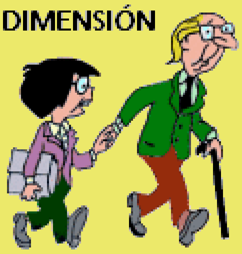
ANÁLISIS FACTORIAL

Fac. Ciencias Económicas y Empresariales

UNIVERSIDAD AUTONOMA DE MADRID

TÉCNICAS DE REDUCCIÓN DE LA DIMENSIÓN



Los métodos de reducción de la dimensión son métodos multivariantes de la independencia en el sentido de que todas sus variables tienen una importancia equivalente.

Técnicas de reducción en variables cuantitativas: Análisis Factorial y Análisis de Componentes Principales.

Técnicas de reducción en variables cualitativas: Análisis de Correspondencias y Escalonamiento Óptimo.

INTRODUCCIÓN ANÁLISIS FACTORIAL

El análisis factorial es una técnica de reducción de datos que sirve para encontrar grupos homogéneos de variables a partir de un conjunto numeroso de variables.

Los grupos homogéneos se forman con las variables que correlacionan mucho entre sí y procurando, inicialmente, que unos grupos sean independientes de otros.

Cuando se recogen un gran número de variables de forma simultánea (por ejemplo, en un cuestionario de satisfacción laboral) se puede estar interesado en averiguar si las preguntas del cuestionario se agrupan de alguna forma característica. Aplicando un análisis factorial a las respuestas de los sujetos se pueden encontrar grupos de variables con significado común y conseguir de este modo reducir el número de dimensiones necesarias para explicar las respuestas de los sujetos.

El Análisis Factorial es, por tanto, una técnica de reducción de la dimensionalidad de los datos. Su propósito último consiste en buscar el número mínimo de dimensiones capaces de explicar el máximo de información contenida en los datos.

A diferencia de lo que ocurre en otras técnicas como el análisis de varianza o el de regresión, en el análisis factorial todas las variables del análisis cumplen el mismo papel: todas ellas son independientes en el sentido de que no existe a priori una dependencia conceptual de unas variables sobre otras.

Fundamentalmente lo que se pretende con el Análisis Factorial (Análisis de Componentes Principales o de Factores Comunes) es simplificar la información que nos da una matriz de correlaciones para hacerla más fácilmente interpretable.
Se pretende encontrar una respuesta al preguntarnos ¿Por qué unas variables se relacionan más entre sí y menos con otras?. Hipotéticamente es porque existen otras variables, otras dimensiones o factores que explican por qué unos ítems se relacionan más con unos que con otros.
En definitiva, ¿se trata de un análisis de la estructura subyacente a una serie de variables?.

CONCEPTOS PREVIOS DEL ANÁLISIS FACTORIAL

Un ejemplo concreto de introducción al concepto de varianza compartida y varianza única: Sean unos ítems de una escala de actitudes, donde la puntuación de cada sujeto encuestado es la suma de las respuestas a todos los ítems, según la clave de corrección diseñada:

1º Me lo paso muy bien en mi casa, con mis padres	Muy de acuerdo = 5 De acuerdo = 4
2º Algunas veces me gustaría marcharme de mi casa	Muy de acuerdo = 1 De acuerdo = 2

- La varianza σ^2 de cada ítem indica la diferencia que crea en las respuestas. Si todos respondieran lo mismo la varianza sería cero, no habría diferencias. Si la mitad estuviera muy a gusto en su casa y la otra mitad muy a disgusto, la varianza sería máxima.

- Cada ítem o variable tiene su varianza (diferencias en las respuestas), la varianza de cada ítem puede ser compartida con la varianza de otros ítems: Algunos individuos encuestados están muy bien en su casa con sus padres (ítem 1) y nunca piensan irse de su casa (ítem 2). Otros individuos responderán con otras variaciones. En este caso, las respuestas señalada a estos dos ítems son coherentes con el significado pretendido de los dos ítems, comparten varianza porque los dos ítems están relacionados positivamente (estoy bien en casa, no me quiero ir).

Esta relación viene expresada por el coeficiente de correlación 'r' de Pearson, donde r^2 expresa la proporción de varianza común o de variación conjunta.

Es decir, si la correlación entre estos dos ítems es de 0,90, esto significa que tienen un 81% de varianza común (variación en las respuestas). El resto de la varianza (19%) no es varianza compartida.

- La varianza no compartida puede descomponerse en otras dos fuentes de varianza: Cada variable tiene una varianza especifica: un encuestado puede responder que 'se lo pasa muy bien con sus padres y que le gustaría irse de casa', simplemente porque le gusta viajar.

El ítem 1 no cuantifica únicamente la integración familiar, también tiene un significado específico que para muchos encuestados no puede coincidir del todo con 'sentirse bien en casa'.

También hay una Varianza de error de medición, ocasionada por cansancio, estilos personales de responder, orden en que se responde, etc.

- La varianza total de un ítem puede descomponerse:

Varianza Total	=	Varianza compartida o común	+	Varianza específica de cada variable	+	Varianza de errores de medición

Uniendo la varianza específica con la varianza debida a errores de medición (toda la varianza única o no compartida de cada ítem o variable), se tiene:

Varianza Total	=	Varianza compartida o común	+	Varianza de errores de medición

¿Qué hace el Análisis Factorial?

Se encarga de analizar la varianza común a todas las variables. Partiendo de una matriz de correlaciones, trata de simplificar la información que ofrece. Se opera con las correlaciones elevadas al cuadrado r^2 (coeficientes de determinación), que expresan la proporción de varianza común entre las variables.

En cada casilla de la matriz de correlaciones se refleja la proporción de varianza común a dos ítems o variables, excepto en la diagonal principal (donde cada ítem coincide consigo mismo). En los 1 de la diagonal principal se refleja la varianza que cada ítem o variable comparte con los demás y también los que no comparte (la específica o única de cada ítem).

Si se desea analizar exclusivamente la varianza compartida habrá que eliminar los unos de la matriz de correlaciones y poner en su lugar la proporción de varianza que cada ítem tiene en común con todos los demás.

En el Análisis Factorial, por tanto, caben dos enfoques:

1. Analizar TODA la varianza (común y no común). En este caso utilizamos los unos de la matriz de correlaciones. El método más usual es el de Análisis de Componentes Principales.

2. Analizar SOLO la varianza común. En este caso, se substituyen los unos de la diagonal por estimaciones de la varianza que cada ítem tiene en común con los demás (y que se denominan Comunalidades). Para la estimación de las comunalidades no hay un cálculo único, existen diversos procedimientos (correlaciones múltiples de cada ítem con todos los demás, coeficientes de fiabilidad si cada variable es un test). El procedimiento por el que se sustituyen los unos por las comunalidades se denomina Análisis de Factores Comunes.

Los dos enfoques caben bajo la denominación genérica de Análisis Factorial, aunque es el Análisis de Factores Comunes al que con más propiedad se le aplica la denominación de Análisis Factorial. Ambos enfoques dan resultados similares y se interpretan de manera casi idéntica.

¿Qué es un FACTOR?

En realidad los factores no existen, lo que existe de cada sujeto es una suma de sus respuestas a una serie de ítems o preguntas, una combinación lineal de variables (ítem a + ítem b + ítem c + …).

La suma total de ítems son distintos para cada sujeto, o pueden serlo, la varianza de los totales nos expresa la diversidad que existe entre los sujetos.

Si hay 'n' factores, se interpreta que el instrumento original se puede descomponer en 'n' instrumentos (cada uno compuesto por todos los ítems), aunque en cada instrumento los ítems tienen un 'peso específico' distinto según sea su relación con el factor:

Si se encuentra, por ejemplo, tres factores, esto quiere decir que se puede descomponer el instrumento original en tres instrumentos; cada uno está compuesto por todos los ítems, pero en cada instrumento los ítems tienen un peso específico distinto según sea su relación con cada factor:

$a_1 a + b_1 b + c_1 c + \ ....$ = Total en el Factor 1	a_1 es el peso específico del ítem a en el Factor 1
$a_2 a + b_2 b + c_2 c + \ ....$ = Total en el Factor 2	a_2 es el peso específico del ítem a en el Factor 2
....	....
$a_n a + b_n b + c_n c + \ ....$ = Total en el Factor n	a_n es el peso específico del ítem a en el Factor n

Las nuevas puntuaciones son las puntuaciones factoriales o factor scores.

Los pesos pueden ser grandes o pequeños, positivos o negativos. Generalmente, en cada factor hay ítems con pesos grandes y otros próximos a cero; los ítems que más pesan en cada factor son los que lo definen.

La varianza (diversidad) de todas las nuevas medidas equivale a la varianza de la medida original (no a toda, pero sí a la máxima que es posible explicar); estos factores indican las fuentes de varianza; si hay diferencias en la medida original es porque las hay en estas nuevas puntuaciones.

El análisis factorial se reduce a la búsqueda de estos pesos para localizar medidas distintas a partir de las variables originales, y de manera que, a poder ser, entre todas las nuevas medidas agoten o expliquen toda la varianza presente en las variables originales.

ESQUEMA DE UN ANÁLISIS FACTORIAL:

FORMULACIÓN DEL PROBLEMA

↓

ANÁLISIS DE LA MATRIZ DE CORRELACIÓN

↓

EXTRACCIÓN DE FACTORES

↓

DETERMINACIÓN DEL NÚMERO DE FACTORES

↓

ROTACIÓN DE FACTORES

↓

INTERPRETACIÓN DE FACTORES

↓

VALIDACIÓN DEL MODELO

↓

CÁLCULO DE PUNTUACIONES FACTORIALES	SELECCIÓN DE LAS VARIABLES REPRESENTATIVAS

↓

ANÁLISIS POSTERIORES: REGRESIÓN, CLUSTER...

Modelo del Análisis Factorial

Sean $(X_1, X_2, \ldots, X_p)$ las p variables objeto de análisis que supondremos en todo lo que sigue, que están tipificadas. Si no lo estuvieran el análisis se realizaría de forma similar pero la matriz utilizada para calcular los factores no sería la matriz de correlación sino la de varianzas y covarianzas.

El investigador mide estas variables sobre n individuos, obteniéndose la siguiente matriz de datos:

Sujetos	Variables X_1	X_2	$\ldots$	X_p
1	x_{11}	x_{12}	$\ldots$	x_{1p}
2	x_{21}	x_{22}	$\ldots$	x_{2p}
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
n	x_{n1}	x_{n2}	$\ldots$	x_{np}

El modelo del Análisis Factorial viene dado habitualmente por las ecuaciones:

$$
\begin{aligned}
X_1 &= a_{11} F_1 + a_{12} F_2 + \cdots + a_{1k} F_k + u_1 \\
X_2 &= a_{21} F_1 + a_{22} F_2 + \cdots + a_{2k} F_k + u_2 \\
&\text{------------------------------------} \\
X_p &= a_{p1} F_1 + a_{p2} F_2 + \cdots + a_{pk} F_k + u_p
\end{aligned}
$$

Donde, $(F_1, F_2, \cdots, F_k)$ $(k < p)$ son los factores comunes, $(u_1, u_2, \cdots, u_p)$ los factores únicos o específicos, y los coeficientes (a_{ij}) $\{i = 1, \cdots, p \; ; \; j = 1, \cdots, k\}$ las cargas factoriales.

Se supone que los Factores Comunes están a su vez estandarizados $[E(F_i) = 0 \; ; \; Var(F_i) = 1]$, los Factores Específicos tienen media 0 y están incorrelados $[E(u_i) = 0 \; ; \; Cov(u_i, u_j) = 0$ sí $i \neq j \; ;$ $(i, j = 1, \cdots, p)]$ y que ambos tipos de factores están incorrelados $Cov(F_i, u_j) = 0$, $\forall i = 1, \ldots, k$; $j = 1, \ldots, p$

Si, además, los Factores Comunes están incorrelados $\left[\text{Cov}(F_i, F_j) = 0 \text{ si } i \neq j \ ; \ j, i = 1, \cdots, k\right]$ estamos ante un modelo con factores ortogonales.

En caso contrario el modelo se dice que es de factores oblicuos.

En forma matricial: $x = A\,f + u \Leftrightarrow X = FA' + U$

$X \equiv$ matriz de datos
$A \equiv$ matriz de cargas factoriales
$F \equiv$ matriz de puntuaciones factoriales

$$\text{donde: } x = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}, \; f = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_k \end{pmatrix}, \; u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_p \end{pmatrix}, \; A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pk} \end{pmatrix}, \; F = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1k} \\ f_{21} & f_{22} & \cdots & f_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ f_{p1} & f_{p2} & \cdots & f_{pk} \end{pmatrix}$$

Utilizando las hipótesis anteriores, se tiene: $\text{Var}(X_i) = \sum_{j=1}^{k} a_{ij}^2 + \psi_i = h_i^2 + \psi_i \quad (i = 1, 2, \cdots, p)$

donde, $h_i^2 = \text{Var}\left(\sum_{j=1}^{k} a_{ij} F_j\right)$ y $\psi_i = \text{Var}(u_i)$ reciben los nombres, respectivamente, de Comunalidad y Especificidad de la variable X_i

En consecuencia, la varianza de cada una de las variables analizadas se puede descomponer en dos partes: la Comunalidad h_i^2 que representa la varianza explicada por los factores comunes y la Especificidad ψ_i que representa la parte de la varianza específica de cada variable. Además se tiene:

$$\text{Cov}(X_i, X_l) = \text{Cov}\left(\sum_{j=1}^{k} a_{ij} F_j \, , \sum_{j=1}^{k} a_{lj} F_j \, , \right) = \sum_{j=1}^{k} a_{ij}\, a_{lj} \quad \forall i \neq l$$

por lo que son los factores comunes los que explican las relaciones existentes entre las variables.

Por este motivo, los factores comunes tienen interés y son susceptibles de interpretación experimental. Los factores únicos se incluyen en el modelo dada la imposibilidad de expresar, en general, p variables en función de un número más reducido k de factores.

Ejemplo.- Unos estudiantes son sometidos a diversos test en distintas materias para medir sus actitudes intelectuales. Como consecuencia, se obtienen una serie de puntuaciones estandarizadas en Matemáticas (Ma), Física (Fi), Química (Qu), Inglés (In), Historia (Hi) y Dibujo (Di).

