Tema 6. Reducción de la dimensión.

6.1. Análisis de componentes principales y análisis factorial.

6.2. Análisis de correspondencias

6.1. Análisis de componentes principales y análisis factorial.

6.1. 1. EL MODELO FACTORIAL (Y EL MODELO DE LAS COMPONENTES PRINCIPALES)
Básicamente pretendemos poder explicar n variables X_j ($j=1,2,\dots,n$) en función de un número menor m de factores inobservables comunes y n factores específicos (uno por cada variable). La pretensión de reducción nos obliga a que $m < n$.¹

Sin perder generalidad, consideraremos como variables iniciales a las correspondientes variables tipificadas (a media 0 y desviación típica 1), Z_j ($j=1,2,\dots,n$). De esta forma todas sus medias serán cero, todas sus varianzas unos y la matriz de varianzas coincidirá con la matriz de correlación².

Así pues, nuestro objetivo es representar cada una de las Z_j como una función de m factores comunes y un factor único para cada una de ellas. La forma de esta función se presume lineal, de manera que el modelo factorial obedece al siguiente esquema :

$$Z_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jm}F_m + u_j Y_j \quad (\text{con } j= 1,2,\dots,n)$$

Esquema que, en realidad, se traduce en un sistema de n ecuaciones; una para cada variable.

En este esquema las Z_j son las variables originales tipificadas. Las F_p son los factores comunes (variables que **no son directamente observables**), donde p recorre los valores $1,2,\dots,m$, con $m < n$. Las Y son factores únicos o específicos, uno para cada variable, que explican la varianza de la variable j -sima que no resulta explicada por los factores comunes. Cada una de las a_{jp} es cada uno de los coeficientes que afectan al factor p -simo para la explicación de la j -sima variable, a estos coeficientes se les suele llamar "cargas", "pesos" o "saturaciones" de la variable j en el factor p .³

El problema básico del análisis factorial es, precisamente, la estimación de las $n \times m$ cargas (a_{jp}) de los factores comunes, cosa que puede hacerse utilizando distintos métodos. Una vez estimadas las cargas, y dependiendo del método empleado, se podrá proceder a la estimación o determinación de los factores.

Si representamos el esquema del modelo factorial, con todo el sistema de n ecuaciones, éstas pueden muy bien simbolizarse matricialmente del modo siguiente:

$$\mathbf{Z} = \mathbf{A} \mathbf{F} + \mathbf{Y} \mathbf{U}$$

Donde:

$\mathbf{Z}$ es el vector columna formado por las variables originales tipificadas.

$\mathbf{A}$ es la matriz $n \times m$ de las cargas a_{jp} .

¹ (*1) En lo sucesivo emplearemos el subíndice j para denotar genéricamente el ordinal de la variable (j -sima); el subíndice i , para el ordinal del individuo (i -simo); y el subíndice p , para el ordinal del factor común (p -simo). Para referirnos a una observación concreta el orden de los subíndices será "variable- individuo" (o "factor-individuo, su caso); así, por ejemplo, X_{12} es el valor que toma la variable X_1 para el segundo individuo.

² Entre variables originales (X) o entre tipificadas (Z), ya que las correlaciones son independientes de origen y unidades.

³ Suele considerarse que los factores únicos están incorrelacionados entre sí y están incorrelacionados con los factores comunes.

F es el vector columna formado por los m factores comunes.

Y es una matriz n×n que dispondrá de los factores únicos Y_j en sus elementos diagonales y ceros en el resto.

Y, por último:

U será el vector columna formado por los respectivos coeficientes u_j de los factores específicos.

$$\begin{pmatrix} Z_1 \\ Z_2 \\ \dots \\ Z_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & a_{jm} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ \dots \\ F_m \end{pmatrix} + \begin{pmatrix} Y_1 & 0 & 0 & \dots & 0 \\ 0 & Y_2 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & Y_n \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}$$

COMPONENTES PRINCIPALES.

Un posible método de solución para el análisis factorial se basa en un modelo similar, aunque algo diferente, que constituye el esquema de relación básico del análisis de **componentes principales**.

Esta técnica puede considerarse un método particular de análisis factorial, o bien como un análisis diferenciado. Este método fue desarrollado por Pearson y Hotelling y se basa en un resultado fundamental del álgebra conocido como Teorema Espectral. Basándose en este teorema es siempre posible hacer una transformación lineal sobre el espacio vectorial de las variables de forma que las n variables iniciales puedan considerarse como combinaciones lineales de otras n nuevas variables (factores) o componentes principales que resultan estar incorrelacionadas.

Así pues el esquema del análisis de componentes principales es:

$$Z_j = a_{j1}P_1 + a_{j2}P_2 + \dots + a_{jn}P_n$$

(Con j = 1, 2, ..., n)

o matricialmente : **Z = A P**

Donde cada una de las n variables se describe linealmente en función de las n componentes incorrelacionadas P₁, P₂, ..., P_n.

Una propiedad importante de este método es que cada sucesiva componente hace máxima su contribución a la suma las varianzas de las n variables. Para un problema práctico, bastará quedarse con un pequeño número de componentes que expliquen una proporción suficiente de la varianza inicial total. Este pequeño número de componentes se corresponderían con los factores comunes del modelo factorial general. La parte no explicada por estos factores se correspondería con la parte no explicada por los factores comunes, explicable por factores únicos o específicos.

De esta manera puede relacionarse el modelo de componentes principales y el factorial y puede considerarse método de componentes principales como un sistema particular de solución factorial

Debemos puntualizar, antes de pasar a analizar las primeras consecuencias del modelo, que el modelo de análisis factorial es muy similar al de la regresión en lo que se refiere a que una variable se expresa como una combinación lineal de otro conjunto de variables más un término residual. Sin embargo, mientras en el caso de la regresión las variables independientes son observables, en el análisis factorial son *entes hipotéticos* que únicamente pueden *estimarse* a partir de los datos observados.

1.2.DESCOMPOSICION DE LA VARIANZA

La varianza de cada variable Z_j, S_j², puede expresarse en función de los factores teniendo en cuenta la relación del esquema del modelo factorial

$$S_j^2 = 1 = a_{j1}^2 S_{F_1}^2 + a_{j2}^2 S_{F_2}^2 + \dots + a_{jm}^2 S_{F_m}^2 + \sum_{p=1}^m \sum_{k \neq i} a_p a_k S_{F_p F_k} + u_j^2 S_{Y_j}^2 + u_j \sum_{p=1}^m a_{jp} S_{F_p Y_j}$$

Considerando que los factores (comunes y únicos) también están tipificados, sus varianzas serán unitarias y sus covarianzas coincidirán con los coeficientes de correlación:

$$S_j^2 = 1 = a_{j1}^2 + a_{j2}^2 + \dots + a_{jm}^2 + \sum_{p=1}^m \sum_{k \neq p} a_p a_k r_{F_p F_k} + u_j^2 + u_j \sum_{p=1}^m a_{jp} r_{F_p Y_j}$$

Los factores específicos suelen estar, por hipótesis, incorrelacionados con los comunes, de forma que todos los $r_{F_p Y_j} = 0$ y, por tanto, el último sumando de la expresión anterior se anula.

