de puntos se deforma los menos posible cuando lo proyectamos. Eso se traduce de manera a conseguir el eje para el cual la inercia proyectada es máxima.
- El segundo es ortogonal al primero, es el eje, después del primero sobre el cual la nube de puntos se deforme lo menos posible cuando la proyectamos.
- etc.

Inercia de una nube de puntos

- La inercia de un punto i al punto A se define como

$$I(i, A) = d^2(i, A) \times p_i$$

donde d es una distancia.

- La inercia de una nube de puntos N al punto A es

$$I(N, A) = \sum_{i=1}^{n} d^2(i, A) \times p_i$$

- La inercia de la nube de puntos N al centro de gravedad G es

$$I = \sum_{i=1}^{n} d^2(i, G) \times p_i = \sum_{i=1}^{n} p_i \sum_{j=1}^{p} (x_{ij} - \overline{x}_j)^2 = \sum_{j=1}^{p} Var(x_j)$$

I es un indicador de la cantidad de información, de la dispersión o de la forma de la nube respecto del centro de gravedad: más I será grande, más la nube será dispersa alrededor del centro de gravedad.

- Si $I = 0$ entonces todos los individuos son idénticos.
- Si la matriz es centrada el centro de gravedad de los individuos es el origen.
- Puesto que $\sum_{j=1}^{p} Var(x_j) = tr(\Sigma)$ si las variables son centradas y reducidas entonces $I = p$.
- Para obtener la inercia en R: sum(diag(cov(X))).

Componentes principales

El nuevo sistema de coordenadas pasa por el centro de gravedad G de la nube de puntos.

Queremos buscar el eje a_1 de manera que cuando proyectamos la nube de puntos sobre él la inercia sea máxima.

Sea $z_1 = (c_{11}, c_{21}, \ldots, c_{i1}, \ldots, c_{n1})$ el vector de coordenadas de la proyección ortogonal de los individuos sobre el eje a_1:

$$c_{i1} = \langle \mathbf{x}_i, a_1 \rangle \quad \text{y en definitiva} \quad z_1 = Xa_1$$

Si I_1 es la inercia de la proyección de la nube de puntos:

$$I_1 = \sum_{i=1}^{n} p_i d^2(\mathbf{x}_i, G) = \sum_{i=1}^{n} p_i c_{i1}^2 = z_1^{'} P z_1 = Var(z_1) = a_1^{'} X' P X a_1 = a_1^{'} S a_1$$

siendo $S = X'PX$ la matriz de varianzas y covarianzas de X con la matriz de pesos P.

Si X es centrada entonces $S = \Sigma$ y si X es centrada y reducida entonces $S = R$.

Maximizamos I_1 y por lo tanto $a_1^{'} S a_1$ sujeto a $a_1^{'} a_1 = 1$ cuya solución consiste en tomar a_1 un vector propio unitario asociado al mayor valor propio λ_1 de S.

Entonces:

- $I_1 = a_1^{'} S a_1 = a_1^{'} \lambda_1 a_1 = \lambda_1$
- El vector de coordenadas de los n puntos de la nube sobre el primer eje es $z_1 = Xa_1$.
- $\bar{z}_1 = \mathbf{0}$ y $Var(z_1) = z_1^{'} P z_1 = a_1^{'} S a_1 = \lambda_1$.

Componentes Principales

Recordamos que dado que maximizamos las sucesivas inercias, condicionada a que los vectores sobre los que vamos proyectar tienen norma 1, la solución consiste en obtener $a_1, a_2, \ldots, a_p$ una sucesión de vectores propios asociados a los valores propios $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$ de S.

De la misma manera que

- $I_1 = a_1^{'} S a_1 = a_1^{'} \lambda_1 a_1 = \lambda_1$
- El vector de coordenadas de los n puntos de la nube sobre el primer eje es $z_1 = X a_1$.
- $\bar{z}_1 = \mathbf{0}$ y $Var(z_1) = z_1^{'} P z_1 = a_1^{'} S a_1 = \lambda_1$.