El modelo factorial viene dado por las ecuaciones

$$\begin{array}{ll} Ma = 0{,}8F_1 + 0{,}2F_2 + U_{Ma} & In = 0{,}2F_1 + 0{,}8F_2 + U_{In} \\ Fi = 0{,}7F_1 + 0{,}3F_2 + U_{Fi} & Hi = 0{,}15F_1 + 0{,}82F_2 + U_{In} \\ Qu = 0{,}6F_1 + 0{,}3F_2 + U_{Qu} & Di = 0{,}25F_1 + 0{,}85F_2 + U_{Di} \end{array}$$

Los factores comunes están estandarizados e incorrelados

E[Fi] = 0 $\forall$i=1,2 ; j $\in$ {Ma, Fi, Qu, In, Hi, Di}
Var[Fi] = 1 i=1,2;
Cov(F1,F2) = 0

Los factores específicos tienen media 0 e incorrelados

$E[u_i] = 0 \quad \forall i = 1,2 ; \; j \in \{Ma, Fi, Qu, In, Hi, Di\}$
$Cov(u_1, u_2) = 0 \quad \forall i \neq j \in \{Ma, Fi, Qu, In, Hi, Di\}$

Ambos tipos de factores están incorrelados $\quad \begin{array}{l} \text{Cov}(F_i, u_j) = 0 \quad \forall i \neq j \in \{\text{Ma, Fi, Qu, In, Hi, Di}\} \\ \text{Cov}(u_1, u_2) = 0 \quad \forall i \neq j \in \{\text{Ma, Fi, Qu, In, Hi, Di}\} \end{array}$

Matriz de cargas factoriales : $A = \begin{pmatrix} 0,8 & 0,2 \\ 0,7 & 0,3 \\ 0,6 & 0,3 \\ 0,2 & 0,8 \\ 0,15 & 0,82 \\ 0,25 & 0,85 \end{pmatrix}$

Comunalidad y Especificidad:

Matemáticas

$$\begin{aligned} \text{Var}[Ma] &= 1 = \text{Var}[0,8\, F_1 + 0,2\, F_2 + u_{Ma}] = \\ &= 0,8^2\, \text{Var}[F_1] + 0,2^2\, \text{Var}[F_2] + \text{Var}[u_{Ma}] + \cancel{2(0,8)(0,2)\, \text{Cov}(F_1, F_2)} + \\ &+ \cancel{2(0,8)\, \text{Cov}(F_1, u_{Ma})} + \cancel{2(0,2)\, \text{Cov}(F_2, u_{Ma})} = 0,68 + \psi_{Ma} \end{aligned}$$

La Comunalidad en Matemáticas es $h^2_{Ma} = 0,68$ y la Especificidad $\psi_{Ma} = 0,32$

Dibujo

$$\begin{aligned} \text{Var}[Di] &= 1 = \text{Var}[0,25\, F_1 + 0,85\, F_2 + u_{Di}] = \\ &= 0,25^2\, \text{Var}[F_1] + 0,85^2\, \text{Var}[F_2] + \text{Var}[u_{Di}] + \cancel{2(0,25)(0,85)\, \text{Cov}(F_1, F_2)} + \\ &+ \cancel{2(0,25)\, \text{Cov}(F_1, u_{Di})} + \cancel{2(0,85)\, \text{Cov}(F_2, u_{Di})} = 0,785 + \psi_{Di} \end{aligned}$$

La Comunalidad en Dibujo es $h^2_{Di} = 0,785$ y la Especificidad $\psi_{Di} = 0,215$

Análogamente,

Comunalidades

Matemáticas	0,68
Física	0,42
Química	0,55
Inglés	0,215
Historia	0,36
Dibujo	0,785

Como las puntuaciones están estandarizadas, la matriz de varianzas y covarianzas coincide con la matriz de correlaciones:

$$\begin{array}{c} \\ \text{Ma} \\ \text{Fi} \\ \text{Qu} \\ \text{In} \\ \text{Hi} \\ \text{Di} \end{array} \begin{array}{c} \begin{array}{cccccc} \text{Ma} & \text{Fi} & \text{Qu} & \text{In} & \text{Hi} & \text{Di} \end{array} \\ \begin{pmatrix} 1 & 0,62 & 0,54 & 0,32 & 0,284 & 0,37 \\ 0,62 & 1 & 0,51 & 0,38 & 0,351 & 0,43 \\ 0,54 & 0,51 & 1 & 0,36 & 0,336 & 0,405 \\ 0,32 & 0,38 & 0,36 & 1 & 0,686 & 0,73 \\ 0,284 & 0,351 & 0,336 & 0,686 & 1 & 0,7345 \\ 0,37 & 0,43 & 0,405 & 0,73 & 0,7345 & 1 \end{pmatrix} \end{array}$$

Adviértase que,

$\text{Cov}(\text{Ma, Fi}) = \text{Cov}(0{,}8\, F_1 + 0{,}2\, F_2 + u_{Ma}\,,\; 0{,}7\, F_1 + 0{,}3\, F_2 + u_{Fi}) =$
$= (0{,}8)\,(0{,}7)\,\text{Var}(F_1) + (0{,}8)\,(0{,}3)\,\text{Cov}(F_1,F_2) + (0{,}8)\,\text{Cov}(F_1,u_{Fi}) + (0{,}2)\,(0{,}7)\,\text{Cov}(F_2,F_1) +$
$+ (0{,}2)\,(0{,}3)\,\text{Var}(F_2) + (0{,}2)\,\text{Cov}(F_2,u_{Fi}) + (0{,}7)\,\text{Cov}(u_{Ma},F_1) + (0{,}3)\,\text{Cov}(u_{Ma},F_2) + \text{Cov}(u_{Ma},\, u_{Fi}) =$
$= 0{,}56 + 0{,}06 = 0{,}62$

$\text{Cov}(\text{Fi, Qu}) = \text{Cov}(0{,}7\, F_1 + 0{,}3\, F_2 + u_{Fi}\,,\; 0{,}6\, F_1 + 0{,}3\, F_2 + u_{Qu}) =$
$= (0{,}7)\,(0{,}6)\,\text{Var}(F_1) + (0{,}6)\,(0{,}3)\,\text{Cov}(F_1,F_2) + (0{,}7)\,\text{Cov}(F_1,u_{Qu}) + (0{,}3)\,(0{,}6)\,\text{Cov}(F_2,F_1) +$
$+ (0{,}3)\,(0{,}3)\,\text{Var}(F_2) + (0{,}3)\,\text{Cov}(F_2,u_{Qu}) + (0{,}6)\,\text{Cov}(u_{Fi},F_1) + (0{,}3)\,\text{Cov}(u_{Fi},F_2) + \text{Cov}(u_{Fi},\, u_{Qu}) =$
$= 0{,}42 + 0{,}09 = 0{,}51$

ANÁLISIS DE LA MATRIZ DE CORRELACIÓN

La finalidad de analizar la matriz de las correlaciones muestrales $R = (r_{ij})$, donde r_{ij} es la correlación muestral observada entre las variables (X_i, X_j), es comprobar si sus características son las adecuadas para realizar un Análisis Factorial.

Uno de los requisitos que deben cumplirse es que las variables se encuentran altamente intercorrelacionadas. También se espera que las variables que tengan correlación muy alta entre sí la tengan con el mismo factor o factores.

En consecuencia, si las correlaciones entre todas las variables son bajas, tal vez no sea apropiado el Análisis Factorial.

Existen varios indicadores para analizar la matriz de correlación:

- Test de esfericidad de Barlett

Contrasta, bajo la hipótesis de normalidad multivariante, si la matriz de correlación de las p variables observadas (R_p) es la identidad.

Si una matriz de correlación es la identidad significa que las intercorrelaciones entre las variables son cero. Si se confirma la hipótesis nula $H_0 : |R_p| = 1$ o $R_p = I$, las variables no están intercorrelacionadas.

El test de esfericidad de Barlett se obtiene mediante una transformación del determinante de la matriz de correlación. El estadístico del test viene dado por:

$$d_R = -\left[n - 1 - \frac{1}{6}(2p+5)\right]\log|R| = -\left[n - \frac{(2p+11)}{6}\right]\sum_{j=1}^{p}\log(\lambda_j)$$

donde n es el número de individuos de la muestra y $\lambda_j \;\; (j = 1, \dots, p)$ son los valores propios de R.

Bajo la hipótesis nula, el estadístico se distribuye asintóticamente según una $\chi^2_{p(p-1)/2}$

Si la hipótesis nula es cierta, los valores propios valdrán uno, o su logaritmo será nulo y, por tanto, el estadístico del test valdría cero.

Por el contrario, si con el test de Barlett se obtienen valores altos de χ^2, o un determinante bajo, hay variables con correlaciones altas (un determinante próximo a cero indica que una o más variables podrían ser expresadas como combinación lineal de otras variables).

En definitiva, si el estadístico del test toma valores grandes (o un determinante próximo a cero) se rechaza la hipótesis nula con cierto grado de significación. En caso de aceptarse la hipótesis nula, las variables no están intercorreladas y debería reconsiderarse la aplicación de un Análisis Factorial.

Medidas de adecuación de la muestra

El coeficiente de correlación parcial es un indicador del grado de relaciones entre dos variables, eliminando la influencia del resto.

Si las variables comparten factores comunes, el coeficiente de correlación parcial entre pares de variables es bajo, puesto que se eliminan los efectos lineales de las otras variables.

Las correlaciones parciales son estimaciones de las correlaciones entre los factores únicos, debiendo ser próximas a cero cuando el Análisis Factorial es adecuado, dado que se supone que los factores únicos están incorrelados entre sí.

En definitiva, si existe un número elevado de coeficientes de correlación parcial distintos *de cero*, se interpreta que las hipótesis del modelo factorial no son compatibles con los datos.

Una manera de cuantificar este hecho es con la Media de Adecuación de la Muestra KMO propuesta por Kaiser-Meyer-Olkin:

$$\text{KMO} = \frac{\sum_{j\neq i}\sum_{i\neq j} r_{ij}^2}{\sum_{j\neq i}\sum_{i\neq j} r_{ij}^2 + \sum_{j\neq i}\sum_{i\neq j} r_{ij(p)}^2} \qquad 0 \leq \text{KMO} \leq 1$$

donde $r_{ij(p)}$ es el coeficiente de correlación parcial entre (X_i, X_j) eliminando la influencia del resto de las variables.

El índice KMO se utiliza para comparar las magnitudes de los coeficientes de correlación parcial, de forma que cuánto más pequeño sea su valor, mayor será el valor de los coeficientes de correlación parciales $r_{ij(p)}$ y, en consecuencia, menos apropiado es realizar un Análisis Factorial.

Kaiser-Meyer-Olkin para realizar un Análisis Factorial, proponen:

$$\begin{array}{l} \text{KMO} \geq 0{,}75 \Rightarrow \text{Bien} \\ \text{KMO} \geq 0{,}5 \Rightarrow \text{Aceptable} \\ \text{KMO} < 0{,}5 \Rightarrow \text{Inaceptable} \end{array}$$

La experiencia práctica aconseja que es precipitado tomar el índice KMO como única medida de adecuación de la muestra a las hipótesis del modelo de Análisis Factorial, sobre todo si hay un número pequeño de variables consideradas.

Para tomar la decisión de eliminar una variable del estudio es aconsejable complementar la información con otras fuentes: las comunalidades de cada variable, los residuos del modelo, e interpretar los factores obtenidos.

EXTRACCIÓN DE FACTORES

El objetivo del Análisis Factorial (AF) es determinar un número reducido de factores que puedan representar a las variables originales.

Una vez que se ha determinado que el AF es una técnica apropiada para analizar los datos, hay que seleccionar el método adecuado para la extracción de factores. Existen diversos métodos, cada uno de ellos con sus ventajas e inconvenientes.

El modelo factorial en forma matricial: $X = FA' + U$, teniendo que cuantificar la matriz A de cargas factoriales que explica X en función de los factores.

Partiendo de $X = FA' + U$, se deduce la llamada Identidad Fundamental del Análisis Factorial:

$$R_p = AA' + \psi$$

donde R_p es la matriz de correlación poblacional de las variables $(X_1, X_2, \ldots, X_p)$ y $\psi = \text{diag}(\psi_i)$ es la matriz diagonal de las especificidades.

En este sentido, surgen dos problemas:

(a) Problemas de Grados de Libertad.- Igualando cada elemento de la matriz R_p con el correspondiente de la combinación lineal $(AA' + \psi)$, resultan (p x p) ecuaciones, que es el número de elementos de R.

Ahora bien, la matriz R_p es simétrica y, en consecuencia, está integrada por $\dfrac{p\,(p+1)}{2}$ elementos distintos, que es el número real de ecuaciones.

En el segundo miembro de la igualdad, los parámetros a estimar con (p x k) elementos de la matriz A y los p-elementos de la matriz ψ.

En consecuencia, para que pueda efectuarse el proceso de estimación se requiere que el número de ecuaciones sea mayor o igual que el número de parámetros a estimar: $\dfrac{p(p+1)}{2} \geq p(k+1)$, o lo que es equivalente, $k \leq \dfrac{p-1}{2}$.

(b) No Unicidad de la Solución.- Las soluciones dadas por la matriz A no son únicas, puesto que cualquier transformación ortogonal de A es también solución.

Si T es una matriz ortogonal, $TT' = T'T = I$, al aplicar una solución ortogonal de A se obtiene una solución distinta al sistema anterior. Esta es la base de los métodos de rotación de factores.

En consecuencia, si T es una matriz ortogonal $\Rightarrow \; A^* = AT$ es solución

Se define $F^* = FT$ (F* es el vector F rotado por la matriz ortogonal T).

Se verifica que X y R_p verifican también las ecuaciones del modelo, es decir:

$$R_p = A^*A^{*'} + \psi = (AT)(T'A') + \psi = AA' + \psi$$

$$X = F^*A^{*'} + U = (FT)(T'A') + U = FA' + U$$

Por tanto, el modelo es único salvo rotaciones ortogonales, es decir, se pueden realizar rotaciones de la matriz de las ponderaciones o cargas factoriales sin alterar el modelo.

Ejemplo.- En el modelo factorial definido anteriormente, se tenía:

$$\begin{pmatrix} Ma \\ Fi \\ Qu \\ In \\ Hi \\ Di \end{pmatrix} = \begin{pmatrix} 0,8 & 0,2 \\ 0,7 & 0,3 \\ 0,6 & 0,3 \\ 0,2 & 0,8 \\ 0,15 & 0,82 \\ 0,25 & 0,85 \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} + \begin{pmatrix} U_{Ma} \\ U_{Fi} \\ U_{Qu} \\ U_{In} \\ U_{Hi} \\ U_{Di} \end{pmatrix} \Rightarrow \begin{array}{ll} Ma = 0,8F_1 + 0,2F_2 + U_{Ma} & In = 0,2F_1 + 0,8F_2 + U_{In} \\ Fi = 0,7F_1 + 0,3F_2 + U_{Fi} & Hi = 0,15F_1 + 0,82F_2 + U_{In} \\ Qu = 0,6F_1 + 0,3F_2 + U_{Qu} & Di = 0,25F_1 + 0,85F_2 + U_{Di} \end{array}$$

Si se definen los factores: $\begin{cases} F_1' = \dfrac{1}{\sqrt{2}} F_1 + \dfrac{1}{\sqrt{2}} F_2 \\ F_2' = -\dfrac{1}{\sqrt{2}} F_1 + \dfrac{1}{\sqrt{2}} F_2 \end{cases}$, siendo la matriz ortogonal $T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$

$$\begin{pmatrix} F_1' \\ F_2' \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} F_1' \\ F_2' \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \Rightarrow \begin{cases} F_1 = \dfrac{1}{\sqrt{2}} F_1' - \dfrac{1}{\sqrt{2}} F_2' \\ F_2 = \dfrac{1}{\sqrt{2}} F_1' + \dfrac{1}{\sqrt{2}} F_2' \end{cases}$$

de donde,

$$Ma = 0{,}8F_1 + 0{,}2F_2 + U_{Ma} = \frac{1}{\sqrt{2}} F_1' - \frac{0{,}6}{\sqrt{2}} F_2' + U_{Ma} = 0{,}71\, F_1' - 0{,}42\, F_2' + U_{Ma}$$

$$Fi = 0{,}7F_1 + 0{,}3F_2 + U_{Fi} = \frac{1}{\sqrt{2}} F_1' - \frac{0{,}4}{\sqrt{2}} F_2' + U_{Ma} = 0{,}71\, F_1' - 0{,}28\, F_2' + U_{Fi}$$

$$Qu = 0{,}6F_1 + 0{,}3F_2 + U_{Qu} = \frac{0{,}9}{\sqrt{2}} F_1' - \frac{0{,}3}{\sqrt{2}} F_2' + U_{Ma} = 0{,}64\, F_1' - 0{,}21\, F_2' + U_{Qu}$$

$$In = 0{,}71\, F_1' + 0{,}42\, F_2' + U_{In} \qquad Hi = 0{,}69\, F_1' + 0{,}47\, F_2' + U_{Hi} \qquad Di = 0{,}78\, F_1' + 0{,}42\, F_2' + U_{Di}$$

verificándose que $Cov(F_1', F_2') = 0$, por lo que las nuevas cargas factoriales serán las correlaciones de los nuevos factores con las variables originales.

Las comunalidades, especificidades y matrices de correlación permanecen igual.

La nueva matriz de cargas factoriales será: $B = \begin{pmatrix} 0{,}71 & -0{,}42 \\ 0{,}71 & -0{,}28 \\ 0{,}64 & -0{,}21 \\ 0{,}71 & 0{,}42 \\ 0{,}69 & 0{,}47 \\ 0{,}78 & 0{,}42 \end{pmatrix}$

La forma de calcular la matriz de rotación T y la de nueva cargas factoriales B da lugar a los distintos métodos de rotación ortogonales, siendo los métodos más utilizados: Varimax, Quartimax y Equamax.

MÉTODOS DE EXTRACCIÓN DE FACTORES

Existen diferentes métodos para obtener los factores comunes, los implantados en SPSS son: Método de las Componentes Principales, Método de los Ejes principales y Método de Máxima Verosimilitud.