Por otro lado si los factores comunes están incorrelacionados entre sí, cosa que ocurre, por ejemplo, cuando se aplica la técnica de componentes principales y algunas otras, la expresión anterior se reduce a :

$$S_j^2 = 1 = a_{j1}^2 + a_{j2}^2 + \dots + a_{jm}^2 + u_j^2$$

Donde cada uno de los términos del segundo miembro representa las proporciones de la varianza unitaria de Z_j adjudicada o explicada por cada factor.

De la descomposición de la varianza unitaria de cada variable se derivan dos conceptos fundamentales del análisis factorial : la **comunalidad** y la **unicidad**.

La suma de los m primeros términos de la expresión anterior se conoce con el nombre de comunalidad de la j-sima variable original, se representa por h_j^2 y es la parte (proporción) de la varianza de Z_j que resulta explicada por los factores comunes.

Como $S_j^2 = 1$, Comunalidad + Unicidad = 1 $S_j^2 = 1 = h_j^2 + u_j^2$

Si al realizar un análisis factorial todas las comunalidades de todas las variables fueran altas, próximas a uno, todas las variables resultarían bastante bien explicadas por los factores comunes. En consecuencia, cuando más comunalidades sean más altas mejor será el análisis.

La unicidad es el cuadrado del coeficiente del factor único, u_j^2 , y viene a expresar la proporción de la varianza que queda explicada por el factor único, o, dicho de otro modo, la que no puede explicarse por los factores comunes.

En ocasiones es conveniente separar la unicidad en dos partes: una que dé cuenta del resto inexplicado que resulta debido a la selección particular de las variables en estudio y otra debida a la falta de fiabilidad de las medidas. Se llaman respectivamente especificidad b^2 , y varianza de error o falta de fiabilidad, e^2 .

Por otra parte, la contribución total del factor F_p a la varianza total (VarianzaTotal= $n \cdot \sum S_j^2$) será:

$$V_p = \sum_{j=1}^n a_{jp}^2$$

Y la contribución total de todos los factores comunes a la varianza total de todas las variables vendrá

dada por: $V = \sum_{p=1}^m V_p = \sum_{p=1}^m \sum_{j=1}^n a_{jp}^2 = \sum_{j=1}^n h_j^2$ Es decir, la suma de todas las comunalidades.

La fracción V/n se emplea, a veces, como indicador de la completitud del análisis factorial. Si el análisis fuera absolutamente perfecto este factor valdría 1.

$$\left. \begin{aligned}
\frac{1}{N} \sum_{i=1}^N Z_{1i} F_{1i} &= a_{11} \frac{1}{N} \sum_{i=1}^N F_{1i}^2 + a_{12} \frac{1}{N} \sum_{i=1}^N F_{1i} F_{2i} + \dots + a_{1p} \frac{1}{N} \sum_{i=1}^N F_{1i} F_{pi} + \dots + a_{1m} \frac{1}{N} \sum_{i=1}^N F_{1i} F_{mi} + u_1 \frac{1}{N} \sum_{i=1}^N F_{1i} Y_{1i} \\
\frac{1}{N} \sum_{i=1}^N Z_{1i} F_{2i} &= a_{11} \frac{1}{N} \sum_{i=1}^N F_{2i} F_{1i} + a_{12} \frac{1}{N} \sum_{i=1}^N F_{2i}^2 + \dots + a_{1p} \frac{1}{N} \sum_{i=1}^N F_{2i} F_{pi} + \dots + a_{1m} \frac{1}{N} \sum_{i=1}^N F_{2i} F_{mi} + u_1 \frac{1}{N} \sum_{i=1}^N F_{2i} Y_{1i} \\
&\dots \\
\frac{1}{N} \sum_{i=1}^N Z_{1i} F_{pi} &= a_{11} \frac{1}{N} \sum_{i=1}^N F_{pi} F_{1i} + a_{12} \frac{1}{N} \sum_{i=1}^N F_{pi} F_{2i} + \dots + a_{1p} \frac{1}{N} \sum_{i=1}^N F_{pi}^2 + \dots + a_{1m} \frac{1}{N} \sum_{i=1}^N F_{pi} F_{mi} + u_1 \frac{1}{N} \sum_{i=1}^N F_{pi} Y_{1i} \\
&\dots \\
\frac{1}{N} \sum_{i=1}^N Z_{1i} F_{mi} &= a_{11} \frac{1}{N} \sum_{i=1}^N F_{mi} F_{1i} + a_{12} \frac{1}{N} \sum_{i=1}^N F_{mi} F_{2i} + \dots + a_{1p} \frac{1}{N} \sum_{i=1}^N F_{mi} F_{pi} + \dots + a_{1m} \frac{1}{N} \sum_{i=1}^N F_{mi}^2 + u_1 \frac{1}{N} \sum_{i=1}^N F_{mi} Y_{1i} \\
&\dots \\
&\dots \\
\frac{1}{N} \sum_{i=1}^N Z_{ji} F_{1i} &= a_{j1} \frac{1}{N} \sum_{i=1}^N F_{1i}^2 + a_{j2} \frac{1}{N} \sum_{i=1}^N F_{1i} F_{2i} + \dots + a_{jp} \frac{1}{N} \sum_{i=1}^N F_{1i} F_{pi} + \dots + a_{jm} \frac{1}{N} \sum_{i=1}^N F_{1i} F_{mi} + u_j \frac{1}{N} \sum_{i=1}^N F_{1i} Y_{ji} \\
\frac{1}{N} \sum_{i=1}^N Z_{ji} F_{2i} &= a_{j1} \frac{1}{N} \sum_{i=1}^N F_{2i} F_{1i} + a_{j2} \frac{1}{N} \sum_{i=1}^N F_{2i}^2 + \dots + a_{jp} \frac{1}{N} \sum_{i=1}^N F_{2i} F_{pi} + \dots + a_{jm} \frac{1}{N} \sum_{i=1}^N F_{2i} F_{mi} + u_j \frac{1}{N} \sum_{i=1}^N F_{2i} Y_{ji} \\
&\dots \\
\frac{1}{N} \sum_{i=1}^N Z_{ji} F_{pi} &= a_{j1} \frac{1}{N} \sum_{i=1}^N F_{pi} F_{1i} + a_{j2} \frac{1}{N} \sum_{i=1}^N F_{pi} F_{2i} + \dots + a_{jp} \frac{1}{N} \sum_{i=1}^N F_{pi}^2 + \dots + a_{jm} \frac{1}{N} \sum_{i=1}^N F_{pi} F_{mi} + u_j \frac{1}{N} \sum_{i=1}^N F_{pi} Y_{ji} \\
&\dots \\
\frac{1}{N} \sum_{i=1}^N Z_{ji} F_{mi} &= a_{j1} \frac{1}{N} \sum_{i=1}^N F_{mi} F_{1i} + a_{j2} \frac{1}{N} \sum_{i=1}^N F_{mi} F_{2i} + \dots + a_{jp} \frac{1}{N} \sum_{i=1}^N F_{mi} F_{pi} + \dots + a_{jm} \frac{1}{N} \sum_{i=1}^N F_{mi}^2 + u_j \frac{1}{N} \sum_{i=1}^N F_{mi} Y_{ji} \\
&\dots \\
&\dots \\
\frac{1}{N} \sum_{i=1}^N Z_{ni} F_{1i} &= a_{n1} \frac{1}{N} \sum_{i=1}^N F_{1i}^2 + a_{n2} \frac{1}{N} \sum_{i=1}^N F_{1i} F_{2i} + \dots + a_{np} \frac{1}{N} \sum_{i=1}^N F_{1i} F_{pi} + \dots + a_{nm} \frac{1}{N} \sum_{i=1}^N F_{1i} F_{mi} + u_n \frac{1}{N} \sum_{i=1}^N F_{1i} Y_{ni} \\
\frac{1}{N} \sum_{i=1}^N Z_{ni} F_{2i} &= a_{n1} \frac{1}{N} \sum_{i=1}^N F_{2i} F_{1i} + a_{n2} \frac{1}{N} \sum_{i=1}^N F_{2i}^2 + \dots + a_{np} \frac{1}{N} \sum_{i=1}^N F_{2i} F_{pi} + \dots + a_{nm} \frac{1}{N} \sum_{i=1}^N F_{2i} F_{mi} + u_n \frac{1}{N} \sum_{i=1}^N F_{2i} Y_{ni} \\
&\dots \\
\frac{1}{N} \sum_{i=1}^N Z_{ni} F_{pi} &= a_{n1} \frac{1}{N} \sum_{i=1}^N F_{pi} F_{1i} + a_{n2} \frac{1}{N} \sum_{i=1}^N F_{pi} F_{2i} + \dots + a_{np} \frac{1}{N} \sum_{i=1}^N F_{pi}^2 + \dots + a_{nm} \frac{1}{N} \sum_{i=1}^N F_{pi} F_{mi} + u_n \frac{1}{N} \sum_{i=1}^N F_{pi} Y_{ni} \\
&\dots \\
\frac{1}{N} \sum_{i=1}^N Z_{ni} F_{mi} &= a_{n1} \frac{1}{N} \sum_{i=1}^N F_{mi} F_{1i} + a_{n2} \frac{1}{N} \sum_{i=1}^N F_{mi} F_{2i} + \dots + a_{np} \frac{1}{N} \sum_{i=1}^N F_{mi} F_{pi} + \dots + a_{nm} \frac{1}{N} \sum_{i=1}^N F_{mi}^2 + u_n \frac{1}{N} \sum_{i=1}^N F_{mi} Y_{ni}
\end{aligned} \right\}$$