Entonces:

- $I_k = a_k^{'} S a_k = a_k^{'} \lambda_k a_k = \lambda_k$
- El vector de coordenadas de los n puntos de la nube sobre el eje k-esimo es $z_k = X a_k$.
- $\bar{z}_k = \mathbf{0}$ y $Var(z_k) = z_k^{'} P z_k = a_k^{'} S a_k = \lambda_k$.
- $Cov(z_k, z_{k'}) = 0$ si $k \neq k'$.
- Cada eje factorial a_k (loading) representa una nueva variable z_k, de dimensión n, que se construye como combinación lineal de las variables originales. A z_k la llamamos *componente principal*. La coordenada (score) $c_{ik} = \langle \mathbf{x_i}, a_k \rangle = x_{i1} a_{k1} + x_{i2} a_{k2} + \cdots + x_{ip} a_{kp}$ de un individuo i sobre este eje corresponde al valor de la componente principal que toma este individuo.
- La inercia total es $I = \sum I_k = \sum \lambda_k$. Si la matriz de datos es centrada y reducida entonces $S = R$ y $\sum \lambda_k = p$ y por lo tanto la inercia es p.

Resumen

El análisis directo pasa por los siguientes pasos:

- Diagonalización de S (S es positiva definido de orden p, no tiene valor propio 0 y por lo tanto hay p vectores propios ortgonales).
- Ordenación de los valores propios en orden decreciente (todos ellos son menores o iguales a 1). Los vectores propios asociados determinan los ejes del nuevo referencial.

Cada eje factorial k, del vector de dirección a_k, representa una nueva variable z_k de dimensión n, construida como una combinación lineal de las variables (ejes) de salida, llamada componente principal. Estas nuevas variables creadas como combinación lineal de las variables de partida tienen varianza máxima. La coordenada c_{ik} de un individuo i dada en este eje corresponde al valor del componente principal tomado por este individuo.

Los componentes principales se construyen para restaurar la mayor parte de la información en la tabla y de manera a distorsionar la información lo menos posible. El primer componente principal será una combinación lineal de las variables de varianza máxima.

Los componentes principales no están correlacionados (los ejes son ortogonales).

Etapas de un Análisis de Componentes Principales

- **Preparación de la matriz de datos:** en general se trabaja con una matriz de datos centrada (siempre) y reducida (por lo general). Si la matriz de datos no se reduce, la importancia de las variables en el cálculo de los componentes principales depende de su orden de magnitud; una variable con desvío grande tendrá más peso que una variable con desvío bajo. Las variables con desvío grande contribuyen a la construcción de las primeros componentes principales: los cálculos no son falsos y conducen a la misma interpretación, pero la lectura de los resultados puede ser engorrosa.
- **Análisis directo:** construcción del espacio factorial de la nube de los individuos asociados a la matriz de datos. Mantenemos por el momento los ejes factoriales.
- **Análisis dual:** Construcción del espacio factorial de la nube de variables: se deduce del primero.
- **Interpretación de estos análisis:** elección del número de ejes a retener, construcción de las nubes de puntos proyectadas en estos ejes, interpretación de los ejes principales y estudio de las proximidades entre los puntos. Síntesis de los resultados y visualización de las nubes de puntos asociadas.

Análisis Dual

En el proceso anterior, encontramos los ejes factoriales trabajando con la nube de puntos de las filas (individuos). Veamos que pasa si trabajamos ahora con la nube de puntos de las columnas (variables).

Dado que queremos cambiar individuos por variables, es natural sustituir $X'PX \in \mathcal{M}_{p\times p}$ por $XX'P \in \mathcal{M}_{n\times n}$.

Veremos que:

- Los ejes factoriales que se encuentran en este caso son los mismos que los que se hallaron cuando consideramos la nube de puntos filas.
  Se prueba que los valores propios no nulos de esta matriz son los mismos que $X'PX$ ($= R$ en el caso centrado y reducida), y por lo tanto hay p ejes que retendremos.
- Se prueba además que la inercia calculada en este caso es la misma respectivamente para cada uno de los ejes, $I_k = \lambda_k$.