- **Método de las Componentes Principales.-** Consiste en estimar las puntuaciones factoriales mediante las puntuaciones tipificadas de las primeras k-componentes y la matriz de cargas factoriales mediante las correlaciones de las variables originales con dichas componentes.

  Este método tiene la ventaja de que siempre proporciona una solución.

  Tiene el inconveniente de que al no estar basado en el modelo de Análisis Factorial puede llevar a estimadores muy sesgados de la matriz de cargas factoriales, especialmente, si existen variables con Comunalidades bajas.

- **Método de los Ejes Principales.-** Basado en la Identidad Fundamental del Análisis Factorial $R_p = AA' + \psi$, sustituyendo la matriz de las correlaciones poblacionales R_p por las correlaciones muestrales R, con lo que: $R' = R\,\psi' = AA'$

Respetando $R' = R\,\psi' = AA'$, el método es iterativo y consiste en alternar una estimación de la matriz de las especificidades ψ con una estimación de la matriz de las cargas factoriales A.

Se parte de una estimación inicial de la matriz ψ, $\psi^{(0)}$, y en el paso i-ésimo del algoritmo se verifica que $R\psi^{(i)} = A^{(i)}A^{(i)'}$.

La estimación $A^{(i)}$ se obtiene aplicando el método de las componentes principales a la matriz $R - \psi^{(i-1)}$. Posteriormente, se calcula $\psi^{(i)}$ a partir de la igualdad $R\psi^{(i)} = A^{(i)}A^{(i)'}$ y se itera hasta que los valores de dichas estimaciones apenas cambien.

Este método tiene la ventaja de estar basado en el modelo del Análisis Factorial por lo que suele proporcionar mejores estimaciones que el método de componentes principales. Sin embargo, no garantiza su convergencia, sobre todo en muestras pequeñas.

- Método de la Máxima Verosimilitud.- Basado en el modelo $x = Af + u \rightarrow X = FA' + U$, adoptando la hipótesis de normalidad multivariante, aplica el método de la máxima verosimilitud.

  Sobre los anteriores, tiene la ventaja de que las estimaciones obtenidas no dependen de la escala de medida de las variables.

  Por otra parte, como está basado en el método de máxima verosimilitud, tiene todas las propiedades estadísticas de éste y, en particular, es asintóticamente insesgada, eficiente y normal si las hipótesis del modelo factorial son ciertas.
  Además, permite seleccionar el número de factores mediante contrastes de hipótesis.

  Este método también puede ser utilizado en el Análisis Factorial Confirmatorio, donde el investigador puede plantear hipótesis como que algunas cargas factoriales son nulas, que algunos factores están correlacionados con determinados factores, etc., y aplicar tests estadísticos para determinar si los datos confirman las restricciones asumidas.

  El principal inconveniente del método radica en que, al realizarse la optimización de la función de verosimilitud por métodos iterativos, si las variables originales no son normales, puede haber problemas de convergencia sobre todo en muestras finitas.

- Método Mínimos cuadrados no ponderados.- Para un número fijo de factores, genera una matriz de coeficientes que minimiza la suma de las diferencias al cuadrado entre las matrices de correlación observada R y reproducida $\tilde{R} = \tilde{A}\tilde{A}'$, eliminando en las diferencias los elementos de la diagonal.

- Método Mínimos cuadrados generalizados.- Minimiza el mismo criterio - La suma de las diferencias al cuadrado entre las matrices de correlación observada R y reproducida $\tilde{R} = \tilde{A}\tilde{A}'$ - ponderando las correlaciones inversamente por la varianza del factor específico. Este método permite, además, aplicar contraste de hipótesis para determinar el número de factores.

- Método de Factorización por imágenes.- Consiste en aplicar el método de componentes principales a la matriz de correlaciones $\tilde{R}$ obtenida a partir de las partes predichas de las diversas regresiones lineales de cada una de las variables sobre las demás (dicha parte recibe el nombre de imagen de la variable).

- Método Alfa.- Maximiza el alfa de Cronbach para los factores $x = Af + u \rightarrow X = FA' + U$

Comparación entre distintos Métodos

- Cuando las comunalidades son altas (> 0,6) todos los procedimientos tienen a dar la misma solución.
- Cuando las comunalidades son bajas para algunas de las variables, el método de componentes principales tiende a dar soluciones muy diferentes del resto de los métodos, con cargas factoriales mayores.
- Si el número de variables es alto (> 30), las estimaciones de la comunalidad tienen menos influencia en la solución obtenida y todos los métodos tienden a ofrecer el mismo resultado.
- Si el número de variables es bajo, todo depende del método utilizado para estimar las comunalidades y de si éstas son altas más que del método utilizado para estimarlas.
- Es más robusto utilizar un método para el modelo de factores comunes. El único problema puede ser la falta de convergencia del método utilizado.

DETERMINAR EL NÚMERO DE FACTORES

La matriz factorial puede representar un número de factores superior al necesario para explicar la estructura de los datos originales. Generalmente, hay un conjunto pequeño de factores, los primeros, que contienen casi toda la información. El resto de factores suelen contribuir relativamente poco.

Uno de los problemas consiste en determinar el número de factores que conviene conservar, pues se trata de cumplir el principio de parsimonia.

Existen diversas reglas y criterios para determinar el número de factores a conservar, algunos de los más utilizados son:

(a) Determinación "a priori".- Es el criterio más fiable si los datos y las variables están bien elegidos y el investigador conoce la situación, lo ideal es plantear el Análisis Factorial con una idea previa de cuántos factores hay y cuáles son.

(b) Regla de Kaiser.- Calcula los valores propios de la matriz de correlaciones R y toma como número de factores el número de valores propios superiores a la unidad.
Este criterio es una alusión del Análisis de Componentes Principales y se ha verificado en simulaciones que, generalmente, tiende a infraestimar el número de factores por lo que se recomienda su uso para establecer un límite inferior. Un límite superior se calcularía aplicando este mismo criterio tomando como límite 0,7.

(c) Criterio del porcentaje de la varianza.- Es una alusión del Análisis de Componentes Principales y consiste en tomar como número de factores el número mínimo necesario para que el porcentaje acumulado de la varianza explicado alcance un nivel satisfactorio (75%, 80%).
Tiene la ventaja de que se puede aplicar también cuando la matriz analizada es la de varianzas y covarianzas, pero no tiene ninguna justificación teórica o práctica.

(d) Criterio de Sedimentación.- Se trata de la representación gráfica donde los factores están en el eje de abscisas y los valores propios en el de ordenadas.
Los factores con varianzas altas suelen diferenciarse de los factores con varianzas bajas. Se pueden conservar los factores situados antes de este punto de inflexión.
En simulaciones el criterio ha funcionado bien, tiene el inconveniente de que depende del 'ojo' del analista.

(e) Criterio de división a la mitad.- La muestra se divide en dos partes iguales tomadas al azar y se realiza el Análisis Factorial en cada una de ellas.
Solo se conservan los factores que tienen alta correspondencia de cargas de factores en las dos muestras. Antes de aplicarlo, conviene comprobar que no existen diferencias significativas entre las dos muestras en lo que se refiere a las variables estudiadas.

PRUEBAS DE SIGNIFICACIÓN

Para seleccionar el número, consiste en aplicar contrastes de hipótesis de modelos anidados. Este criterio se puede utilizar si el método empleado para estimar los factores es el de *máxima verosimilitud*.

En la mayor parte de los caso exploratorios k no puede ser especificado por adelantado y, en consecuencia, se utilizan procedimientos secuenciales para determinar k.

Se comienza usualmente con $k = 1$ (valor pequeño), los parámetros en el modelo factorial son estimados utilizando el método de máxima verosimilitud. Si el estadístico del test no es significativo, se acepta el modelo con este número de factores, en caso contrario, se aumenta $k = 2$ y se repite el proceso hasta alcanzar una solución aceptable.

El principal inconveniente de este método es que está basado en resultados asintóticos y que, si el tamaño de la muestra es grande, se corre el riesgo de tomar el valor k excesivamente grande puesto que el test detecta cualquier factor por pequeño que sea su poder explicativo.

INTERPRETACIÓN DE LOS FACTORES

La interpretación de los factores se basa en las correlaciones estimadas de los mismos con las variables originales.

El modelo de Análisis Factorial es cierto, si se verifica:

$$\text{Corre}(X_i, F_l) = \text{Cov}(X_i, F_l) = \sum_{j=1}^{k} a_{ij}\,\text{Cov}(F_j, F_l) \qquad \forall i = 1, \cdots, p \;\; ; \;\; l = 1, \cdots, k$$

y, en particular, si los factores son ortogonales: $\text{Corre}(X_i, F_l) = a_{il} \qquad \forall i = 1, \cdots, p \;\; ; \;\; l = 1, \cdots, k$

Como se observa, la matriz de cargas factoriales (A) tiene un papel fundamental en la interpretación. Por otra parte, las cargas factoriales al cuadrado (a_{il}^2) indican si los factores son ortogonales, qué porcentaje de la variable original (X_i) es explicado por el factor F_l.

A efectos prácticos, en la interpretación de los factores, señalar:

- Identificar las variables cuyas correlaciones con el factor son las más elevadas en valor absoluto.
- Intentar dar un nombre a los factores. El nombre se asigna de acuerdo con la estructura de las correlaciones: Cuando es positiva (resp. negativa) la relación entre el factor y dicha variable es directa (resp. inversa).
  Analizando con qué variables tiene una relación fuerte es posible, en muchos casos, tener una idea más o menos clara de cuál es el significado de un factor.
- Una ayuda en la interpretación de los factores puede ser la representación gráfica de los resultados obtenidos. La representación se hace tomando los factores dos a dos. Cada factor representa un je de coordenadas. A estos ejes se les denomina ejes factoriales.
  Sobre los ejes factoriales se proyectan las variables originales.
  Las coordenadas vienen dadas por los respectivos coeficientes de correlación entre la variable y el factor, de forma que las variables saturadas en un mismo factor aparecen agrupadas. Esto puede servir de ayuda para descubrir la estructura latente de este factor.
  - Las variables al final de un eje son aquellas que tienen correlaciones altas sólo en ese factor y, en consecuencia, lo describen.
  - Las variables cerca del origen tienen correlaciones reducidas en ambos factores.
  - Las variables que no están cerca de ninguno de los ejes se relacionan con ambos factores.

- Ordenar la matriz factorial de forma que las variables con cargas altas para el mismo factor aparezcan juntas.
- Eliminar las cargas factoriales bajas y de este modo suprimir información redundante. El investigador decide a partir de qué valor deben eliminarse las cargas factoriales.
  De cara a una mayor facilidad interpretativa, el investigador puede ordenar la matriz factorial y eliminar las cargas factoriales bajas.
  Generalmente, se toma como significativas las cargas superiores a 0,5 en valor absoluto. Aunque, si el factor es más tardío o el número de variables es grande, se eleva el valor mínimo de la carga factorial significativa.

Ejemplo.- En el modelo factorial definido, se tenía:

$$\begin{pmatrix} Ma \\ Fi \\ Qu \\ In \\ Hi \\ Di \end{pmatrix} = \overbrace{\begin{pmatrix} 0,8 & 0,2 \\ 0,7 & 0,3 \\ 0,6 & 0,3 \\ 0,2 & 0,8 \\ 0,15 & 0,82 \\ 0,25 & 0,85 \end{pmatrix}}^{\text{matriz cargas factoriales}} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} + \begin{pmatrix} U_{Ma} \\ U_{Fi} \\ U_{Qu} \\ U_{In} \\ U_{Hi} \\ U_{Di} \end{pmatrix} \rightarrow \begin{array}{ll} Ma = 0,8F_1 + 0,2F_2 + U_{Ma} & In = 0,2F_1 + 0,8F_2 + U_{In} \\ Fi = 0,7F_1 + 0,3F_2 + U_{Fi} & Hi = 0,15F_1 + 0,82F_2 + U_{In} \\ Qu = 0,6F_1 + 0,3F_2 + U_{Qu} & Di = 0,25F_1 + 0,85F_2 + U_{Di} \end{array}$$

$$Corr(Ma, F_1) = Cov(Ma, F_1) = Cov(0,8F_1 + 0,2F_2 + U_{Ma}) = 0,8\,Var(F_1) + 0,2\,Cov(F_2, F_1) + Cov(U_{Ma}, F_1) = 0,8$$

En general, como $F_1 \perp F_2 \quad \Rightarrow$ Las correlaciones de las calificaciones de los test con dichos factores vendrán dadas por las cargas factoriales.

Observando la matriz de las cargas factoriales, se aprecia que el factor F_1 está muy relacionado con la variables Ma, Fi y Qu, pero poco relacionado con In, Hi y Di. De otra parte, el factor F_2 está muy relacionado con In, Hi y Di y poco con las restantes.

Análogamente, analizando la matriz de cargas factoriales correspondientes a los factores F_1 y F_2 :

$$B = \begin{pmatrix} 0,71 & -0,42 \\ 0,71 & -0,28 \\ 0,64 & -0,21 \\ 0,71 & 0,42 \\ 0,69 & 0,47 \\ 0,78 & 0,42 \end{pmatrix}$$

Se observa que el factor $F_1^{'}$ está muy relacionado con todas las variables de forma directa y, en consecuencia, podría interpretarse como un factor de inteligencia general.

Por su parte, el factor $F_2^{'}$ destaca en la aptitud verbal, al estar relacionado de forma inversa con Ma, Fi y Qu.

Cabe preguntarse ¿Cuál es la interpretación más correcta?. Todo dependerá de la teoría que subyace al problema que llevará al analista a hacer más hincapié en una interpretación u otra. De cualquier modo, tendrá que validar el modelo elegido.

ROTACIÓN DE LOS FACTORES

La matriz de cargas factoriales tiene un papel importante para interpretar el significado de los factores. Cuando los factores son ortogonales cuantifican el grado y tipo de la relación entre éstos y las variables originales.

En la práctica, los métodos de extracción de factores pueden no proporcionar matrices de cargas factoriales adecuadas para la interpretación.

Para acometer este problema están los procedimientos de Rotación de Factores que, a partir de la solución inicial, buscan factores cuya matriz de cargas factoriales los hagan más fácilmente interpretables.

Estos métodos intentan aproximar la solución obtenida al Principio de Estructura Simple (Louis Leon Thurstone, 1935), según el cual la matriz de cargas factoriales debe reunir tres características:

a) Cada factor debe tener unos pocos pesos altos y los demás próximos a cero.
b) Cada variable no debe estar saturada más que en un factor.
c) No deben existir factores con la misma distribución, esto es, dos factores distintos deben presentar distribuciones diferentes de cargas altas y bajas.

De esta manera, dado que hay más variables que factores comunes, cada factor tendrá una correlación alta con un grupo de variables y baja con el resto de las variables.

Al examinar las características de las variables de un grupo asociado a un determinado factor se pueden encontrar rasgos comunes que permitan identificar el factor y darle una denominación que responda a esos rasgos comunes.
Si se consigue identificar claramente estos rasgos, además de reducir la dimensión del problema, también se desvela la naturaleza de las interrelaciones existentes entre las variables originales.

Existen dos formas básicas de realizar la Rotación de Factores: $\begin{cases} \text{Rotación Ortogonal} \\ \text{Rotación Oblicua} \end{cases}$

Se elige uno u otro procedimiento según que los factores rotados sigan siendo ortogonales o no.

Señalar que en ambas rotaciones la comunalidad de cada variable no se modifica, esto es, la rotación no afecta a la bondad del ajuste de la solución factorial: aunque cambie la matriz factorial, las especificidades no cambian y, en consecuencia, las comunidades permanecen invariantes. Sin embargo, cambia la varianza explicada por cada factor, por tanto, los nuevos factores no están ordenados de acuerdo con la información que contienen, cuantificada mediante su varianza.

Rotación Ortogonal.- Los ejes se rotan de forma que quede preservada la incorrelación entre los factores. Es decir, los nuevos ejes (ejes rotados) son perpendiculares de igual forma que lo son los factores sin rotar.

La rotación se apoya en el problema de falta de identificabilidad de los factores obtenidos por rotaciones ortogonales, de forma que si T es una matriz ortogonal con $TT' = T'T = I$, entonces:

$$X = FA' + U = FTT'A' + U = GB' + U$$

La matriz G geométricamente es una rotación de F, verificando las mismas hipótesis que ésta.