Desarrollando estas expresiones, estos coeficientes de correlación y prescindiendo de los factores únicos incorrelacionados con los comunes, por hipótesis tendremos que para cada variable original Z_j (con $j=1,2,\dots,n$), tendremos que :

$$\left. \begin{array}{l} r_{Z_j F_1} = a_{j1} + a_{j2} r_{F_1 F_2} + \dots + a_{jp} r_{F_1 F_p} + \dots + a_{jm} r_{F_1 F_m} \\ \dots \\ r_{Z_j F_p} = a_{j1} r_{F_p F_1} + a_{j2} r_{F_p F_2} + \dots + a_{jp} + \dots + a_{jm} r_{F_p F_m} \\ \dots \\ r_{Z_j F_m} = a_{j1} r_{F_m F_1} + a_{j2} r_{F_m F_2} + \dots + a_{jp} r_{F_m F_p} + \dots + a_{jm} \end{array} \right\} \text{ para } j=1,2,\dots,n$$

Aunque puede parecer que las ecuaciones anteriores se utilizan para calcular los elementos de la estructura, lo cierto es que, habitualmente, su utilidad es la contraria: la determinación de los coeficientes del patrón (a_{jp}), a partir del conocimiento de las correlaciones entre factores y entre las variables y factores.

Por lo general, los coeficientes del patrón y los de la estructura serán diferentes. Sin embargo, en el caso en que los factores comunes estén incorrelacionados entre sí, se deduce fácilmente que los elementos de la estructura serán idénticos a los coeficientes del patrón:

$$\text{si todos los factores están incorrelacionados} \Rightarrow r_{Z_j F_p} = a_{jp} \quad \forall j, p$$

No es lo habitual, pero es lo que ocurre cuando se aplica el método de Componentes Principales.

En general, por lo tanto, será necesario para obtener la solución del análisis factorial determinar tanto la estructura como el patrón factorial:

" La estructura revela las correlaciones entre las variables y los factores comunes -- las variables originales con las que se supone estamos familiarizados y los entes hipotéticos (factores) que deseamos comprender-- y por tanto, son útiles para la identificación o reconocimiento de los factores. Los valores de la estructura se necesitan también para estimaciones posteriores de las puntuaciones factoriales. El patrón muestra la composición lineal de las variables en términos de los factores en la forma de ecuaciones de regresión. También pueden utilizarse para reproducir las correlaciones entre las variables para determinar la bondad de las soluciones. Los patrones son también útiles a la hora de comparar diferentes sistemas de factores para un conjunto dado de variables.⁴

EL ANÁLISIS DE COMPONENTES PRINCIPALES COMO MÉTODO DE ANÁLISIS FACTORIAL.

El análisis de componentes principales es, en principio, un método destinado a poder representar un conjunto de n variables en función de otro conjunto de n nuevas variables (componentes principales) que están todas ellas incorrelacionadas entre sí.

El método requiere, en esencia, la rotación de los ejes de coordenadas a un nuevo sistema de referencia en el espacio *total* de las variables. Esto es, una transformación ortogonal en la que cada una de las n variables originales se puede describir en términos de n nuevas componentes principales.