La clave pasa por un resultado de Algebra Lineal que consiste en mostrar que si $A \in \mathcal{M}_{n\times p}$ entonces $A'A \in \mathcal{M}_{p\times p}$ y $AA' \in \mathcal{M}_{n\times n}$ tienen los mismos valores propios no nulos.

Análisis Dual

La demostración es totalmente general, pero veamos como la aplicamos a nuestro caso. Si multiplicamos en ambos lados por X a

$$X'PXa_k = \lambda_k a_k$$

entonces

$$X(X'PX)a_k = X(\lambda_k a_k)$$

y como $z_k = Xa_k$ se tiene que

$$(XX'P)z_k = \lambda_k z_k$$

Entonces los valores propios de $XX'P$ son idénticos a los de $X'PX$, es decir $\lambda_k = \mu_k$ y los restantes valores propios son nulos.

- Recordar que $a_k \in \mathbb{R}^p$ pero $z_k \in \mathbb{R}^n$.
- Lo que probamos anteriormente es bastante notable, porque a pesar que $X'PX \in \mathcal{M}_{p\times p}$ y $XX'P \in \mathcal{M}_{n\times n}$ son matrices con tamaños distintos (¡una puede ser 2×2 y la otra 15×15!), tienen los mismos valores propios no nulos.
- Si bien los valores propios no nulos son los mismos, los vectores propios no lo son: los de $X'PX$ son de $\mathbb{R}^p$ y los de $XX'P$ de $\mathbb{R}^n$.

Análisis Dual

Entonces, y de la misma manera, buscando los ejes que maximicen la inercia proyectada de la nube de puntos de las variables, necesitamos una sucesión de vectores $P-$ unitarios $v \in \mathbb{R}^n$ que maximice

$$v'PXX'Pv$$

de manera que

$$v_k^{'}Pv_k = 1$$

y sean ortogonales dos a dos. Sabemos cual es la solución de este problema, la sucesión es una sucesión de n vectores propios de $XX'P$ y cumple con todo lo que tiene que cumplir pues dicha matriz es simétrica. Por lo tanto, consideramos los vectores propios unitarios $v_1, \ldots, v_p, v_{p+1}, \ldots, v_n \in \mathbb{R}^n$ asociados a los vaps $\lambda_1 \geq \cdots \geq \lambda_p > \lambda_{p+1} = \cdots = \lambda_n = 0$ de $XX'P$.

Como $(XX'P)v_k = \lambda_k v_k$ y $(XX'P)z_k = \lambda_k z_k$ para todo $k = 1, \ldots, p$ entonces z_k y los ejes v_k son ambos vectores propios asociados al mismo valor propio y son por lo tanto colineales. Podemos entonces tomar

$$v_k = \frac{z_k}{||z_k||_P}$$

donde $||z_k||_P^2 = z_k' P z_k = (Xa_k)'P(Xa_k) = \lambda_k$ y por lo tanto

$$v_k = \frac{z_k}{\sqrt{\lambda_k}}$$

Análisis Dual

Si llamamos $w_k = X'Pv_k = (d_{1k}, d_{2k}, \ldots, d_{jk}, \ldots, d_{pk}) \in \mathbb{R}^p$ para todo $k = 1, \ldots, p$ entonces w_k es un vector formado por las proyecciones de los puntos de la nube columnas (las p variables) sobre v_k.