Realmente lo que se realiza es un giro de ejes, de forma que cambian las cargas factoriales y los factores.

Se trata de buscar una matriz T tal que la nueva matriz de cargas factoriales B tenga muchos valores nulos o casi nulos, y unos pocos valores cercanos a la unidad de acuerdo con el principio de estructura simple.

Los métodos empleados en la rotación ortogonal de factores son: Varimax, Quartimax, Equamax, Oblimin y Promax.

- Método Varimax.- Es un método de rotación que minimiza el número de variables con cargas altas en un factor, mejorando así la interpretación de factores.

  El método considera que, si se logra aumentar la varianza de las cargas factoriales al cuadrado de cada factor consiguiendo que algunas de sus cargas factoriales tiendan a acercarse a 1 mientras que otras se aproximan a 0, se obtiene una pertenencia más clara e inteligible de cada variable al factor.

  Los nuevos ejes se obtienen maximizando la suma para los k-factores retenidos de las varianzas de las cargas factoriales al cuadrado dentro de cada factor.

  Para evitar que las variables con mayores comunalidades tengan más peso en la solución final, se efectúa la normalización de Kaiser (dividiendo cada carga factorial al cuadrado por la comunalidad de la variable correspondiente).

  En consecuencia, el método Varimax determina la Matriz B de forma que maximice la suma de las varianzas:

$$V = p\sum_{i=1}^{k}\sum_{j=1}^{p}\left(\frac{b_{ij}}{h_j}\right)^2 - \sum_{i=1}^{k}\left(\sum_{j=1}^{p}\frac{b_{ij}^2}{h_j^2}\right)^2$$

- Método Quartimax.- El objetivo es que cada variable tenga correlaciones elevadas con un pequeño número de factores. Para ello, maximiza la varianza de las cargas factoriales al cuadrado de cada variable en los factores, es decir, se trata de maximizar la función:

$$S = \sum_{i=1}^{p} k \sum_{j=1}^{p}\left(b_{ij}^2 - b_i^2\right)^2 \quad \text{donde,} \quad b_i^2 = \frac{1}{k}\sum_{j=1}^{k} b_{ij}^2$$

  Con ello, se logra que cada variable concentre su pertenencia en un determinado factor, esto es, presente una carga factorial alta mientras que, en los demás factores, sus cargas factoriales tienden a ser bajas.

  De este modo, la interpretación gana en claridad por cuanto la comunalidad total de cada variable permanece constante, quedando más evidente hacia qué factor se inclina con más fuerza cada variable.

  El método será más clarificador, cuanto mayor número de factores se hayan calculado. Este método tiende a producir un primer factor general, conocido con el nombre de tamaño, y el resto de factores presentan ponderaciones menores que las dadas por el método Varimax.

- Método Equamax.- Trata de maximizar la media de los criterios anteriores. Con un comportamiento similar al de los métodos anteriores.

Rotación oblicua.- En este caso la matriz T de rotación no tiene que ser ortogonal (cuando una matriz multiplicada por su transpuesta es la matriz identidad $TT' = I$) sino únicamente no singular (matriz cuadrado cuyo determinante no es cero)

De esta manera, los factores rotados no tienen por qué ser ortogonales y tener, por tanto, correlaciones distintas de cero entre sí.

La rotación oblicua puede utilizarse cuando es probable que los factores en la población tengan una correlación muy fuerte.

Es necesario ir con mucha atención en la interpretación de las rotaciones oblicuas, pues la superposición de factores puede confundir la significación de los mismos.

De esta forma, el análisis gana más flexibilidad y realismo pero a riesgo de perder robustez, por lo que conviene aplicar estos métodos si el número de observaciones por factor es elevada.

Ejemplo.- En el modelo factorial definido, se tenía:

$$\begin{pmatrix} Ma \\ Fi \\ Qu \\ In \\ Hi \\ Di \end{pmatrix} = \overbrace{\begin{pmatrix} 0,8 & 0,2 \\ 0,7 & 0,3 \\ 0,6 & 0,3 \\ 0,2 & 0,8 \\ 0,15 & 0,82 \\ 0,25 & 0,85 \end{pmatrix}}^{\text{matriz cargas factoriales}} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} + \begin{pmatrix} U_{Ma} \\ U_{Fi} \\ U_{Qu} \\ U_{In} \\ U_{Hi} \\ U_{Di} \end{pmatrix} \rightarrow \begin{matrix} Ma = 0,8F_1 + 0,2F_2 + U_{Ma} & In = 0,2F_1 + 0,8F_2 + U_{In} \\ Fi = 0,7F_1 + 0,3F_2 + U_{Fi} & Hi = 0,15F_1 + 0,82F_2 + U_{In} \\ Qu = 0,6F_1 + 0,3F_2 + U_{Qu} & Di = 0,25F_1 + 0,85F_2 + U_{Di} \end{matrix}$$

Si se definen los factores: $\begin{cases} F_1'' = \dfrac{4}{\sqrt{17}} F_1 + \dfrac{1}{\sqrt{17}} F_2 \\ F_2'' = \dfrac{1}{\sqrt{17}} F_1 + \dfrac{4}{\sqrt{17}} F_2 \end{cases}$ $\quad \begin{pmatrix} F_1'' \\ F_2'' \end{pmatrix} = \begin{pmatrix} 4/\sqrt{17} & 1/\sqrt{17} \\ 1/\sqrt{17} & 4/\sqrt{17} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$

$\text{Corr}(F_1'', F_2'') = \dfrac{8}{17} = 0,47 \neq 0 \rightarrow$ Los nuevos factores estarán correlacionados.

$$\begin{pmatrix} F_1'' \\ F_2'' \end{pmatrix} = \begin{pmatrix} 4/\sqrt{17} & 1/\sqrt{17} \\ 1/\sqrt{17} & 4/\sqrt{17} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \Rightarrow \frac{15}{17} \begin{pmatrix} 4/\sqrt{17} & -1/\sqrt{17} \\ -1/\sqrt{17} & 4/\sqrt{17} \end{pmatrix} \begin{pmatrix} F_1'' \\ F_2'' \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \Rightarrow \begin{cases} F_1 = \dfrac{4\sqrt{17}}{15} F_1'' - \dfrac{\sqrt{17}}{15} F_2'' \\ F_2 = -\dfrac{\sqrt{17}}{15} F_1'' + \dfrac{4\sqrt{17}}{15} F_2'' \end{cases}$$

de donde,

$$Ma = 0,8F_1 + 0,2F_2 + U_{Ma} = 0,8\left[\frac{4\sqrt{17}}{15} F_1'' - \frac{\sqrt{17}}{15} F_2''\right] + 0,2\left[-\frac{\sqrt{17}}{15} F_1'' + \frac{4\sqrt{17}}{15} F_2''\right] + U_{Ma} =$$

$$= 0,82\, F_1'' + 0\, F_2'' + U_{Ma}$$

$$Fi = 0,7F_1 + 0,3F_2 + U_{Fi} = 0,7\left[\frac{4\sqrt{17}}{15} F_1'' - \frac{\sqrt{17}}{15} F_2''\right] + 0,3\left[-\frac{\sqrt{17}}{15} F_1'' + \frac{4\sqrt{17}}{15} F_2''\right] + U_{Fi} =$$

$$= 0,69\, F_1'' + 0,14\, F_2'' + U_{Fi}$$

y así sucesivamente.

En este caso, se tiene que la matriz de rotación: $T = \dfrac{15}{17} \begin{pmatrix} 4/\sqrt{17} & -1/\sqrt{17} \\ -1/\sqrt{17} & 4/\sqrt{17} \end{pmatrix} = \begin{bmatrix} \dfrac{4\sqrt{17}}{15} & -\dfrac{\sqrt{17}}{15} \\ -\dfrac{\sqrt{17}}{15} & \dfrac{4\sqrt{17}}{15} \end{bmatrix}$

La matriz de configuración: $B = \overbrace{\begin{pmatrix} 0{,}82 & 0{,}00 \\ 0{,}69 & 0{,}14 \\ 0{,}58 & 0{,}17 \\ 0{,}00 & 0{,}82 \\ -0{,}06 & 0{,}86 \\ 0{,}04 & 0{,}87 \end{pmatrix}}^{\text{matriz de cargas factoriales}}$

La matriz de la estructura será aquella que contiene las correlaciones de las variables originales con los nuevos factores:

$$\begin{pmatrix} 0{,}82 & 0{,}00 \\ 0{,}69 & 0{,}14 \\ 0{,}58 & 0{,}17 \\ 0{,}00 & 0{,}82 \\ -0{,}06 & 0{,}86 \\ 0{,}04 & 0{,}87 \end{pmatrix} \begin{pmatrix} 1 & 8/17 \\ 8/17 & 1 \end{pmatrix} = \overbrace{\begin{pmatrix} 0{,}82 & 0{,}39 \\ 0{,}76 & 0{,}46 \\ 0{,}66 & 0{,}44 \\ 0{,}39 & 0{,}82 \\ 0{,}34 & 0{,}83 \\ 0{,}45 & 0{,}89 \end{pmatrix}}^{\text{matriz estructura}}$$

Se observa de nuevo que F_1'' se puede interpretar de nuevo como un factor donde destaca en (Matemáticas, Física y Química), mientras que F_2'' destaca en (Inglés, Historia y Dibujo), a diferencia de los factores F_1 y F_2 que son incorrelados, en esta nueva estimación ambos factores tienen una correlación positiva significativa, lo que proporciona más realismo al análisis realizado.

- Método Oblimin.- Busca minimizar la expresión: $\displaystyle\sum_{s<q=1}^{k} \left[\alpha \sum_{i=1}^{p} b_{is}^2\, b_{iq}^2 + (1-\alpha) \sum_{i=1}^{p} (b_{is}^2 - \overline{b_s^2})(b_{is}^2 - \overline{b_s^2}) \right]$

$$\sum_{s<q=1}^{k} \sum_{i=1}^{p} b_{is}^2\, b_{iq}^2 \equiv \text{controla la interpretabilidad de los factores}$$

$$\sum_{s<q=1}^{k} \sum_{i=1}^{p} (b_{is}^2 - \overline{b_s^2})(b_{is}^2 - \overline{b_s^2}) \equiv \text{controla la ortogonalidad de los factores}$$

  - Para $\alpha = 1$ se alcanza el máximo grado de oblicuidad.
  - Cuánto más α se aproxima a 0, más ortogonales son los factores.

En la rotación oblicua, como los factores están correlacionados entre sí, las cargas factoriales no coinciden con las correlaciones entre el factor y la variable.

Por este motivo, los paquetes estadísticos calculan dos matrices:

1. La matriz de cargas factoriales que muestra la contribución única de cada variable al factor.
2. La matriz de estructura factorial que muestra las correlaciones entre los factores y las variables, mostrando información acerca de la contribución única y de las correlaciones entre factores.

Además de estas dos matrices, conviene analizar la matriz de correlaciones entre factores.
Si las correlaciones entre los factores son muy pequeñas es más robusto aplicar rotaciones ortogonales.

De otra parte, si dos factores están muy correlacionados puede ser porque estén midiendo el mismo concepto y que, por tanto, haya que reducir el número de factores.

- Método Promax.- Altera los resultados de una rotación ortogonal hasta crear una solución con cargas factoriales lo más próximas a la estructura ideal.

  La estructura ideal se obtiene elevando a una potencia (entre 2 y 4) las cargas factoriales obtenidas en una rotación ortogonal. Cuanto mayor sea la potencia, más oblicua es la solución obtenida.

  Sea H la matriz de cargas buscada por el método Promax, busca una matriz T tal que $A\,T = H$.

  Multiplicando ambos miembros por la matriz $(A'A)^{-1}A'$, se tiene: $T = (A'A)^{-1}A'H$

CÁLCULO DE PUNTUACIONES FACTORIALES

Habiendo determinado los factores rotados, se calcula las matrices de puntuaciones factoriales F.

Son variadas las posibilidades de analizar las puntuaciones factoriales de los sujetos:

- Conocer qué sujetos son los más raros o extremos, es decir, la representación gráfica de las puntuaciones factoriales para cada par de ejes factoriales facilita detectar casos atípicos.
- Conocer dónde se ubican ciertos grupos o subcolectivos de la muestra (ejemplo; clase alta frente a clase baja, una provincia frente a las otras provincias, jóvenes frente a mayores, etc.)
- Conocer en qué factor sobresalen unos sujetos y n qué factor no.
- Explicar, atendiendo las informaciones anteriores, por qué han aparecido dichos factores en el análisis factorial realizado.

Es necesario conocer los valores que toman los factores en cada observación, pues en ocasiones, el Análisis Factorial es un paso previo a otros análisis: Regresión Múltiple o Análisis Cluster, en los que sustituye el conjunto de variables originales por los factores obtenidos.

Métodos del Cálculo de las Puntuaciones.- Existen diversos métodos de estimación de la matriz F, las propiedades deseables que verificasen los factores estimados son:

— Cada factor estimado presente una correlación alta con el verdadero factor.

— Cada factor estimado tenga correlación nula con los demás factores verdaderos.

— Los factores estimados son incorrelados dos a dos (mutuamente ortogonales si son ortogonales).

— Los factores estimados sean estimadores insesgados de los verdaderos factores.

Señalar que el problema de estimación es complejo por la propia naturaleza de los factores comunes. Se puede demostrar que los factores no son, en general, combinación lineal de las variables originales.

Por otra parte, en la mayoría de las situaciones, no existirá una solución exacta ni siquiera será única.

Todos los métodos de obtención de puntuaciones factoriales parten de la expresión $X = F\,A' + U$, con $E[U] = 0$, $Var[U] = \psi$, buscando estimar el valor de F.

Los métodos de estimación más utilizados: Regresión, Barlett, Anderson-Rubin

Método de Regresión.- Estima F por el método de los mínimos cuadrados: $\hat{F} = (A'A)^{-1}A'X$

Método de Barlett.- Utiliza el método de los mínimos cuadrados generalizados estimando las puntuaciones factoriales mediante: $\hat{F} = (A'\psi^{-1}A)^{-1}A'\psi^{-1}X$

Método de Anderson-Rubin.- Estima F mediante el método de los mínimos cuadrados generalizados, imponiendo la condición $F'F = I$

$$\hat{F} = (A'\psi^{-1}R\,\psi^{-1}A)^{-1}A'\psi^{-1}X$$

Análisis de los tres métodos:

El Método de Regresión da lugar a puntuaciones con máxima correlación con las puntuaciones teóricas. Sin embargo, el estimador no es insesgado, ni unívoco y, en caso de que los factores sean ortogonales, puede dar lugar a puntuaciones correladas.

El Método de Barlett da lugar a puntuaciones correladas con las puntuaciones teóricas, insesgadas y unívocas. Sin embargo, en caso de que los factores sean ortogonales, puede dar lugar a puntuaciones correladas.

El Método de Anderson-Rubin da lugar a puntuaciones ortogonales que están correladas con las puntuaciones teóricas. Sin embargo, el estimador no es insesgado ni unívoco.

Selección de Variables.- El investigador en ocasiones desea seleccionar las variables más representativas de los factores, en lugar de calcular sus puntuaciones.

Por ejemplo, si se utiliza el Análisis Factorial para reducir el número de datos, por razones de economía, si se quieren aplicar los resultados obtenidos a objetos diferentes de los estudiados en el análisis, es más interesante seleccionar algunas de las variables originalmente medidas - dada la dificultad del cálculo de las puntuaciones factoriales para las que se necesitaría medir todas las variables utilizadas en el estudio -.

Una forma de llevar a cabo la selección de variables es estudiar la matriz de correlaciones de las variables con los factores, seleccionando como representante de cada factor la variable con la correlación más elevada en éste, que sea más fácil de medir y que tenga más sentido desde un punto de vista teórico.

En cualquier caso, conviene elegir las variables de forma que una misma variable no se utilice para medir dos factores distintos.

Una vez elegidas las variables, se les asigna pesos basados en su correlación con el factor, y se comprueba su validez estimando su correlación con los factores que desea estimar mediante la fórmula $R_{fs} = A'W\,\text{diag}(R_{ss})$ donde R_{ss} es la matriz de correlaciones de las puntuaciones estimadas.