Si partimos de las variables originales tipificadas, el esquema del modelo factorial nos viene dado por las expresiones, vistas ya:

$$Z_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jm}F_m + u_j Y_j \quad (\text{con } j= 1,2,\dots,n)$$

O bien $Z = A F + Y U$

Pero ahora consideraríamos que el número de factores comunes (en principio) sería de n y no existirían los factores únicos

$$Z = A P$$

⁴ Harman, H.H: "Análisis Factorial moderno"

La posibilidad de expresar las variables originales en función de factores (componentes) incorrelacionados se basa en un resultado del álgebra conocido como Teorema Espectral. Según este teorema, cualquier forma cuadrática definida positiva puede transformarse en otra forma cuadrática definida positiva cuya matriz asociada sólo tenga elementos en su diagonal principal, si se lleva a cabo una determinada transformación del sistema de referencia; esto es, un determinado cambio de base del espacio vectorial.

Este resultado tiene como consecuencia que una matriz definida positiva puede transformarse, en otro sistema de referencia, en una matriz diagonal (también definida positiva).

Si se aplica este resultado a la matriz de varianzas (o de correlación) de las variables originales, se obtiene que *es posible encontrar una transformación lineal, que nos exprese estas variables en términos de otras (nueva base) variables incorrelacionadas.*

En nuestro caso partamos de las variables originales tipificadas Z, que pueden expresarse en forma vectorial como:

$$Z = \begin{pmatrix} Z_1 \\ Z_2 \\ \dots \\ Z_n \end{pmatrix}$$

Y su matriz de varianzas , que al estar las variables tipificadas, coincidirá con la matriz de correlación , como matriz definida positiva a diagonalizar.

La matriz de varianzas de Z coincide con la matriz de correlación, R :

$$R = \begin{pmatrix} 1 & r_{12} & r_{13} & \dots & r_{1n} \\ r_{21} & 1 & r_{23} & \dots & r_{2n} \\ \dots & \dots & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & \dots & 1 \end{pmatrix}$$

Sigamos, a la vez un sencillo ejemplo (**EJEM1**) para ilustrar el método: Supongamos que disponemos de información sobre el peso y la estatura de 20 individuos. El análisis factorial no está pensado precisamente para aplicarlo a conjuntos de datos de tan pocas variables, pero como ilustración nos será útil:

individuo	Estatura (cm)	Peso Kg
1	170	65
2	173	65
3	189	78
4	190	82
5	175	89
6	168	78
7	174	70
8	179	73
9	185	79
10	193	90
11	180	78

12	179	75
13	182	95
14	172	85
15	173	69
16	179	76
17	182	80
18	197	85
19	183	80
20	182	78

Tipificando , peso y altura obtendríamos las variables originales ,que hemos llamado Z:

Z1	Z2
-1,47753578	-1,9421646
-1,08255096	-1,9421646
1,0240347	-0,25895528
1,1556963	0,25895528
-0,81922776	1,16529876
-1,74085898	-0,25895528
-0,95088936	-1,2947764
-0,29258134	-0,90634348
0,49738828	-0,12947764
1,55068111	1,2947764
-0,16091974	-0,25895528
-0,29258134	-0,6473882
0,10240347	1,9421646
-1,21421257	0,6473882
-1,08255096	-1,42425404
-0,29258134	-0,51791056
0,10240347	0
2,07732753	0,6473882
0,23406507	0
0,10240347	-0,25895528

Y obteniendo , previamente ,medias, varianzas y covarianzas podríamos obtener la matriz de correlación R que en este caso sería una matriz 2×2:

$$R = \begin{pmatrix} 1 & 0.5050243 \\ 0.5050243 & 1 \end{pmatrix}$$

Volvamos al desarrollo:

Aplicando el teorema espectral, podremos encontrar una transformación lineal de matriz **T** tal que :

$$\mathbf{Z} = \mathbf{T} \mathbf{P} \quad \text{o bien} \quad \mathbf{P} = \mathbf{T}' \mathbf{Z}^5$$

y que $\mathbf{R} = \mathbf{T} \mathbf{D} \mathbf{T}'$

de forma que la forma cuadrática $\mathbf{Q}(\mathbf{Z}) = \mathbf{Z}' \mathbf{R} \mathbf{Z}$ se transforme en el nuevo espacio de referencia en : $\mathbf{Q}'(\mathbf{P}) = \mathbf{P}' \mathbf{D} \mathbf{P}$ siendo **D** una matriz diagonal.⁶

⁵ Por ser T la matriz de cambio de base ser una matriz ortogonal que verificar que $T = T'$

⁶ Las formas cuadráticas Q(Z) y Q'(P) no son más que las matrices de correlación de las Z y las P (componentes principales) , respectivamente

El teorema determina también que la matriz diagonal obtenida tiene la forma :

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & \dots & \lambda_j & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

donde los λ_j , todos ellos positivos, son los **valores propios** (autovalores o eigenvalues) de la matriz definida original, en nuestro caso R. Esto es, las n raíces de la *ecuación característica*, en nuestro caso:

$$\det(R - \lambda I) = 0$$

Para el ejemplo (**EJEM1**) tendremos:

$$\det(R - \lambda I) = 0 \Rightarrow \det \begin{pmatrix} 1 - \lambda & 0.5050243 \\ 0.5050243 & 1 - \lambda \end{pmatrix} = 0 \Rightarrow \lambda^2 - 2\lambda + 0.74495046 = 0$$

Las soluciones de esta ecuación de segundo grado son $\lambda_1 = 1.505$ y $\lambda_2 = 0.495$

Las columnas de la matriz de transformación **T**, son los vectores propios (autovectores o eigenvectors) de la matriz original, en nuestro caso, R, y se obtienen de la solución de los sistemas de ecuaciones:

$$(R - \lambda_j I)X = 0 \quad (\text{para } j=1,2,\dots,n)$$

En el ejemplo:

$$\begin{pmatrix} 1 - 1.505 & 0.5050243 \\ 0.5050243 & 1 - 1.505 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.867 \\ 0.867 \end{pmatrix}$$

$$\begin{pmatrix} 1 - 0.495 & 0.5050243 \\ 0.5050243 & 1 - 0.495 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.497 \\ -0.497 \end{pmatrix}$$

Y donde X es un vector columna n-dimensional de incógnitas a determinar (las componentes de la j-sima columna de T).