Los vectores w_k son los equivalentes de los z_k pero para la nube de puntos columnas. Como $v_k = \frac{z_k}{\sqrt{\lambda_k}}$ entonces multiplicando por $X'P$ se tiene:

$$w_k = X'Pv_k = X'P\left(\frac{z_k}{\sqrt{\lambda_k}}\right) = \frac{X'PXa_k}{\sqrt{\lambda_k}} = \sqrt{\lambda_k}a_k$$

Consecuencias:

- $||w_k||^2 = \sum_{i=1}^{p} d_{ik}^2 = \lambda_k = I_k$
- Para cada $j = 1, \ldots, p$ se tiene que $d_{jk} = \sqrt{\lambda_k}a_{jk}$.
- Recordar que si la matriz está reducida vimos que $Cor(z_k, x_j) = \sqrt{\lambda_k}a_{jk}$ por lo tanto

$$d_{jk} = Cor(z_k, x_j) \qquad \text{lo cual implica que } -1 \leq d_{jk} \leq 1$$

- Como $a_k = \frac{w_k}{\sqrt{\lambda_k}} \Rightarrow Xa_k = X\frac{w_k}{\sqrt{\lambda_k}}$ y por lo tanto $z_k = \frac{1}{\sqrt{\lambda_k}}Xw_k$
- De la misma manera $v_k = \frac{z_k}{\sqrt{\lambda_k}} \Rightarrow (X'P)v_k = (X'P)\frac{z_k}{\sqrt{\lambda_k}}$ y por lo tanto

$$w_k = \frac{1}{\sqrt{\lambda_k}}(X'P)z_k$$

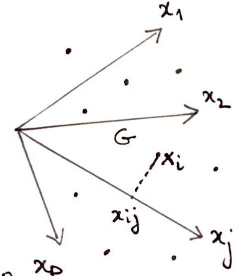
Análisis Dual

Siendo $z_k = (c_{1k}, c_{2k}, \ldots, c_{ik}, \ldots, c_{nk}) \in \mathbb{R}^n$ y $w_k = (d_{1k}, d_{2k}, \ldots, d_{jk}, \ldots, d_{pk}) \in \mathbb{R}^p$ entonces tenemos las fórmulas de pasaje:

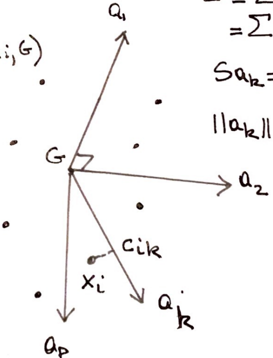
$$z_k = \frac{1}{\sqrt{\lambda_k}} X w_k \ \Rightarrow c_{ik} = \frac{1}{\sqrt{\lambda_k}} \sum_{j=1}^{p} x_{ij} d_{jk} \ \ \forall\, i = 1, \ldots, n$$

$$w_k = \frac{1}{\sqrt{\lambda_k}} (X'P) z_k \ \Rightarrow \ d_{jk} = \frac{1}{\sqrt{\lambda_k}} \sum_{i=1}^{n} x_{ij} p_i c_{ik} \ \ \forall\, i = j, \ldots, p$$

Análisis Directo.



$$I = \sum p_i d^2(x_i, G)$$



$$I = \sum I_k = \sum \lambda_k = \sum Var(z_k)$$

$$S a_k = \lambda_k a_k$$

$$\|a_k\| = 1.$$

individuo i

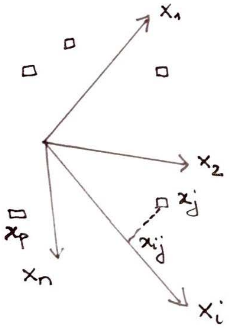
$$X_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$$

$$x_i = (c_{i1}, c_{i2}, \dots, c_{ip})'$$

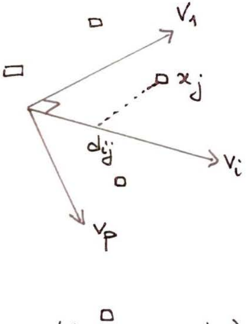
$$z_k = X a_k = (c_{1k}, c_{2k}, \dots, c_{nk})'$$