VALIDACIÓN DEL MODELO.- El último paso en el Análisis Factorial es estudiar la validez del modelo. El proceso debe realizarse en dos direcciones: Analizando la bondad de ajuste y la Generalidad de los resultados.

Bondad de Ajuste.- Una suposición básica subyacente al Análisis Factorial es que la correlación observada entre las variables puede atribuirse a factores comunes.

Por consiguiente, las correlaciones entre variables pueden deducirse o reproducirse a partir de las correlaciones estimadas entre las variables y los factores.

A fin de determinar el ajuste del modelo, pueden estudiarse las diferencias (residuos) entre las correlaciones observadas (matriz de correlación de entrada) y las correlaciones reproducidas (como se estiman a partir de la matriz factorial).

El modelo factorial es adecuado cuando los residuos son pequeños.

Si hay un porcentaje elevado de residuos superiores a una cantidad pequeña prefijada (por ejemplo, 0,05), será una indicación de que el modelo factorial estimado no se ajusta a los datos.

Se sabe además que hay más estabilidad en los resultados si el número de casos por variable es alto.

Generalidad de los resultados.- Es conveniente refrendar los resultados del primer análisis factorial realizando nuevos análisis factoriales sobre nuevas muestras extraídas de la población objeto de estudio y, en caso de no ser posible, sobre submuestras de la muestra original.

En cada caso habrá que estudiar qué factores de los calculados son corroborados en los distintos análisis llevados a cabo.

Otra posibilidad es realizar nuevos análisis factoriales modificando las variables consideradas, bien sea eliminando aquellas variables que no tienen relación con ningún factor o eliminando las variables con relaciones más fuertes tratando de descubrir cómo se comporta el resto de ellas sin su presencia.

Otro de los procedimientos metodológicos y estadísticos que complementan y profundizan las interpretaciones que se deducen del análisis factorial consiste en la realización de otros análisis factoriales en base, no al conjunto total de la muestra o población, sino referido a subcolectivos o grupos que están presentes en la muestra y que pueden formarse utilizando las categorías de las variables primarias (sexo, clase social, tipo de centro, tipo de metodología pedagógica, tipos de actitud, etc.).

Lo que se desprende de los trabajos e investigaciones que han utilizado este procedimiento es que generalmente la interpretación que se da y que es válida para el conjunto total de sujetos debe modificarse, en algunos casos sustancialmente, cuando se refiere a esos subcolectivos. En caso de ser así, se deriva una doble conclusión:

(a) Las variables se comportan en el Análisis Factorial de distinta forma según de qué muestra se trate.

(b) No existe el sujeto 'tipo' sino que existen diferentes 'tipos' de sujetos en la muestra global.

Finalmente, se debería plantear un Análisis Factorial Confirmatorio para comprobar los resultados obtenidos en la versión de Análisis Factorial Exploratorio.

Resumen.- El Análisis Factorial es una técnica estadística multivariante cuya finalidad es analizar las relaciones de interdependencia existentes entre un conjunto de variables, calculando un conjunto de variables latentes, denominadas *factores*, que explican con un número menor de dimensiones, dichas relaciones.

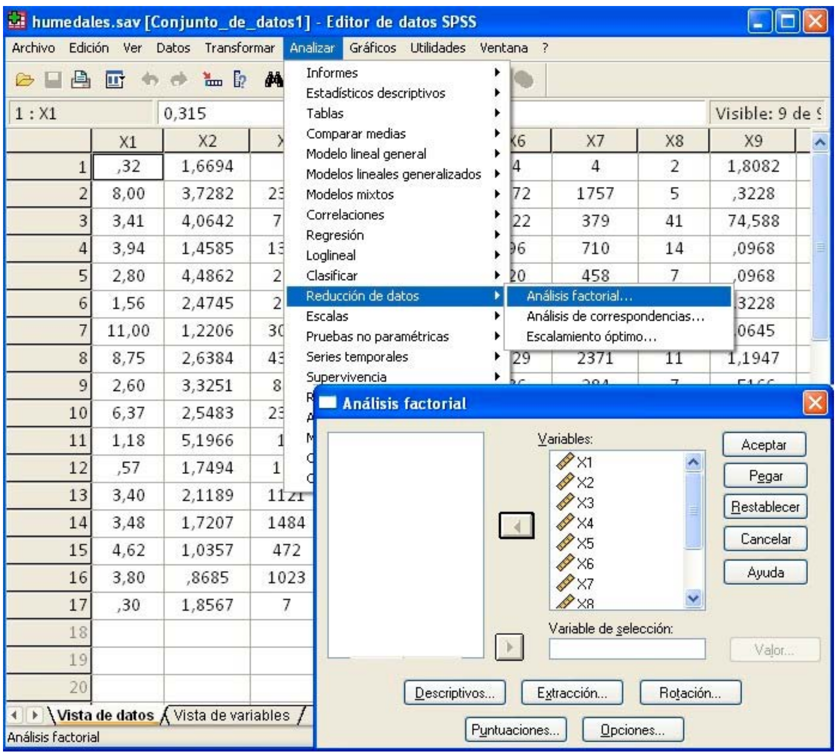
Por este motivo, el Análisis Factorial es una técnica de reducción de datos con un número menor de variables sin distorsionar dicha información, lo que aumenta el grado de manejo e interpretación de la misma.

📖 SPSS 1.- Los datos adjuntos corresponden a la medición de 17 humedales en determinada época del año. Las variables medidas han sido:

1. V1: Conductividad eléctrica
2. V2: Contenido en bicarbonatos
3. V3: Contenido en cloruros
4. V4: Contenido en sulfatos
5. V5: Contenido en calcio
6. V6: Contenido en magnesio
7. V7: Contenido en sodio
8. V8: Contenido en potasio
9. V9: Contenido en fosfatos

Humedal	V1	V2	V3	V4	V5	V6	V7	V8	V9
Caja	0,315	1,6694	5	86	55	4	4	2	1,8082
Camuñas	8	3,7282	2388	7638	2123	972	1757	5	0,3228
Capacete	3,41	4,0642	732	881	218	122	379	41	74,588
Cerero	3,94	1,4585	1359	772	251	96	710	14	0,0968
Chica	2,8	4,4862	220	2510	572	20	458	7	0,0968
Dulce	1,56	2,4745	269	495	157	38	162	9	0,3228
FP Salinas	11	1,2206	3038	923	233	226	1488	11	0,0645
FP Vicaria	8,75	2,6384	4325	456	234	229	2371	11	1,1947
Grande	2,6	3,3251	840	2270	609	86	284	7	0,5166
Gualdal. May	6,37	2,5483	2320	1040	1294	192	485	23	0,4843
Hoyos1	1,18	5,1966	13	499	202	20	5	18	6,7807
Lobón	0,57	1,7494	110	42	21	12	60	6	0,5812
Marcela	3,4	2,1189	1121	866	157	115	643	4	0,7426
Ratosa	3,48	1,7207	1484	554	151	151	708	7	0,1291
Redonda	4,62	1,0357	472	2964	752	160	652	34	0,1291
Salada	3,8	0,8685	1023	2274	1946	360	430	23	0,5489
Viso	0,3	1,8567	7	15	39	3	4	2	4,4882

Las variables están medidas en distintas unidades, teniendo que tipificar en su momento
Para realizar en SPSS el Análisis Factorial por el método de Componentes Principales:
Analizar/Reducción de Datos/Análisis Factorial



En el botón [Extracción] se puede cambiar la opción de método deseado, SPSS realiza por defecto el método de Componentes principales. Los métodos disponibles son: Componentes principales, Mínimos cuadrados no ponderados, Mínimos cuadrados generalizados, Máxima verosimilitud, Factorización de Ejes principales, Factorización Alfa y Factorización Imagen.

Lo primero que se realiza es determinar la estructura factorial necesaria, en la opción [Extraer] se utiliza el método de Kaiser que determina tantos factores como autovalores mayores que 1. Es el método por defecto que realiza SPSS.

La regla de Kaiser proporciona una estructura factorial con tres factores que explican el 81,946% de la varianza total.

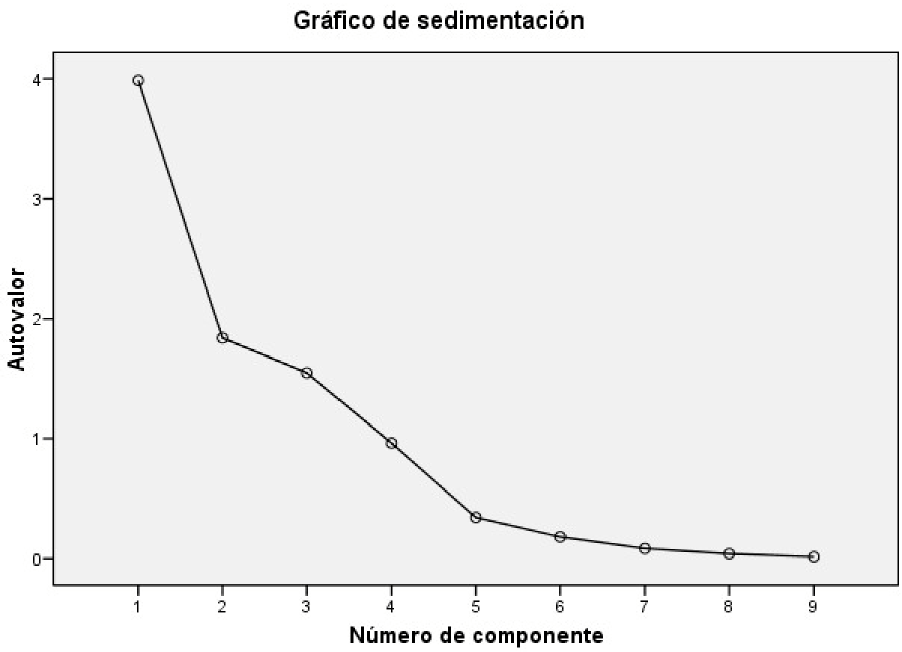
Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3,986	44,291	44,291	3,986	44,291	44,291
2	1,840	20,448	64,739	1,840	20,448	64,739
3	1,549	17,207	81,946	1,549	17,207	81,946
4	,962	10,693	92,639			
5	,343	3,807	96,446			
6	,180	2,005	98,451			
7	,084	,933	99,384			
8	,041	,452	99,836			
9	,015	,164	100,000			

Método de extracción: Análisis de Componentes principales.

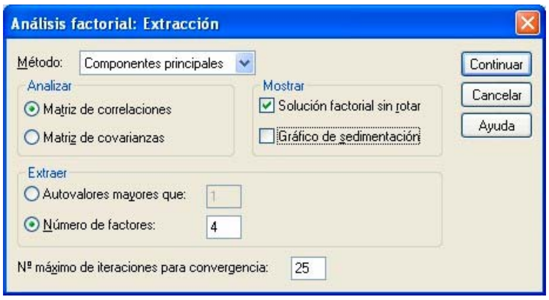
No obstante, el cuarto valor se encuentra muy próximo a 1, proporciona un factor que determina el 10,963% de la varianza, por lo que se decide incluirlo también en la estructura factorial.

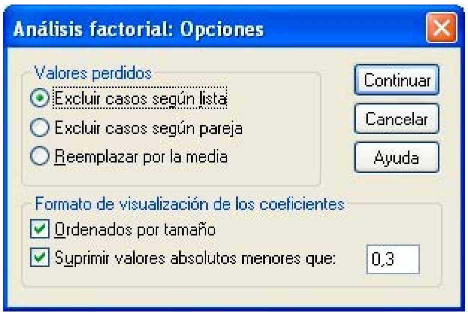
Finalmente, se elige una estructura factorial de cuatro factores que explicarían el 92,639% de la varianza. Esta decisión se observa también en el Gráfico de Sedimentación:



El análisis se enfoca en las Comunalidades (que muestran que porcentaje de cada variable es explicado por la nueva estructura factorial), y en la matriz de Componentes de la nueva estructura (eliminando los valores menores de 0,3).

Para ello, en el botón [Extracción] se eligen 4 factores





En el botón [Opciones] se elige *Ordenar las coeficientes por tamaño* y *Suprimir valores absolutos menores que* 0,3.

El Visor de SPSS presenta:

Comunalidades

	Inicial	Extracción
X1	1,000	,936
X2	1,000	,942
X3	1,000	,959
X4	1,000	,930
X5	1,000	,915
X6	1,000	,920
X7	1,000	,951
X8	1,000	,922
X9	1,000	,864

Método de extracción: Análisis de Componentes principales.

Comunalidades

	Inicial	Extracción
C_ELECTR	1,000	,936
BICARB	1,000	,942
CLORUROS	1,000	,959
SULFATOS	1,000	,930
CALCIO	1,000	,915
MAGNESIO	1,000	,920
SODIO	1,000	,951
POTASIO	1,000	,922
FOSFATOS	1,000	,864

Método de extracción: Análisis de Componentes principales.

Las Comunalidades son muy altas, lo que implica que todas las variables están muy bien representadas en el espacio de los factores (la Comunalidad representa el coeficiente de correlación lineal múltiple de cada variable con los factores).

Matriz de componentes[a]

	Componente			
	1	2	3	4
X6	,894			
X1	,868		,398	
X7	,858		,334	
X3	,795	-,318	,461	
X4	,733	,369	-,506	
X5	,705	,350	-,449	-,306
X9		,738	,536	
X8		,688	,488	-,455
X2		,578		,777

Método de extracción: Análisis de componentes principales.
a. 4 componentes extraídos

Matriz de componentes[a]

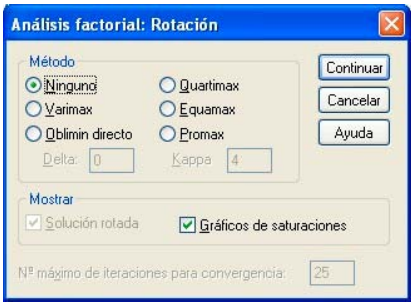
	Componente			
	1	2	3	4
MAGNESIO	,894			
C_ELECTR	,868		,398	
SODIO	,858		,334	
CLORUROS	,795	-,318	,461	
SULFATOS	,733	,369	-,506	
CALCIO	,705	,350	-,449	-,306
FOSFATOS		,738	,536	
POTASIO		,688	,488	-,455
BICARB		,578		,777

Método de extracción: Análisis de componentes principales.
a. 4 componentes extraídos

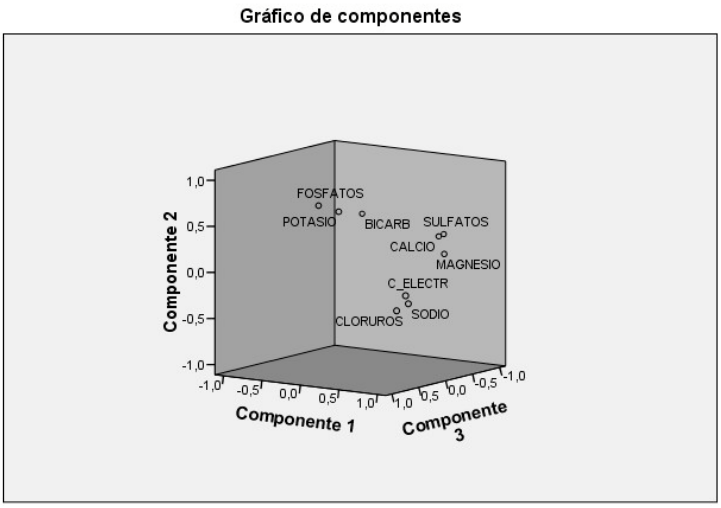
La estructura factorial no está muy clara en principio, ya que diversos factores comparten variables.