En el ejemplo:
$$T = \begin{pmatrix} 0.867 & 0.497 \\ 0.867 & -0.497 \end{pmatrix}$$

Así pues, aplicando el Teorema Espectral podemos obtener que las variables originales (tipificadas), representadas por el vector **Z** pueden expresarse como una combinación lineal de otras n variables $P_1, P_2, \dots$, representadas por el vector :

$$P = \begin{pmatrix} P_1 \\ P_2 \\ \dots \\ P_n \end{pmatrix}$$
 mediante la expresión $Z = TP$, de forma que las P están incorrelacionadas; de manera

que la transformación, **T**, nos venga dada (por columnas) por la matriz formada por los vectores propios de **R**, e igualmente que la matriz de varianzas de las variables P nos venga dada por:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & \dots & \lambda_j & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

En el ejemplo: $D = \begin{pmatrix} 1.505 & 0 \\ 0 & 0.495 \end{pmatrix}$

Si pretendemos que la obtención de las componentes principales nos conduzca a factores, no sólo incorrelacionados, sino, también, tipificados, nos bastará cambiar la escala de la transformación $Z = TP$, de manera que los nuevos factores tipificados **F**:

$F = D^{-1/2}P$ de donde $P = D^{1/2}F$ de forma que: $Z = TP = TD^{1/2}F$ será la transformación buscada como solución factorial y la matriz **A** de **patrón factorial** será: $A = TD^{1/2}$

Como consecuencia de que se conserva el carácter ortogonal de la matriz **A** tendremos que $Z = AF$, pero también $F = A'Z$; de manera que las variables originales pueden ponerse como la siguiente combinación lineal de las componentes principales pero, también, las componentes principales pueden ponerse como combinación lineal de las variables originales con los mismos coeficientes, pero tomados en un caso, por filas y, en otro, por columnas.

Este hecho de poder poner las componentes principales como combinación lineal de las variables posibilita el cálculo exacto de las puntuaciones factoriales cuando se emplea el método. (No como ocurre al emplear otros métodos que sólo pueden estimarse)

Por último, y antes de pasar a ver cómo el método de componentes principales puede verse como un método de análisis factorial, calculemos la matriz de correlación entre las componentes y las variables originales, esto es, la matriz de **estructura factorial** (si consideramos las componentes como factores comunes).

El coeficiente de correlación entre Z_j y F_p vendrá dado por:

$$r_{jp} = r_{Z_j F_p} = \frac{1}{N} \sum_{i=1}^N Z_{ji} F_{pi} = \frac{1}{N} \sum_{i=1}^N \left(F_{pi} \left(\sum_{r=1}^n a_{jr} F_{ri} \right) \right) = \frac{1}{N} \sum_{r=1}^n a_{jr} \sum_{i=1}^N F_{ri} F_{pi} = \sum_{r=1}^n a_{jr} \cdot \left(\frac{1}{N} \sum_{i=1}^N F_{ri} F_{pi} \right) = \sum_{r=1}^n a_{jr} \cdot r_{F_r F_p}$$

Como las componentes principales están incorrelacionadas todas las correlaciones

$r_{F_r F_p} = 0$ para $r \neq p$ y $r_{F_p F_p} = 1$ de modo que $r_{jp} = a_{jp}$ para todos los casos y, por tanto la matriz **patrón** y la de **estructura** coinciden.

Con el método de componentes principales, tal como lo hemos desarrollado, hemos conseguido disponer de un sistema para expresar las n variables originales en función de n factores

incorrelacionados. Sin embargo, no hemos conseguido la reducción de la dimensionalidad a un número más reducido de factores m , ($m < n$).

Pero podemos obtener esta reducción quedándonos sólo con un número de factores que nos produzca una mínima pérdida de información.

La forma de conseguir ésto es sencilla:

Consideremos las componentes principales antes de su tipificación. Recordemos que habíamos diagonalizado la matriz de varianzas de las variables originales, obteniendo una matriz diagonal D que tenía en su diagonal principal los valores propios de R (λ_j).

La matriz D será obviamente, la matriz varianzas de las componentes principales (antes de tipificar). De forma que el elemento j de la diagonal será la varianza de la j -sima componente.

La suma de todas las varianzas de las componentes valdrá lo mismo que la suma de todas las varianzas de las variables originales, es decir n ; de forma que:

$$\text{VarianzaTotal} = n = \sum_{j=1}^n \lambda_j$$

Y la contribución de cada componente a la variabilidad total vendrá dada por el cociente λ_j/n que si se quiere puede expresarse en tanto por cien.

Pues bien, si ordenamos las componentes de forma que tengan varianza decreciente tendremos que las componentes explicarán, por turno, la máxima cantidad posible de la varianza total. La primera componente será la combinación lineal de las variables originales que contribuirá a un máximo de su varianza total; la segunda componente, incorrelada con la primera, contribuirá a un máximo de la varianza residual, y así sucesivamente hasta que se analice la varianza total.

En el ejemplo: si sólo conserváramos la primera componente principal nos daría cuenta del $1,505/2=75,25\%$ de la varianza (variabilidad) inicial. Con las dos componentes tendríamos el 100%, claro.

Teniendo en cuenta esto, si queremos explicar las n variables originales con un número menor de factores comunes incorrelacionados, un método adecuado es coger las primeras componentes principales que nos explicarán un mayor porcentaje de la varianza total.

De esta forma podemos encajar la técnica de componentes principales con el principal objetivo del análisis factorial: la reducción de la dimensionalidad.

De forma práctica, suele emplearse uno de los siguientes tres criterios:

a) Prefijamos a priori un número m de factores comunes a extraer. Consideramos entonces como factores comunes las m primeras componentes principales.

b) Prefijamos un determinado porcentaje p de la varianza total que queremos explicar con el análisis. Entonces, consideraremos como factores explicativos las m primeras componentes que verifiquen que:

$$\sum_{j=1}^m \lambda_j \geq p \equiv \text{fracción mínima de la varianza total que se quiere explicar}$$

c) Por último, podemos considerar, para el análisis, la selección de todos aquellos factores que verifiquen siguiente criterio razonable: cualquiera de los factores escogidos debe ser capaz de explicar más proporción de la variabilidad total que cualquiera de las variables originales. Este criterio nos llevará a seleccionar todas aquellas componentes cuyo valor propio asociado sea mayor que 1.

Debemos matizar que si conservamos la totalidad de las componentes como factores comunes explicativos, la parte de la varianza explicada por todas las n componentes será su totalidad, y, en consecuencia, las comunalidades serán todas 1. Sin embargo, si utilizamos sólo una parte de las componentes principales, m de ellas, con $m < n$, la comunalidad será, en general, inferior a uno y tendrá pleno sentido de indicador de la bondad del análisis.