Resúmenes

Análisis Dual



$$x_j = (x_{1j}, x_{2j}, \dots, x_{ij}, x_{nj})'$$



$$x_j = (d_{1j}, d_{2j}, \dots, d_{pj})$$

$$v_k = \frac{X a_k}{\sqrt{\lambda_k}} \quad 1 \le k \le p$$

$$w_k = \sqrt{\lambda_k}\, a_k = (d_{1k}, d_{2k}, \dots, d_{pk})'$$

Ejemplo (con princomp)

```
> #Consumición anual en franco de 8 tipo de comida/bebida (variables) por 8 categorias socio-profesionales.
>
> #Variables: 1 Pan común, 2 Otro tipo de pan, 3 Vino común, 4 Otro tipo de vino,
> # 5 Papas, 6 Vegetales, 7 Uva, 8 Plato preparado
>
> #Individus 1 Productor rural, 2 Asalariado rural, 3 Profesional independiente,
> #4 Ejecutivo superior, 5 Ejecutivo medio, 6 Empleado, 7 Obrero, 8 Desocupado
>
>
> X=t(matrix(c(167,1,163,23,41,8,6,6,162,2,141,12,40,12,4,15,119,6,69,56,39,5,
+ 13,41,87,11,63,111,27,3,18,39,103,5,68,77,32,4,11,30,111,4,72,66,34,6,10,28,130,
+ 3,76,52,43,7,7,16,138,7,117,74,53,8,12,20),nrow=8) )
> colnames(X)=c("PC", "OP", "VC","OV","P","Veg","Uva", "Platos")
> rownames(X)=c("PRodRu", "Asalrur","Prof","Ejsup","Ejmoy","Emp","Obr", "Des")
> X
         PC OP  VC  OV  P Veg Uva Platos
PRodRu  167  1 163  23 41   8   6      6
Asalrur 162  2 141  12 40  12   4     15
Prof    119  6  69  56 39   5  13     41
Ejsup    87 11  63 111 27   3  18     39
Ejmoy   103  5  68  77 32   4  11     30
Emp     111  4  72  66 34   6  10     28
Obr     130  3  76  52 43   7   7     16
Des     138  7 117  74 53   8  12     20
>
> princomp(X,cor=T)
Call:
princomp(x = X, cor = T)

Standard deviations:
      Comp.1       Comp.2       Comp.3       Comp.4       Comp.5       Comp.6       Comp.7       Comp.8
2.491575e+00 9.379133e-01 6.449505e-01 5.535835e-01 4.104162e-01 1.344162e-01 5.870919e-02 2.257138e-08

 8  variables and  8 observations.
```

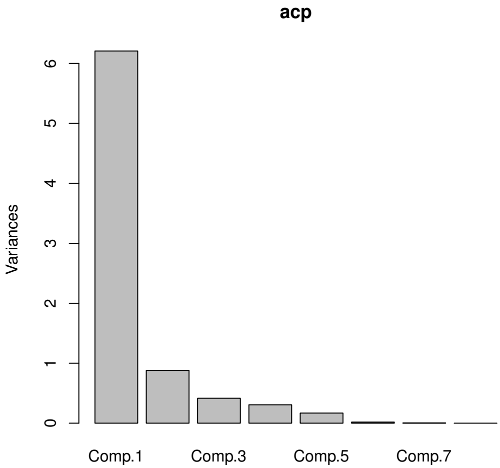
La salida es $\sqrt{\lambda_k}$ de cada una de las componentes

Ejemplo (con princomp)

```
> acp=princomp(X,cor=T)
> summary(acp)
Importance of components:
                          Comp.1    Comp.2     Comp.3     Comp.4     Comp.5      Comp.6       Comp.7       Comp.8
Standard deviation     2.4915752 0.9379133 0.64495048 0.55358348 0.41041625 0.134416177 0.0587091892 2.257138e-08
Proportion of Variance 0.7759934 0.1099602 0.05199514 0.03830683 0.02105519 0.002258464 0.0004308461 6.368342e-17
Cumulative Proportion  0.7759934 0.8859535 0.93794867 0.97625550 0.99731069 0.999569154 1.0000000000 1.000000e+00
```

- En el primer renglón aparece los $\sqrt{\lambda_k}$
- En el segundo renglón la proporción de varianza $\frac{I_k}{I}$
- En el tercer renglón esta proporción cumulada.
- $I = \sum \lambda_k = 8.$