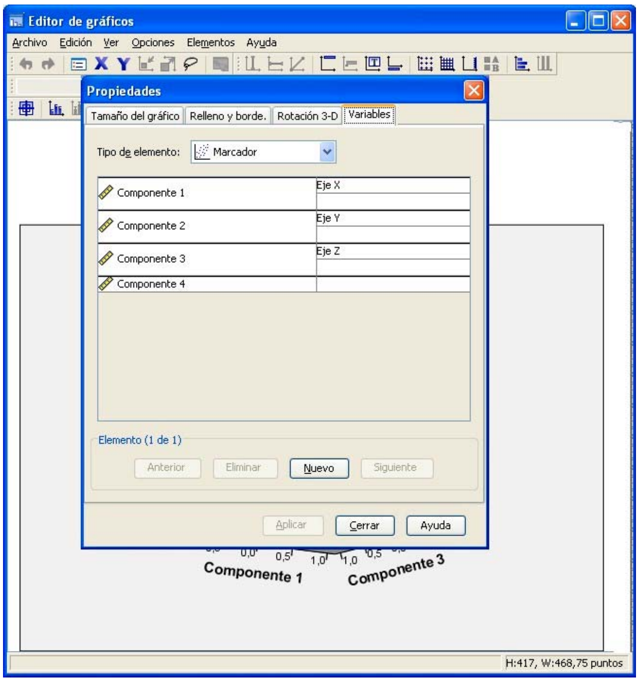
Por ejemplo, la variable Potasio (X8) está relacionada con los factores segundo, tercero y cuarto. La variable Fosfatos (X9) aparece tanto en el segundo factor como en el tercero. Lo mismo ocurre para las variables Sulfatos (X4) y Calcio (X5) respecto a los ejes primero y tercero.

Gráficamente se representan las variables en el plano de los factores (primero, tercero). Para ello, en el botón [Rotación] se elige la opción Gráficos de saturaciones.

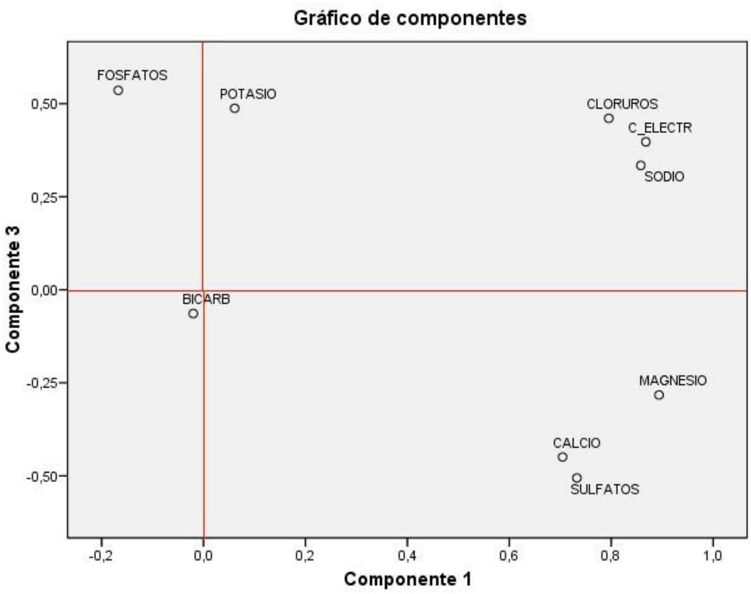


En el Visor de SPSS sale el Gráfico de componentes tridimensional de los factores:

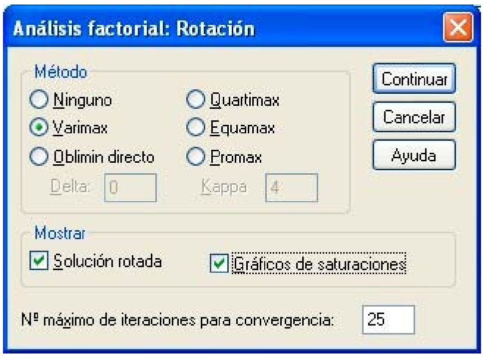




Haciendo dos click en el Gráfico, o bien con el botón de la izquierda del ratón seleccionando Objeto Gráfico de SPSS, se selecciona Propiedades, y se eligen las Variables que se desean representar.



Se observa que las dos variables (Sulfatos, Calcio) forman un ángulo próximo a 45° con cada eje, lo cual no permite asociarlas a ninguno de ellos (las saturaciones representan en este caso las correlaciones de las variables con cada eje y por lo tanto el coseno del ángulo que forman con ellos).
Al mismo tiempo, la variable Bicarbonato esta cerca del eje de coordenadas, indica que no está relacionada con ninguno de los dos ejes.



Con la idea de clarificar la estructura factorial sin perder poder explicativo, se realiza una rotación de ejes. Se elige el botón [Rotación] y el método Varimax (método de rotación ortogonal que minimiza el número de variables que tienen saturaciones altas en cada factor).
La interpretación simplifica de los factores optimando la solución por columna produce la siguiente matriz de componentes (las comunalidades no varían):

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3,986	44,291	44,291	3,986	44,291	44,291	2,922	32,471	32,471
2	1,840	20,448	64,739	1,840	20,448	64,739	2,638	29,308	61,779
3	1,549	17,207	81,946	1,549	17,207	81,946	1,633	18,144	79,923
4	,962	10,693	92,639	,962	10,693	92,639	1,144	12,716	92,639
5	,343	3,807	96,446						
6	,180	2,005	98,451						
7	,084	,933	99,384						
8	,041	,452	99,836						
9	,015	,164	100,000						

poder explicativo varianza total (ejes sin rotar)

poder explicativo varianza total (ejes rotados) VARIMAX

Método de extracción: Análisis de Componentes principales.

Matriz de componentes[a]

	Componente 1	2	3	4
MAGNESIO	,894			
C_ELECTR	,868		,398	
SODIO	,858		,334	
CLORUROS	,795	-,318	,461	
SULFATOS	,733	,369	-,506	
CALCIO	,705	,350	-,449	-,306
FOSFATOS		,738	,536	
POTASIO		,688	,488	-,455
BICARB		,578		,777

Método de extracción: Análisis de componentes principales.
a. 4 componentes extraídos

Matriz de componentes rotados[a]

	Componente 1	2	3	4
CLORUROS	,971			
SODIO	,944			
C_ELECTR	,917			
CALCIO		,939		
SULFATOS		,927		
MAGNESIO	,447	,845		
POTASIO			,947	
FOSFATOS			,834	,379
BICARB				,956

Método de extracción: Análisis de componentes principales.
Método de rotación: Normalización Varimax con Kaiser.
a. La rotación ha convergido en 5 iteraciones.

Factor 1: Asociado a las variables de Cloruros, Sodio, Conductividad Eléctrica y en menor proporción a Magnesio. Tiene un poder explicativo del 44,291% de la varianza total (el porcentaje de inercia se refiere a los ejes que se han obtenido en primer lugar y no tienen por qué coincidir con los porcentajes de inercia una vez rotados, aunque sí coincide con el total explicado, SSPS muestra el porcentaje en la rotación Varimax: 32,471%, 29,308%, 18,144% y 12,716%).

- La variable Conductividad Eléctrica queda explicada por el total de los factores en un 93,6% (Comunalidad 0,936), mientras que representa el 84,08% ($0,917^2$= 84,08%) de la varianza total, es decir, el 89,83% (0,8408/0,936 = 89,83%) del total del espacio de los factores.
- La estructura factorial completa determina a la variable Cloruros una varianza total de 94,28%, esto es, el 98,31% del total del espacio de los factores.
- La variable Sodio tiene una Comunalidad de 0,951, con un 95,1% de la varianza explicada (89,11% por este factor y 93,7% en el espacio de los factores).

- La variable Magnesio queda explicada por la estructura factorial en un 92%, con menos carga factorial que las anteriores (0,447), lo que representa casi el 20% de su varianza (21,71% en la estructura factorial).

Factor 2: Asociado a las variables Calcio, Sulfatos y Magnesio. Con un poder explicativo de 20,448% de inercia.

- La variable Magnesio, representada por una estructura factorial de 92% (Comunalidad de 0,92), está más representada por este factor, su saturación (carga factorial) es de 0,845, con lo que representa el $0,845^2 = 71,40\%$ de su varianza total, es decir el 77,61% (0,714/0,92 = 77,61%) de la explicada por todos los factores.

- La variable Sulfatos que tiene una Comunalidad de 0,93, una saturación de 0,927, es explicada por este eje con un 85,93% ($0,927^2 = 0,8593$), lo que es un 92,4% en el espacio de los factores (0,8593/0,93 = 92,397%).

- La variable Calcio, con una Comunalidad de 0,915 (representa el 91,5%), tiene una carga factorial de 0,939, por lo que el 88,17% de su varianza total [$0,939^2 = 88,17\%$] viene representada por este eje (96,36% de lo explicado por la estructura factorial total $\equiv$ 0,8817/0,915 = 96,36%)

Factor 3: Asociado a las variables Potasio y Fosfatos, con un porcentaje de inercia explicada del 17,207%. (18,144% con ejes rotados).

- La variable Potasio, con una Comunalidad de 0,922, y este factor aporta el 89,68%, es decir, un 97,27% de lo explicado por la estructura factorial.

- La variable Fosfatos está representada en el espacio de los factores por una Comunalidad de 0,864, que atribuible al tercer factor es el 65,55%, con una saturación de 0,834, esto es, el 80,5% del espacio de los factores.

Factor 4: Representado principalmente por la variable Bicarbonato, representada por una estructura factorial de 94,2% (Comunalidad de 0,942), tiene una carga factorial de 0,956. La varianza explicada por el factor es 91,39%, lo que equivale al 97,02% de lo determinado por los cuarto factores.

La estructura factorial ha quedado clarificada y solamente la variable Magnesio parece que comparte parte de su varianza con dos factores. El siguiente paso sería interpretar en términos geológicos el significado de los factores, o sea, intentar resumir el porqué se unen esas variables e incluso intentar dar un nombre a cada factor.

A partir de una matriz de correlaciones, el Análisis Factorial extrae otra matriz que reproduce la primera de forma más sencilla. Esta nueva matriz se denomina matriz factorial y adopta la siguiente forma:

	1	2
1	P_{11}	P_{21}
2	P_{12}	P_{22}
3	P_{13}	P_{23}
..	..	..
l	P_{1l}	P_{2l}

Cada columna es un factor y hay tantas filas como variables originales.

Los elementos P_{ij} pueden interpretarse como índices de correlación entre el factor i-ésimo y la variable j-ésima, aunque estrictamente sólo son correlaciones cuando los factores no están correlacionados entre sí, es decir, son ortogonales.

Estos coeficientes reciben el nombre de pesos, cargas, ponderaciones o saturaciones factoriales. Los pesos factoriales indican el peso de cada variable en cada factor. Lo ideal es que cada variable cargue alto en un factor y bajo en los demás.

Eigenvalues

El cuadrado de una carga factorial indica la proporción de la varianza explicada por un factor en una variable particular.

La suma de los cuadrados de los pesos de cualquier columna de la matriz factorial es lo que denominamos eigenvalues (λ), indica la cantidad total de varianza que explica ese factor para las variables consideradas como grupo.

Las cargas factoriales pueden tener como valor máximo 1, por tanto el valor máximo que puede alcanzar el valor propio es igual al número de variables.

Si dividimos el valor propio entre el número de variables nos indica la proporción de las varianza de las variables que explica el factor.

$$\begin{cases} \lambda_1 = P_{11}^2 + P_{12}^2 + P_{13}^2 + \cdots + P_{1j}^2 \\ \lambda_2 = P_{21}^2 + P_{22}^2 + P_{23}^2 + \cdots + P_{2j}^2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \lambda_i = P_{i1}^2 + P_{i2}^2 + P_{i3}^2 + \cdots + P_{ij}^2 \end{cases} \qquad \begin{array}{l} \dfrac{\lambda_1}{n} \equiv \text{varianza explicada por primer factor} \\ \dfrac{\lambda_2}{n} \equiv \text{varianza explicada por segundo factor} \end{array}$$

Comunalidades

Se denomina Comunalidad a la proporción de la varianza explicada por los factores comunes en una variable.

La Comunalidad (h) es la suma de los pesos factoriales al cuadrado en cada una de las filas.
El Análisis Factorial comienza sus cálculos a partir de lo que se conoce como matriz reducida compuesta por los coeficientes de correlación entre las variables y con las comunalidades en la diagonal.

Como la comunalidad no se puede saber hasta que se conocen los factores, este resulta ser uno de los problemas del Análisis Factorial.

En el Análisis de Componentes Principales no se supone la existencia de ningún factor común la comunalidad toma como valor inicial 1.

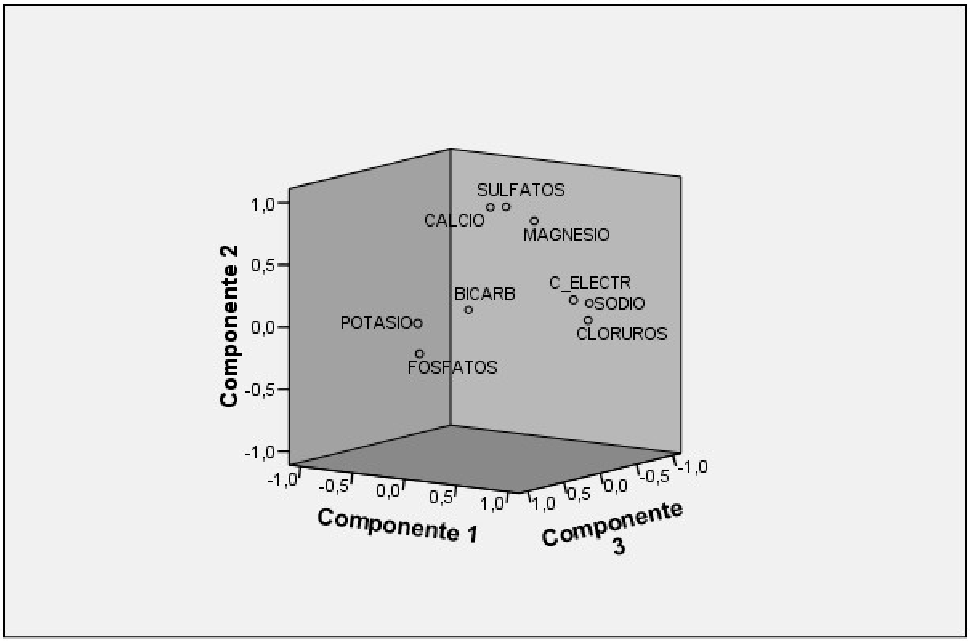
En los otros métodos se utilizan diferentes modos de estimar la comunalidad inicial:

- Estimando la comunalidad por la mayor correlación en la fila i-ésima de la matriz de correlaciones.
- Estimando la comunalidad por el cuadrado del coeficiente de correlación múltiple entre x y las demás variables. (SPSS por defecto).
- El promedio de los coeficientes de correlación de una variable con todas las demás.
- Calculando a partir de los dos coeficientes de correlación mayores de esa variable la siguiente operación: $h^2 = \dfrac{r_{xy}\, r_{xz}}{r_{yz}}$

La comunalidad final de cada variable viene dada por: $h^2 = P_{1j}^2 + P_{2j}^2 + \cdots + P_{kj}^2$

La Gráfica tridimensional de las variables en el espacio de los factores permiten visualizar la estructura factorial matriz de las cargas factoriales correspondientes a los factores

Gráfico de componentes en espacio rotado



SPSS 2.- Se realiza un análisis factorial de todas las variables del fichero ratios.sav que contiene ratios relativos a las ventas de las empresas españolas con mayores ventas : Beneficios/Recursos propios (R1), Cash-Flow/Ventas (R2), Inmovilizado/Activos Totales (R3), Ventas/Activos Totales (R4), Ventas/Plantilla (R5), Beneficios/Capital Social (R6) y Beneficios/Ventas (R7)

El objetivo es resumir estos ratios por un número menor de factores con mínima pérdida de información que tengan la suficiente calidad para seguir agrupando a las empresas según sus ventas.
Se quiere analizar si es coherente identificar tres factores (un factor financiero, un factor estructural y un factor de rentabilidad).