Considerado como método factorial, el análisis de componentes principales es el método más sencillo y permite obtener algunos resultados ventajosos como la incorrelación de los factores comunes, la sencillez del cálculo de las puntuaciones factoriales, y la máxima explicación de la varianza total. Es un método conveniente para obtener una solución factorial directa cuando circunstancias como la naturaleza de las variables o el carácter muestral de los datos no aconsejen la utilización de otro.

EL PROBLEMA DE LA ROTACION //MEJORA DE LA SOLUCIÓN FACTORIAL

Una vez obtenida una solución factorial por cualquiera de los métodos posibles, habremos obtenido un conjunto de m factores comunes que podrán darnos cuenta de una considerable proporción de las variables originales.

De esta forma en el espacio formado por los m factores comunes podremos representar la parte común de variables originales. Sin embargo, un espacio m -dimensional admite más de un sistema de referencia, para decirlo con exactitud, infinitos.

De forma que partiendo de la posibilidad de representar la realidad que investigamos en ese espacio m -dimensional, podemos referir su representación m -dimensional de infinitas maneras distintas.

Evidentemente, a cada orientación del espacio los factores comunes le corresponde una configuración diferente. Algunas de estas configuraciones podrán ser más sugestivas que otras, más fáciles de interpretar o más **simples**.

La posibilidad de que la **solución directa** (que hayamos obtenido por el método que hayamos seguido) no sea la más descriptiva o la mejor interpretable, aconseja, en ocasiones, buscar una solución derivada de la primera que podremos obtener por un cambio del sistema de referencia.

Este curso de acción plantea el problema de la determinación de una ventajosa solución derivada. Problema que también se conoce como problema de la **rotación**, pues de una rotación (adecuada) del sistema de referencia (de los factores obtenidos) es de lo que se trata.

Una vez obtenido nuestro **patrón A**, tratamos de buscar un nuevo patrón **B**, que verifique ciertas condiciones ventajosas y pueda extraerse por transformación lineal, de forma que :

$$\mathbf{B} = \mathbf{A} \mathbf{G}$$

Una vez especificadas las características buscadas en **B**, si la matriz **A** fuera invertible la determinación de la matriz de transformación (de rotación) sería sencilla:

$$\mathbf{G} = \mathbf{A}^{-1} \mathbf{B}$$

Pero normalmente la matriz **A** no es invertible, ya que el número de factores, m , es menor que el de variables, n . (A no ser que se hayan extraído las n componentes principales.)

El problema se resuelve planteando unos *criterios de estructura simple* sobre **B**, y, a continuación, obteniendo una transformación **G** adecuada, a partir de la expresión anterior por ajuste mínimo-cuadrático:

$$\mathbf{G} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{B}$$

Los criterios de estructura simple fueron propuestos, inicialmente, por Thurstone como tres condiciones

que debía cumplir la solución factorial final:⁷(*1)

- 1) Cada fila de la estructura factorial debe tener un cero, por lo menos.
- 2) Cada columna debe tener m ceros, por lo menos donde m es el número de factores comunes).
- 3) Para todo par de columnas debe haber, al menos, m variables cuyos elementos se anulen en una columna pero no en la otra.

Posteriormente, el propio Thurstone propuso dos condiciones adicionales:⁸

- 4) Para todo par de columnas, una gran proporción de las variables deben tener ceros en ambas, cuando haya cuatro o más factores.
- 5) Para todo par de columnas debe haber sólo un número pequeño de variables con entradas no nulas en ambas.

La idea fundamental de la estructura simple es tender a que las variables se aglomeren lo más posible en torno de los factores comunes y lo hagan de una manera lo más discriminativa. Con ello lo que se consigue es una mejor interpretabilidad de éstos y, por lo general, mayor significado teórico.

Básicamente, podemos plantearnos dos tipos de rotaciones a seguir: rotaciones ortogonales y rotaciones oblicuas.

En la solución final obtenida por rotación ortogonal, los factores estarán incorrelacionados entre sí. Patrón y estructura coincidirán. Además este tipo de transformaciones tiene la muy conveniente propiedad de conservar la comunalidad.

En la solución final obtenida por rotación oblicua los factores están correlacionados. Como consecuencia de ello, necesitamos especificar la solución con dos matrices: una patrón y una estructura. Por otro lado, para una solución oblicua ya no será cierto que la parte explicada por cada factor sea independiente de los demás factores. Sin embargo, en muchos casos prácticos puede ser muy deseable que ésto sea así.

Entre los principales métodos de rotación ortogonal destacan los métodos *varimax*, *cuartimax* y *ecumax*; y entre los oblicuos el *cuartimín*, el *oblimax*, y el *oblimín*.

5.LA INTERPRETACION DE LOS RESULTADOS Y OBTENCIÓN DE LAS PUNTUACIONES FACTORIALES

Un análisis factorial completo incluir una primera solución directa y, eventualmente, una solución derivada posterior, pretendidamente más explicativa. A la hora de interpretar la solución factorial (tanto si se trata de la inicial, como de la final) estaremos interesados en:

- 1) Detectar la estructura factorial subyacente.
- 2) Interpretar la naturaleza de los factores obtenidos y, a ser posible, definirlos en términos de alguna categoría teórica de interés científico.
- 3) Ser capaces de poner en relación funcional las variables originales y los factores comunes.
- 4) Obtener o evaluar los factores para cada individuo.
- 5) Disponer las conclusiones para posteriores análisis.

Con lo comentado en secciones anteriores, podemos muy bien señalar cómo acometer con los tres primeros puntos. Para discutir el cuarto, necesitaremos algún comentario adicional. Y, por último, el quinto requiere, como fácilmente se comprende, tener presente el contexto de la investigación global.

Obtener una solución factorial (ya sea directa o derivada) supone:

⁷ Thurstone, L.L.: "The Vectors of Mind" University of Chicago Press, 1935.

⁸ Thurstone, L.L.: "Multiple Factor Analysis" University of Chicago Press, 1947.

A) Si se trata de una solución de factores ortogonales, haber extraído la matriz factorial que nos dará cuenta tanto del patrón como de la estructura; pues recordemos que si los factores están incorrelacionados (componentes principales, p.ej.) patrón y estructura coinciden.

B) Si se trata de una solución oblicua (de factores correlacionados), haber extraído la matriz patrón, que nos informará de la relación funcional lineal entre Z y F; la matriz de estructura, que dará cuenta de las correlaciones entre variables y factores; y la matriz de correlación de los factores, que nos informará de cómo se correlacionan los factores entre sí.

Teniendo en cuenta esto pasemos a examinar los puntos relatados anteriormente:

Determinar la estructura factorial subyacente supone determinar el número, naturaleza y significado de los factores comunes que nos explican la realidad que estudiamos. Este propósito se consigue mediante el análisis del patrón y la estructura factorial.