Ejemplo (con princomp)



plot(*acp*)

Ejemplo (con princomp)

Las coordenadas de los vectores propios (son los a_k).

```
> loadings(acp)

Loadings:
       Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8
PC      0.391 -0.138  0.162 -0.119 -0.294  0.398 -0.107  0.729
OP     -0.349 -0.441  0.320 -0.218  0.265  0.521  0.423 -0.118
VC      0.349 -0.202  0.681        -0.246 -0.465  0.254 -0.180
OV     -0.374 -0.260         0.397  0.346 -0.423         0.575
P       0.246 -0.744 -0.558        -0.176 -0.108        -0.135
Veg     0.365 -0.128        -0.519  0.669 -0.185 -0.313
Uva    -0.373 -0.326  0.254        -0.272        -0.766 -0.159
Platos -0.362        -0.162 -0.708 -0.333 -0.360  0.225  0.219

               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
>
```

Estas coordenadas están expresadas en el referencial original, el de las viejas variables.

Ejemplo (con princomp)

Las coordenadas de los individuos sobre los ejes (los $c_{ik} = \mathbf{x_i}' a_k$):

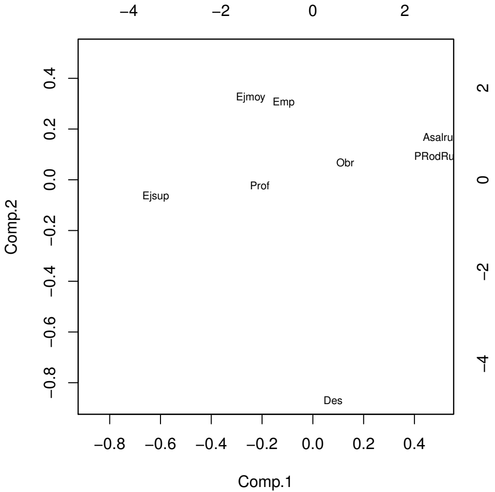
```
> acp$scores
           Comp.1      Comp.2     Comp.3      Comp.4      Comp.5       Comp.6       Comp.7        Comp.8
PRodRu   3.3715788  0.24581608  0.8395890  0.62172682 -0.57655700  0.021785204 -0.022311599  1.498801e-15
Asalrur  3.5217117  0.44739860  0.3515271 -0.91617942  0.49365924  0.004996441  0.031635398 -2.220446e-15
Prof    -1.4720309 -0.05851415 -0.5529570 -0.85448454 -0.74930243  0.058582704 -0.004644449 -1.609823e-15
Ejsup   -4.3587865 -0.17610682  1.0291875 -0.01517950  0.25877162  0.126514563 -0.015935125 -1.443290e-15
Ejmoy   -1.7180777  0.85664744 -0.1746349  0.41188554 -0.03988644 -0.139997633  0.116965074 -2.595146e-15
Emp     -0.8065346  0.80852679 -0.3448490  0.06912202  0.20594611 -0.195710003 -0.109502701  9.339751e-15
Obr      0.8991001  0.18303912 -0.9776683  0.55082419  0.29317809  0.233721990 -0.005887291 -3.441691e-15
Des      0.5630391 -2.30680707 -0.1701944  0.13228491  0.11419083 -0.109893267  0.009680694  4.024558e-16
```

Estas coordenadas están expresadas en el nuevo referencial, el de las componentes principales.