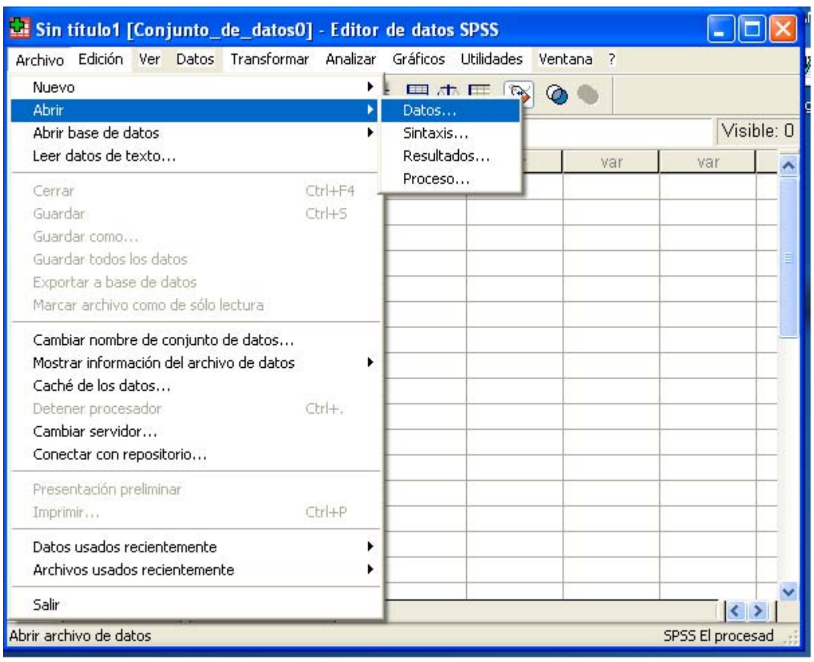
Antes de realizar el análisis factorial se plantea si las 7 variables originales están correlacionadas entre sí o no lo están. En caso de no estar no existen factores comunes y, por lo tanto, no tiene sentido aplicar el análisis factorial.

Esta cuestión suele probarse utilizando el contraste de esfericidad de Barlett que se basa en que la matriz de correlación poblacional R recoge la relación entre cada par de variables mediante sus elementos ρ_{ij} situados fuera de la diagonal principal.

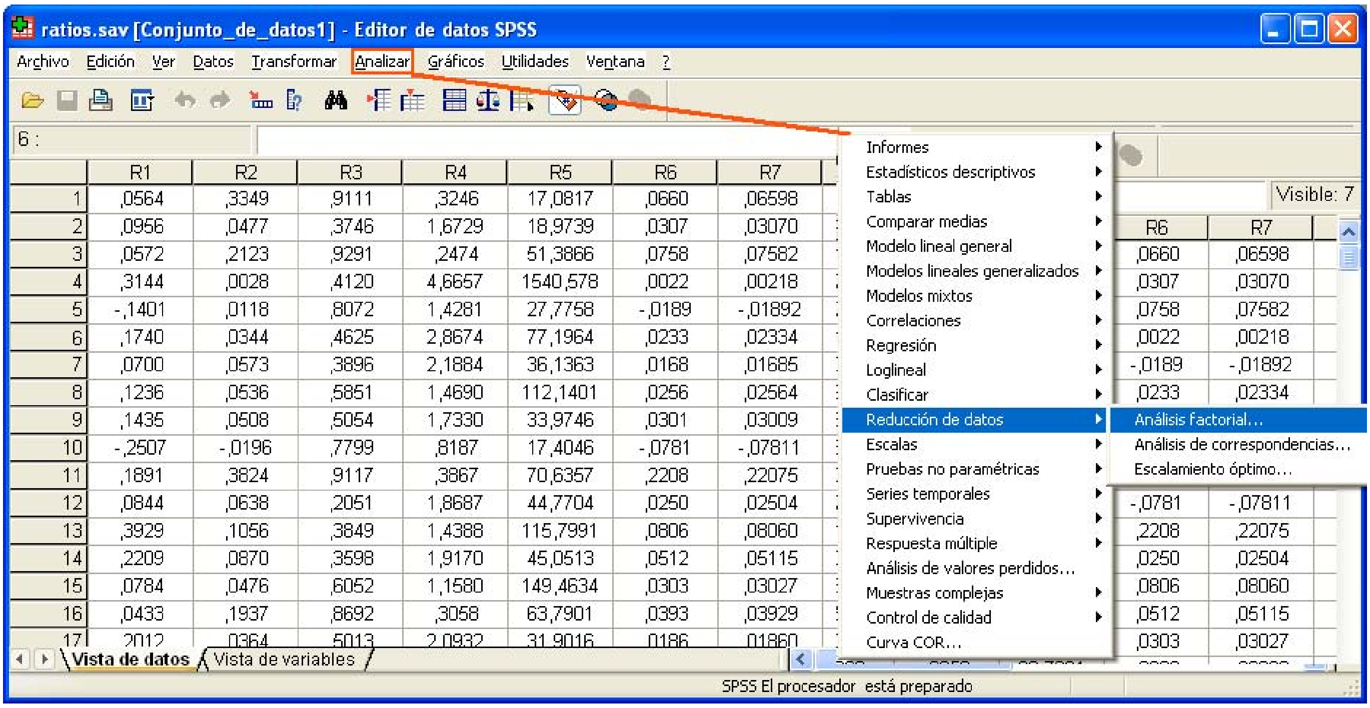
Los elementos de la diagonal principal son unos, ya que toda variable está totalmente relacionada consigo misma.
En consecuencia, para decidir la ausencia o no de la relación ente las 7 variables puede plantearse el contraste: $H_0 : |R| = 1 \quad H_1 : |R| \neq 1$

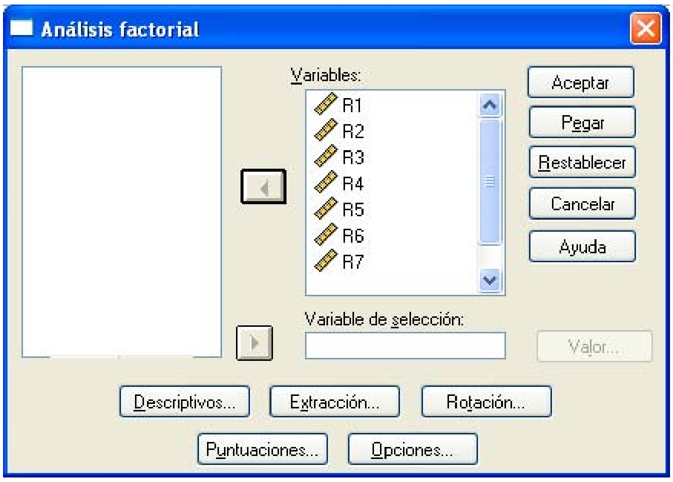
Por ora parte, Kayser-Meyer y Olkin define la medida KMO de adecuación muestral global al modelo factorial basa en los coeficientes de correlación observados de cada para de variables y en sus coeficientes de correlación parcial. Cuando los valores de KMO están por debajo de 0,5 no serán aceptabes, considerando inadecuados los datos a un modelo de análisis factorial.

Se carga el fichero: Archivo / Abrir / Datos / ratios.sav



Analizar / Reducción de datos / Análisis Factorial

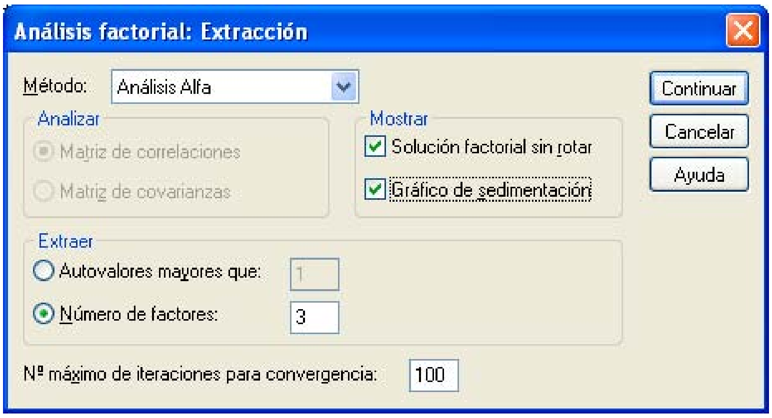




Se incluyen las siete variables en el análisis.

Después de seleccionar las opciones de los botones Descriptivos, Extracción, Puntuaciones factoriales y Rotación se pulsa Aceptar para obtener la salida del procedimiento.

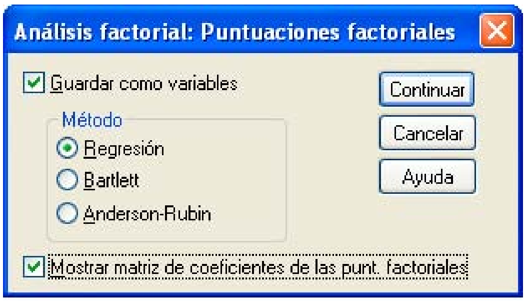
Se elige la Solución inicial, el determinante de la matriz de correlaciones y el estadístico KMO y la prueba de esfericidad de Barllet.



Para extraer los factores se elige el método Alfa, se extraen los tres primeros factores y se muestran la Solución inicial sin rotar y el Gráfico de sedimentación.

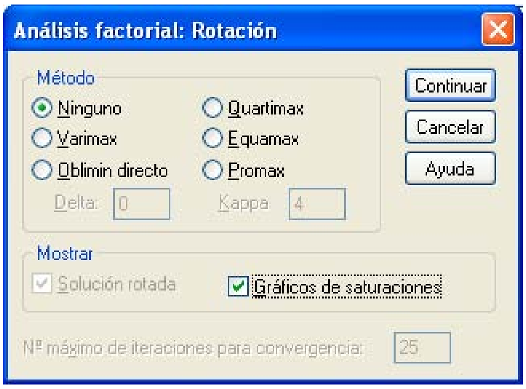
Alfa es un método de extracción factorial que considera a las variables incluidas en el análisis como una muestra del universo de las variables posibles. Este método maximiza el Alfa de Cronbach para los factores.

El Gráfico de Sedimentación es una representación gráfica donde los factores están en el eje de abscisas y los valores propios en el de ordenadas. Los factores con varianzas altas suelen diferenciarse de los factores con varianzas bajas. Se pueden conservar los factores situados antes de este punto de inflexión. En simulaciones el criterio ha funcionado bien, tiene el inconveniente de que depende del 'ojo' del analista.



Las puntuaciones factoriales se guardarán como variables, se muestra su matriz de coeficientes y se calculan mediante el método de Regresión.

El método de Regresión da lugar a puntuaciones con máxima correlación con las puntuaciones teóricas. Sin embargo, el estimador no es insesgado, ni unívoco y, en caso de que los factores sean ortogonales, puede dar lugar a puntuaciones correladas.



Inicialmente no se realiza rotación y se muestran los Gráficos de saturaciones.

Matriz de correlaciones[a]

a. Determinante = 1,26E-009

El determinante de la matriz de correlaciones es muy bajo.

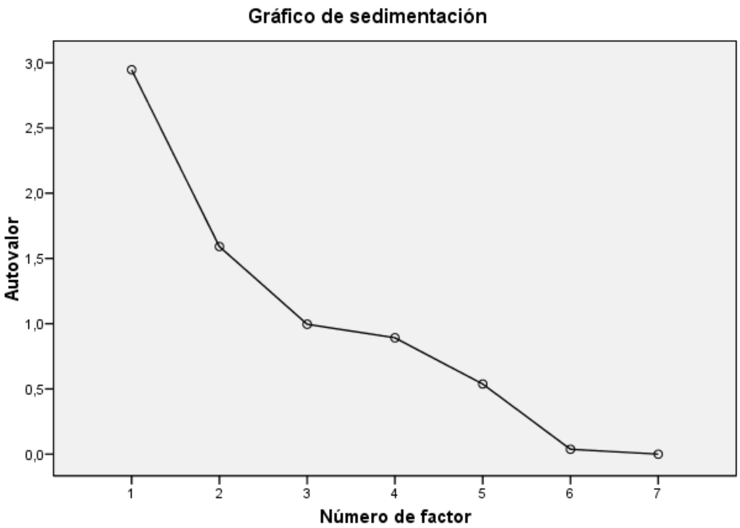
El estadístico KMO es aceptable $(0,676 > 0,5)$.

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,676
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	4565,538
	gl	21
	Sig.	,000

La Prueba de esfericidad de Barlett tiene un p-valor $= 0,000 < 0,05$ rechazando la hipótesis nula de ausencia de relación entre las variables.

Hay una adecuación muestral alta de los datos para el análisis factorial.



El Gráfico de sedimentación tiene sólo dos valores propios mayores que uno con un tercero muy próximo.

Varianza total explicada

Factor	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,946	42,092	42,092	1,636	23,370	23,370
2	1,591	22,727	64,819	2,381	34,014	57,384
3	,996	14,223	79,042	,125	1,790	59,174
4	,892	12,740	91,782			
5	,538	7,682	99,464			
6	,038	,536	100,000			
7	1,51E-008	2,15E-007	100,000			

Método de extracción: Factorización Alfa.

Los dos primeros factores explican el 64,819% de la varianza.

Los tres primeros factores explican el 79,042% de la varianza.

Una posición conservadora tomaría los tres primeros factores.

Por esta razón, en la Extracción de Factores se ha desechado la opción Autovalores mayores que 1 y se ha elegido Número de factores igual a 3.

Comunalidades

	Inicial	Extracción
R1	,004	,004
R2	,944	,964
R3	,420	,459
R4	,292	,601
R5	,093	,124
R6	1,000	,995
R7	1,000	,995

Método de extracción: Factorización Alfa.

La Comunalidad es la parte de variabilidad de cada variable explicada por los factores.

Antes de la extracción de los factores la Comunalidad de cada variable debe de ser alta, e interesa que después de la extracción siga siendo alta.

Matriz factorial[a]

	Factor		
	1	2	3
R1	,054	-,025	-,020
R2	,389	,901	-,012
R3	-,512	,381	,226
R4	,659	-,385	,139
R5	,280	-,056	,207
R6	,595	,797	-,075
R7	,595	,797	-,075

Método de extracción: Factorización Alfa.

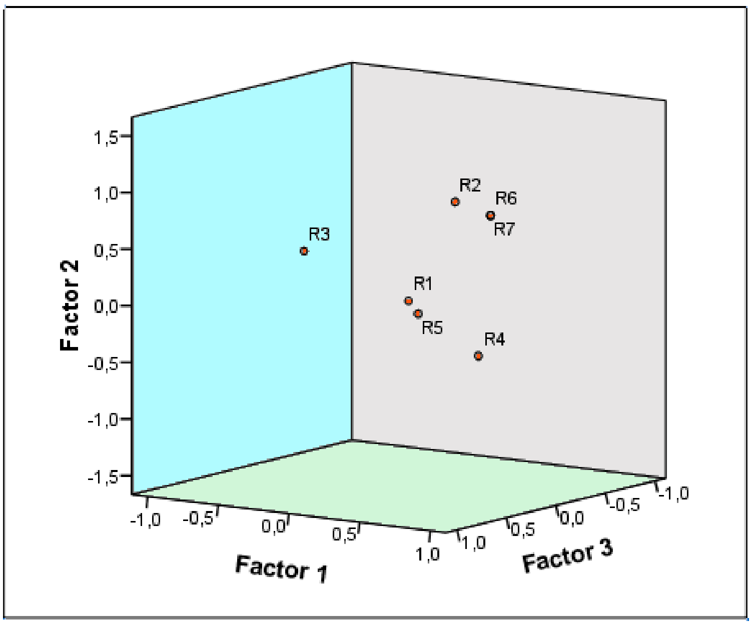
a. 3 factores extraídos. Requeridas 10 iteraciones.

Cargas factoriales altas en valor absoluto de una variable sobre un factor indican que hay mucho en común entre la variable y el factor.

Hay investigadores que sostienen que cargas mayores que 0,6 asocian a la variable con el factor, mientras que otros sostienen que es suficiente un valor superior a 0,4

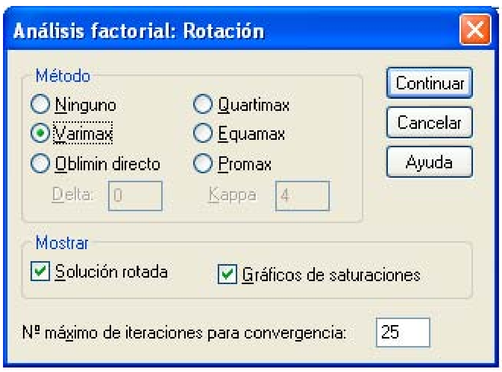
En este caso, no hay forma clara de asociar las variables a los factores, por lo que se hace una rotación.