Para interpretar la naturaleza de los factores obtenidos y, a ser posible, definirlos en términos de alguna categoría teórica de interés científico, tendremos que basarnos en las relaciones entre las variables y los factores y entre los factores entre sí, y comparar estas relaciones con los esquemas y categorías teóricas de nuestra parcela científica.

Esto nos lleva a estudiar las correlaciones entre variables y factores para arrojar cierta luz sobre la naturaleza de los factores e intentar su definición. Debemos pues trabajar con la estructura factorial.

Pero si los factores están correlacionados, su correcta interpretación teórica y su eventual definición no podrá desprejiciarse el análisis de estas correlaciones (entre factores).

Ser capaces de poner en relación las variables en función de los factores comunes es la primera pretensión del análisis factorial y, obviamente, la relación nos vendrá dada por el patrón extraído.

No será, sin embargo, tan sencillo, en general darle la vuelta a la relación y poner los factores en función de las variables para, de esta forma, poder, en su caso, evaluar los valores que toma cada factor para cada individuo: las llamadas **puntuaciones factoriales**.

En efecto, según la solución factorial obtenida: $Z = A F$, A es conocida tras la obtención de la solución. Si la matriz A fuera invertible, la relación de F con Z sería fácil de obtener, y podríamos, después, determinar puntuaciones. F sería sencillamente: $F = A^{-1}Z$

Pero A no es, por lo general, invertible, ya que se trata de una matriz no cuadrada.

Sin embargo, si la extracción factorial se por componentes principales esta matriz sí puede invertirse y recordando que A es una matriz ortogonal, tendremos que $Z = A F$, pero también $F = A'Z$

Lo que nos lleva al hecho, ya comentado, de que los mismos coeficientes que nos servían para obtener las variables en función de los factores, nos sirven, también, para obtener los factores en función de las variables.

En el caso de la solución por componentes principales será, por lo tanto, perfectamente posible determinar las puntuaciones factoriales de una manera *exacta*, sin más que sustituir en la expresión anterior Z los valores que toma cada variable para cada individuo.

En el caso de que la solución no sean las componentes principales, los factores y sus puntuaciones no podrán evaluarse sino, tan sólo, estimarse.

Existen distintos métodos de estimación de las puntuaciones factoriales, entre los que destacan por su aplicación fundamentalmente tres:

La estimación por regresión mínimo-cuadrática o método de Horst.

La estimación máximo verosímil o método de Anderson-Rubin, que presupone la normalidad de las variables originales y acaba estimando F mediante :

Y la estimación por minimización de un error específico o método de Bartlett

OTROS METODOS DE ANALISIS FACTORIAL

En otros métodos de AF distintos del Análisis de componentes principales la determinación de la matriz patrón **A** debe partir del denominado Teorema Thurston. Tal teorema partiendo de la expresión del modelo factorial : $\mathbf{Z} = \mathbf{AF} + \mathbf{YU}$ y considerando la incorrelación entre los distintos factores comunes (F) , los distintos factores específicos (Y) y la existente entre unos y otros, prueba que la matriz de correlación entre las variables originales R acaba pudiéndose expresar como :

$$\mathbf{R} = \mathbf{E}(\mathbf{ZZ}') = \mathbf{AA}' + \mathbf{U}^2 = \mathbf{R}^* + \mathbf{U}^2 \quad (\text{y por tanto: } \mathbf{R}^* = \mathbf{A A}')$$

Donde $\mathbf{U}^2$ es una matriz con las unicidades al cuadrado en la diagonal y ceros en los elementos no diagonales y $\mathbf{R}^*$ es la llamada **matriz de correlación reducida** en la que los unos de la diagonal se sustituyen por la comunalidades y los elementos no diagonales siguen siendo las correlaciones observadas entre las diferentes variables originales.

Sin embargo la ecuación

$\mathbf{R}^* = \mathbf{A A}'$ no tiene una solución única que nos permita determinar los coeficientes de **A**

Se hace necesario, para su obtención, imponer alguna condición adicional sobre la naturaleza o número de los factores a partir de una estimación de las comunalidades.

La indeterminación de la solución factorial lleva consigo que no haya un único método de análisis o una única solución.

Una vez efectuadas las hipótesis y restricciones (elegido el esquema de trabajo) es posible factorizar la matriz de correlación unívocamente. Esta solución será única en el sentido del criterio particular asumido: dos investigadores cualesquiera que acepten los mismos criterios y sigan los mismos procedimientos, llegarán a un resultado idéntico para el mismo conjunto de datos. Sin embargo, la elección de otros criterios nos conducirá generalmente a otra solución *única*.⁹

Ahora vamos a referirnos a los principales métodos directos de solución factorial. La elección de uno, en concreto, de los distintos métodos vendrá en función de las propiedades deseadas de la solución que más nos interesen: su simplicidad, su capacidad reductiva, la manera de contribuir los factores a la variación de las variables,

⁹ *Aún hay otro tipo de indeterminación posterior a la selección del criterio y método a emplear. En efecto, obtenida ya la solución, ésta puede transformarse cambiando el sistema de referencia, a través de una rotación del espacio de los factores. Actuar de esta manera nos conduce a una solución derivada o indirecta, que puede, de hecho, resultar más interesante desde el punto de vista científico, o más fácil de interpretar. Las soluciones derivadas resultarán siempre de la transformación de un método directo, y la determinación de la conveniencia y/o naturaleza del tipo de esta transformación es lo que habitualmente se conoce como problema de la rotación*

las exigencias de agrupación de las variables, las características deseadas del sistema de referencia, etc.

De acuerdo con (Harman, 1980), los distintos métodos de solución directa pueden clasificarse en dos grandes bloques:

- a) Métodos que requieren estimaciones de las comunalidades.
- b) Métodos que requieren la estimación del número de factores

Sin encajar en ninguno de estos dos bloques, puede considerarse al Análisis de Componentes Principales, como un método directo de solución factorial. De hecho, un enfoque factorial típico y seguido en muchas investigaciones, es la obtención de las componentes principales y su posterior rotación (generalmente ortogonal).

En cualquier caso hagamos ahora una breve reseña de las demás posibilidades metodológicas:

Entre los métodos que requieren la estimación de las comunalidades para la obtención de la solución factorial, cabe señalar el método de los Factores Principales (Ejes Principales), el método del Centroide, y la Descomposición Triangular.

El método de los Factores o Ejes Principales, es con mucho, el más utilizado en la actualidad, habiendo desplazado al del Centroide, muy popular en los años 30 y 40. Por último, la descomposición triangular no es, en absoluto, competitivo con relación a los otros dos.