Ejemplo

Coordenadas de los individuos sobre los ejes factoriales:

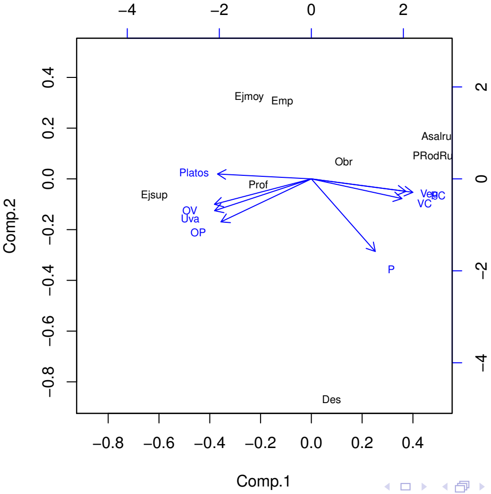
```
> biplot(acp, cex=0.7,col=c(1,0))
```



Eje

Coordenadas de los individuos sobre los ejes factoriales y proyección de las variables originales:

```
> biplot(acp, cex=0.7,col=c(1,4))
```



Ejemplo

Las coordenadas de las variables sobre los ejes factoriales es

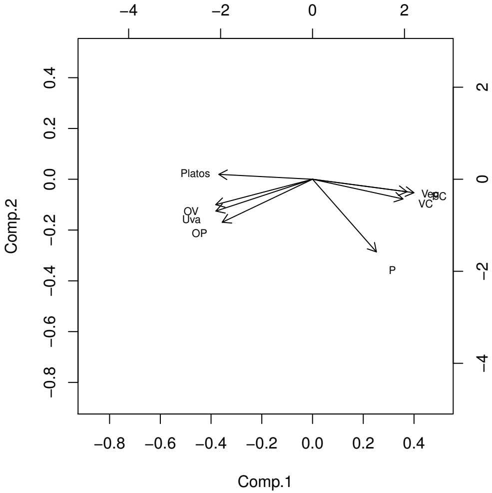
$$d_{jk} = r(x_j, z_k)$$

```
> cor(X,acp$scores)
           Comp.1      Comp.2      Comp.3      Comp.4     Comp.5       Comp.6       Comp.7       Comp.8
PC      0.9749797 -0.12926598  0.10429757 -0.06606998 -0.1206810  0.053463727 -0.006277188 -0.050156302
OP     -0.8687483 -0.41323074  0.20635173 -0.12063082  0.1089416  0.069991068  0.024838650 -0.123257316
VC      0.8700402 -0.18916036  0.43897378  0.01598936 -0.1008460 -0.062470185  0.014907633  0.015617014
OV     -0.9309151 -0.24414749  0.04739248  0.21952071  0.1418418 -0.056840057  0.001957682  0.032209797
P       0.6138529 -0.69764474 -0.35966296  0.04096049 -0.0721205 -0.014482890  0.005485109 -0.082240510
Veg     0.9089814 -0.12007291  0.02089707 -0.28724855  0.2746472 -0.024859138 -0.018382280 -0.022042647
Uva    -0.9294859 -0.30574089  0.16397854 -0.03526677 -0.1114413  0.002186234 -0.044965573 -0.003946343
Platos -0.9011429  0.04710881 -0.10428318 -0.39199413 -0.1366334 -0.048422745  0.013207567 -0.023521898
```

Ejemplo

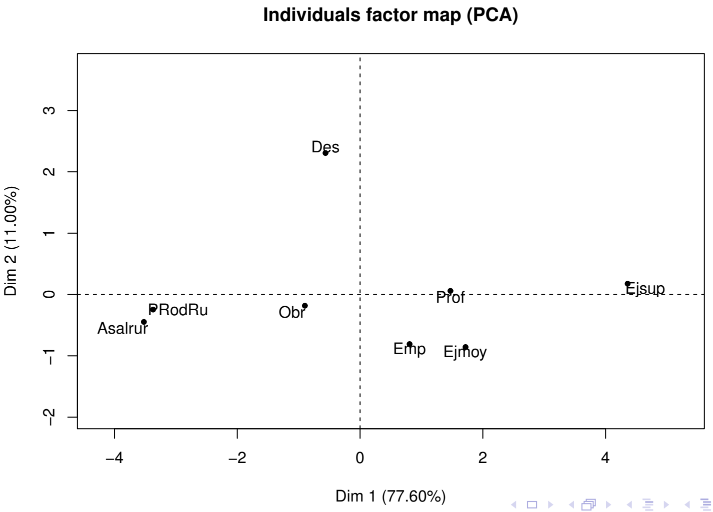
```
>  biplot(acp, cex=0.7,col=c(0,1))
```