Entre los métodos que requieren la estimación del número de factores comunes pueden incluirse el método de Máxima Verosimilitud, el Método Minres, los métodos psicométricos (método alfa y método imagen) y los primeros modelos factoriales, (los modelos de 1 y 2 factores de Spearman y el modelo bi-factorial de Holzinger).

Los tres primeros pueden aplicarse a cualquier matriz de correlación y suponen, de hecho, una alternativa al procedimiento de factores principales. Recurren a criterios estadísticos típicos (estimación máximo verosímil o mínimo cuadrática, en el caso de los dos primeros, respectivamente) o a consideraciones inferenciales particulares (los psicométricos). El método de grupos múltiples tiene un campo de aplicación más restringido y los últimos métodos (primeros, cronológicamente hablando) tan sólo tiene interés histórico.

METODO DE LOS FACTORES PRINCIPALES.

El método de los factores principales se basa, al igual que el de componentes principales, en el Teorema Espectral. A partir de este teorema se puede diagonalizar cualquier matriz definida positiva, mediante una transformación lineal de las variables. Un caso típico de matriz definida positiva es el de una matriz de correlación. Si se parte de la matriz de correlación observada cuyo rango es n (en general), la diagonalización exigirá una transformación lineal a un espacio vectorial de dimensión n (en general); ésto es lo que ocurre en el caso de las componentes principales, donde, en principio, no se produce una reducción de la dimensionalidad.

Sin embargo, si la matriz que se pretende diagonalizar es la correlación reducida (es decir, con las estimaciones de las comunalidades en la diagonal), su rango será m (un valor inferior a n), y se podrá obtener la diagonalización a través de una transformación lineal sobre un espacio de dimensión m ($m < n$).

En esencia, la diagonalización de la matriz de **correlación reducida** puede expresarse por la relación:

$$\mathbf{R}^* = \mathbf{T} \mathbf{D} \mathbf{T}'$$

Donde $\mathbf{R}^*$ es la matriz de correlación reducida, $\mathbf{D}$ es la matriz diagonal obtenida que tendrá en sus elementos diagonales los sucesivos valores propios de la matriz $\mathbf{R}^*$: λ_p , con $p = 1, 2, \dots, m$

(las soluciones de la ecuación: $\det [\mathbf{R}^* - \lambda \mathbf{I}] = 0$) y $\mathbf{T}$ es la matriz de cambio de base cuyas m columnas son los vectores propios de $\mathbf{R}^*$, con módulo unidad

Una vez diagonalizada la matriz $\mathbf{R}$, podemos obtener el patrón factorial $\mathbf{A}$. Teniendo en cuenta que por el teorema de Thurstone:

$$\mathbf{R}^* = \mathbf{A} \mathbf{A}'$$

y la expresión anterior $\mathbf{R}^* = \mathbf{T} \mathbf{D} \mathbf{T}'$

Es fácil obtener que $\mathbf{A} = \mathbf{T} \mathbf{D}^{1/2}$

donde $\mathbf{D}^{1/2}$ es una matriz diagonal cuyos elementos diagonales serán las respectivas raíces cuadradas de los valores propios.

Si el proceso de diagonalización se lleva a cabo de forma que consideremos los sucesivos valores propios de forma decreciente: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ encontraremos que los factores principales darán cuenta de una proporción decreciente de la comunalidad total. En otras palabras, el primer factor principal explica la mayor varianza posible, el segundo explica un máximo en el espacio residual del que se ha eliminado el primer factor, etc.

Básicamente el método es idéntico al de componentes principales, excepto en que, como se ha dicho, parte de la diagonalización de $\mathbf{R}^*$ en vez de la de $\mathbf{R}$.

Es necesario, por tanto estimar previamente las comunalidades, lo que suele hacerse partiendo de la matriz de correlaciones observadas, $\mathbf{R}$, por un proceso iterativo de reducción de la dimensionalidad.

METODO DEL CENTROIDE

El método del Centroide tiene por objeto aproximar los resultados que se obtienen con el método de los factores principales, con un ahorro considerable de esfuerzo, si los cálculos se hacen a mano.

También pretende explicar la mayor parte de la varianza total con cada uno de los factores sucesivos. Sin embargo la solución del centroide no es única y carece de las propiedades matemáticas interesantes de la solución factorial principal.

METODO DE LA MAXIMA VEROSIMILITUD

Este método se basa en consideraciones estadísticas, considerando explícitamente las diferencias entre las correlaciones observadas y los valores que deberían tomar estas correlaciones en la población de la que fue extraída la muestra.

La introducción de este método se debe a (Lawley, 1940)¹⁰ y fue mejorado y desarrollado por (Jöreskog, 1967)¹¹

Este método no requiere la estimación previa de las comunalidades, pero sí es necesaria una hipótesis referente al número de factores comunes. A partir de esta hipótesis se deduce una solución factorial y las comunalidades correspondientes.

Lo habitual en este método es no considerar las variables originales tipificadas y plantearnos el problema de factorizar la matriz de varianzas, C .

En este contexto el teorema de Thurstone adopta la expresión : $C=AA'+D^2$

Nos proponemos, entonces, la estimación estadística A , y nos planteamos obtener la estimación A que maximice la verosimilitud de las covarianzas muestrales observadas. Para realizar esta estimación suele postularse la normalidad de las variables originales. Obviamente, la estimación se realiza por el método de la máxima verosimilitud.

Una solución máximo verosímil tiene el mismo aspecto general que una solución factorial principal, pero no tiene la propiedad de explicar una cantidad máxima de varianza para un número especificado de factores.

EL METODO MINRES

Este método estima las cargas factoriales de forma que se minimicen los cuadrados de los residuos no diagonales de la matriz de correlación. El término "minres" es la contracción de " Minimum Residuals " (residuales mínimos),que es, como se ha dicho, la restricción impuesta para la estimación de la solución.

El método fue desarrollado por (Harman & Jones, 1966)¹².Parte y depende de una estimación previa del número de factores y, como el de máxima verosimilitud, no precisa la estimación de las comunalidades que se obtiene como subproducto del análisis, tras la estimación de las cargas.

¹⁰ Lawley, D. (1940). The estimation of Factor Loadings by the method of maximum Likelihood. *Proceedings of the Royal Society of Edinburgh*, 64-82.

¹¹ Jöreskog, K. (1967). Some contributions to Maximum Likelihood in Factor Analysis. *Psychometrika*, 31, 433-482.

¹² Harman, H., & Jones, W. (1966). Factor Analysis by Minimizing Residuals (Minres). *Psychometrika*, 31, 351-368.