(con $c(0,1)$ activo las variables y escondo los individuos)



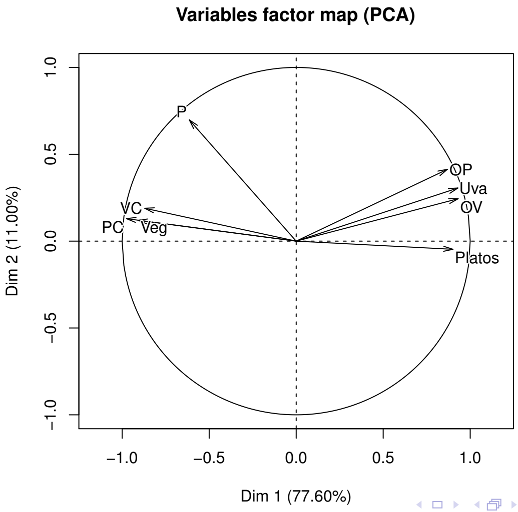
Mismo ejemplo con FactoMineR

```
> library(FactoMineR)
> acp1=PCA(X)
```



Mismo ejemplo con FactoMineR

```
> library(FactoMineR)
> acp1=PCA(X)
```



Mismo ejemplo con FactoMineR

Valores propios y varianza explicada por los ejes.

```
> acp1$eig
       eigenvalue percentage of variance cumulative percentage of variance
comp 1 6.207946839            77.59933549                          77.59934
comp 2 0.879681393            10.99601741                          88.59535
comp 3 0.415961123             5.19951404                          93.79487
comp 4 0.306454670             3.83068337                          97.62555
comp 5 0.168441497             2.10551872                          99.73107
comp 6 0.018067709             0.22584636                          99.95692
comp 7 0.003446769             0.04308461                         100.00000
```

observar que son bien los cuadrados de los valores que se obtienen con la función princomp.

Las coordenadas de los individuos sobre los ejes (los $c_{ik} = \mathbf{x_i}' a_k$):

```
> acp1$ind$coord
             Dim.1       Dim.2      Dim.3       Dim.4       Dim.5
PRodRu  -3.3715788 -0.24581608  0.8395890 -0.62172682  0.57655700
Asalrur -3.5217117 -0.44739860  0.3515271  0.91617942 -0.49365924
Prof     1.4720309  0.05851415 -0.5529570  0.85448454  0.74930243
Ejsup    4.3587865  0.17610682  1.0291875  0.01517950 -0.25877162
Ejmoy    1.7180777 -0.85664744 -0.1746349 -0.41188554  0.03988644
Emp      0.8065346 -0.80852679 -0.3448490 -0.06912202 -0.20594611
Obr     -0.8991001 -0.18303912 -0.9776683 -0.55082419 -0.29317809
Des     -0.5630391  2.30680707 -0.1701944 -0.13228491 -0.11419083
```

Mismo ejemplo con FactoMineR

Coordenadas de las variables sobre los ejes factoriales (los d_{jk}):

```
> acp1$var$coord
            Dim.1       Dim.2       Dim.3       Dim.4      Dim.5
PC     -0.9749797  0.12926598  0.10429757  0.06606998  0.1206810
OP      0.8687483  0.41323074  0.20635173  0.12063082 -0.1089416
VC     -0.8700402  0.18916036  0.43897378 -0.01598936  0.1008460
OV      0.9309151  0.24414749  0.04739248 -0.21952071 -0.1418418
P      -0.6138529  0.69764474 -0.35966296 -0.04096049  0.0721205
Veg    -0.9089814  0.12007291  0.02089707  0.28724855 -0.2746472
Uva     0.9294859  0.30574089  0.16397854  0.03526677  0.1114413
Platos  0.9011429 -0.04710881 -0.10428318  0.39199413  0.1366334
```