Gráfico de factores



Por otra parte, el Gráfico tridimensional de factores no despeja las dudas.

Se realiza una rotación de los factores por el método de Varimax.



Es un método de rotación que minimiza el número de variables con cargas altas en un factor, mejorando así la interpretación de factores.

El método considera que, si se logra aumentar la varianza de las cargas factoriales al cuadrado de cada factor consiguiendo que algunas de sus cargas factoriales tiendan a acercarse a 1 mientras que otras se aproximan a 0, se obtiene una pertenencia más clara e inteligible de cada variable al factor.

Los nuevos ejes se obtienen maximizando la suma para los 7-factores retenidos de las varianzas de las cargas factoriales al cuadrado dentro de cada factor.

Para evitar que las variables con mayores comunalidades tengan más peso en la solución final, se efectúa la normalización de Kaiser (dividiendo cada carga factorial al cuadrado por la comunalidad de la variable correspondiente).

Matriz de factores rotados[a]

	Factor		
	1	2	3
R1	,007	,060	,015
R2	,971	-,143	,029
R3	,052	-,661	-,136
R4	-,025	,575	,519
R5	,065	,116	,327
R6	,989	,084	,099
R7	,989	,084	,098

Método de extracción: Factorización Alfa.
Método de rotación: Normalización Varimax con Kaiser.
a. La rotación ha convergido en 5 iteraciones.

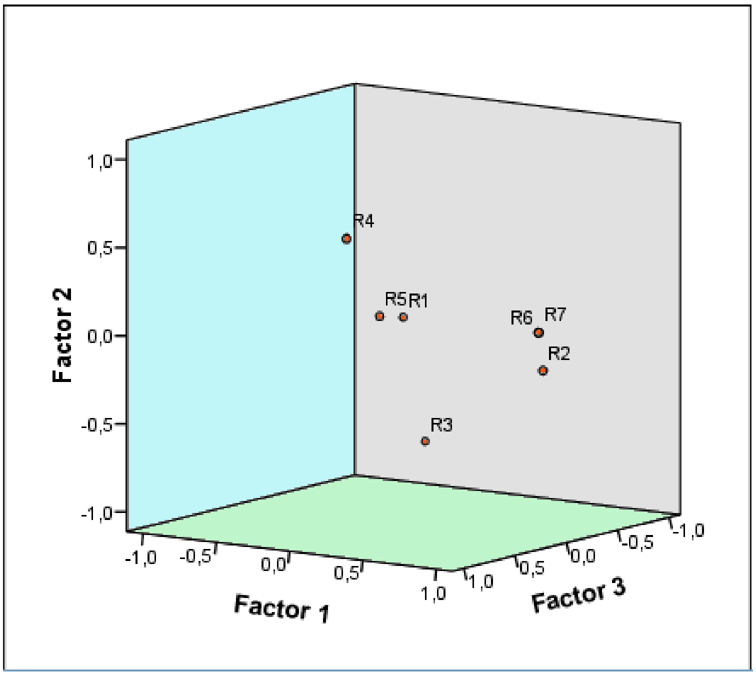
La matriz de los factores rotados muestra que al primer factor se asocian las variables R2, R6 y R7 con cargas factoriales mayores que 0,9.

Al segundo factor se asocia la variable R3

Al tercer factor se asocian las variables R4 y R5. También podría haberse asociado la variable R4 al segundo factor, pero está asociación está más clara para el tercer factor, ya que es su única carga realmente alta.

Queda suelta la variable R1, se la asocia al segundo factor para el que presenta mayor carga.

Gráfico de saturaciones en espacio factorial rotado



Con el Gráfico de saturaciones en espacio factorial rotado se llega a las mismas conclusiones que con la matriz de factores rotados.

El primer factor (R2, R6 y R7) es un factor financiero relativo a la distribución de los Beneficios y Flujo de caja.

El segundo factor (R1 y R3) es un factor estructural relativo a Recursos propios, Inmovilizado y Activos totales.

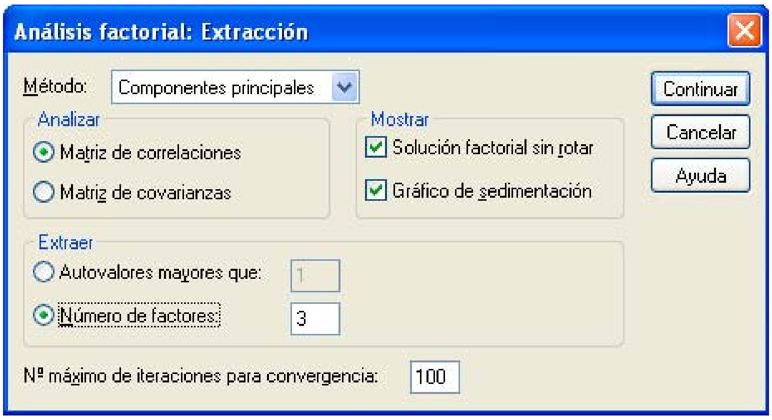
El tercer factor (R4 y R5) es una factor de rentabilidad relativo a la distribución de las Ventas.

Los datos han de tener una distribución normal bivariada para cada par de variables, y las observaciones han de ser independientes.

El modelo de análisis factorial especifica que las variables vienen determinadas por los factores comunes (factores estimados por el modelo) y por factores únicos (no se superponen entre las distintas variables observadas).

Las estimaciones calculadas se basan en el supuesto de que ningún factor único está correlacionado con los demás, ni con los factores comunes.

Se puede resolver el ejercicio mediante un análisis de Componentes principales. Para ello, en el botón de Extracción se introduce el método de Componentes principales.



Para extraer los factores se elige el método Componentes principales, se extraen los tres primeros factores y se muestran la Solución inicial sin rotar y el Gráfico de sedimentación.

Es un método de extracción de factores utilizado para formar combinaciones lineales no correlacionadas de las variables observadas.

La primera componente tiene la varianza máxima. Las componentes sucesivas explican progresivamente proporciones menores de la varianza y no están correlacionadas las unas con las otras.
El análisis de componentes principales se utiliza para obtener la solución factorial inicial.
Puede utilizarse cuando una matriz de correlaciones es singular.

Matriz de componentes[a]

	Componente		
	1	2	3
R1	,014	,128	,984
R2	,974	-,162	,008
R3	,035	-,785	,010
R4	,038	,829	-,046
R5	,140	,491	-,160
R6	,994	,042	,001
R7	,994	,042	,001

Método de extracción: Análisis de componentes principales.
a. 3 componentes extraídos

La primera componente está formada por (R2, R6 y R7) resultando ser la componente financiera relativa a la distribución de Beneficios y Flujo de caja.

La segunda componente está formada por (R3, R4 y R5) y es una componente de rentabilidad relativa especialmente a la distribución de Ventas.

La tercera componente está formada por R1 que puede interpretarse como una componente estructural relativa a Recursos propios.

Se observa que la solución por componentes principales es más fácil de obtener (sin necesidad de rotar) y es más fácil de interpretar ya que la matriz de componentes discrimina mejor que la matriz factorial.

ELEMENTOS DE ALGEBRA LINEAL

Matriz

Una matriz de orden o dimensión $n \times p$ es una ordenación rectangular de elementos dispuestos en n filas y p columnas de la siguiente forma:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

En general, para designar a una matriz se utiliza una letra mayúscula en negrita. Un elemento genérico de la matriz A se designa mediante a_{ij}, donde el primer subíndice i hace referencia a la fila en que está situado el elemento, mientras que el segundo subíndice j hace referencia a la columna.

Una matriz de orden 1×1 es un escalar.

Ejemplos: $A = \begin{bmatrix} 1 & 3 & 4 \\ 2 & -5 & 6 \end{bmatrix}$ $\quad B = \begin{bmatrix} 1 & 3 & 6 \\ 2 & 4 & 5 \\ -5 & 2 & 8 \end{bmatrix}$

Matriz transpuesta

La transpuesta de una matriz A de orden $n \times p$ es una matriz B de orden $p \times n$, obtenida mediante el intercambio de filas y columnas, de forma que $a_{ij} = b_{ji}$.

En general, a la matriz transpuesta de A se denomina A'

Ejemplos de traspuestas de las matrices anteriores:

$$\mathbf{A'} = \begin{bmatrix} \mathbf{1} & \mathbf{2} \\ \mathbf{3} & \mathbf{-5} \\ \mathbf{4} & \mathbf{6} \end{bmatrix} \qquad B' = \begin{bmatrix} 1 & 2 & -5 \\ 3 & 4 & 2 \\ 6 & 5 & 8 \end{bmatrix}$$

Vector columna y vector fila

Un vector columna de orden n es una ordenación de elementos dispuestos en n filas y 1 columna de la siguiente forma:

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

En general, para designar a un vector columna se utiliza una letra minúscula en negrita.

Un vector fila de orden n es una ordenación de elementos dispuestos en 1 fila y n columnas de la siguiente forma: $b = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}$

El transpuesto de un vector fila es un vector columna: $a' = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$

Matriz cuadrada

Se dice que una matriz es cuadrada si el número de filas es igual al número de columnas. Se dice que una matriz cuadrada es de orden n si tiene n filas.

Ejemplo: $C = \begin{bmatrix} 1 & 2 & 2 \\ 4 & 3 & 7 \\ 5 & 6 & 8 \end{bmatrix}$

Traza de una matriz

En una matriz cuadrada de orden n la diagonal principal está formada por los elementos a_{ii} $(i = 1, 2, \cdots, n)$. La traza de una matriz cuadrada A, a la que se designa por $tr(A)$ o por $traza(A)$, es la suma de los elementos de la diagonal principal. Es decir:

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

Ejemplo: $tr(C) = 1 + 3 + 8 = 12$

Matriz simétrica

Se dice que una matriz cuadrada es simétrica si se verifica que $A = A'$

Ejemplo: $D = \begin{bmatrix} 3 & 4 & 1 \\ 4 & 3 & 5 \\ 1 & 5 & 8 \end{bmatrix}$

Matriz diagonal

Se dice que una matriz cuadrada es diagonal cuando todos los elementos situados fuera de la diagonal principal son nulos. Es decir, en una matriz diagonal se verifica que $a_{ij} = 0$ cuando $i \neq j$.

La matriz diagonal es de la forma:

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & a_{np} \end{bmatrix}$$

Matriz escalar

Se dice que una matriz diagonal es escalar cuando todos los elementos de la diagonal principal son idénticos. Es decir, en una matriz escalar se verifica que $a_{ii} = k$ para todo i.

Matriz identidad

Es una matriz escalar en la que $a_{ii} = 1$, se le denomina $\mathbf{I}$. Una matriz identidad genérica tiene la forma:

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

OPERACIONES CON MATRICES

Igualdad de matrices

La igualdad de dos matrices $A = B$ se cumple si, y sólo si, A y B son del mismo orden y $a_{ij} = b_{ij}$ para todo i y todo j. Para realizar la suma, las matrices A y B deben de ser del mismo orden.

Suma de matrices

La suma de las matrices A y B de orden $n \times p$ es igual a otra matriz C, también de orden $n \times p$, definida de la siguiente forma: $A + B = C$

Los elementos de la matriz C se obtienen: $c_{ij} = a_{ij} + b_{ij} \quad \forall\, i, j$

Ejemplo: $A = \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 4 \\ -2 & 6 \end{bmatrix} \qquad C = \begin{bmatrix} 2+1 & -1+4 \\ 3-2 & 5+6 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 1 & 11 \end{bmatrix}$

Multiplicación escalar

La multiplicación escalar de una matriz A por un escalar λ se efectúa multiplicando cada elemento de A por λ. El producto es designado por λA

Ejemplo: $\lambda = 4 \qquad A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} \qquad \lambda A = 4\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 12 \\ 16 & 4 \end{bmatrix}$

Multiplicación de matrices

Si A es una matriz de orden $n \times m$ y B es una matriz de orden $m \times p$, el producto de estas dos matrices es otra matriz C de orden $n \times p$.

$$AB = C \quad \rightarrow \quad \text{Elemento genérico } c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} (a_{11}b_{11} + a_{12}b_{21}) & (a_{11}b_{12} + a_{12}b_{22}) & (a_{11}b_{13} + a_{12}b_{23}) \\ (a_{21}b_{11} + a_{22}b_{21}) & (a_{21}b_{12} + a_{22}b_{22}) & (a_{21}b_{13} + a_{22}b_{23}) \end{bmatrix}$$

$$\begin{bmatrix} 4 & 2 \\ 3 & 5 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 1 & 9 & 1 \\ 7 & -2 & 6 \end{bmatrix} = \begin{bmatrix} 4x1+2x7 & 4x9-2x2 & 4x1+2x6 \\ 3x1+5x7 & 3x9-5x2 & 3x1+5x6 \\ 2x1+6x7 & 2x9-6x2 & 2x1+6x6 \end{bmatrix} = \begin{bmatrix} 18 & 32 & 16 \\ 38 & 17 & 33 \\ 44 & 6 & 38 \end{bmatrix}$$

Determinante de una matriz

El determinante de una matriz cuadrada A, al que se designa por $|A|$, es un escalar que se obtiene por la suma de $n!$ términos, cada uno de los cuales es el producto de n elementos. Se obtiene mediante la fórmula:

$$|A| = \sum \pm a_{1j}\, a_{2k} \cdots a_{nq}$$

En la expresión anterior cada sumando se obtiene permutando el segundo subíndice. El signo de cada sumando es $+$ o $-$ según que el número de permutaciones realizado a partir de la ordenación original sea par o impar.

Si $|A| = 0$ se dice que u_j u_j la matriz A es singular.

Ejemplos: $|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\,a_{22} - a_{12}\,a_{21}$

$$|B| = \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} = b_{11}\,b_{22}\,b_{33} + b_{21}\,b_{32}\,b_{13} + b_{31}\,b_{23}\,b_{12} - b_{13}\,b_{22}\,b_{31} - b_{23}\,b_{32}\,b_{11} - b_{33}\,b_{21}\,b_{12}$$

Propiedades de los determinantes

(a) El determinante de una matriz cuadrada es igual al determinante de su traspuesta: $|A| = |A'|$

(b) El determinante del producto de matrices cuadradas es igual al producto de los determinantes de cada una de las matrices: $|A\,B\,C| = |A|\,|B|\,|C|$

(c) Si se multiplica una matriz A de orden n por una constante h, se verifica: $|h\,A| = h^n\,|A|$

(d) Si una matriz A tiene dos filas, o dos columnas, idénticas o proporcionales, entonces $|A| = 0$

Matriz inversa

La inversa de una matriz cuadrada A es una matriz B que verifica: $A\,B = B\,A = I$

En general, a la matriz B que cumple esta propiedad se le designa por A^{-1}, con lo cual:

$$\mathbf{A\,A^{-1} = A^{-1}\,A = I}$$

Ejemplo: $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$ $\quad A' = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$ $\quad |A| = |A'| = 10$

$$A^{-1} = \frac{1}{10}\begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 3/10 & -1/10 \\ -2/10 & 4/10 \end{bmatrix}$$

$$A^{-1}\,A = \frac{1}{10}\begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$A\,A^{-1} = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}\frac{1}{10}\begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} = I$$

Propiedades de las matrices

(a) La inversa de un producto de matrices ABC es igual a $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$

(b) La traspuesta de una inversa es igual a la inversa de la traspuesta: $(A^{-1})' = (A')^{-1}$

(c) El determinante de la inversa de una matriz es igual al recíproco del determinante de la matriz original, es decir: $\left|A^{-1}\right| = \dfrac{1}{|A|}$

RAÍCES Y VECTORES PROPIOS

Determinación de las raíces y vectores característicos

Se plantea el problema de la determinación de unos escalares (λ_j) y de unos vectores (u_j) tales que satisfagan la ecuación $Au_j = \lambda_j u_j$, donde A es una matriz cuadrada.

A los escalares (λ_j) que satisfacen la ecuación se les denomina raíces características, y a los correspondientes vectores (u_j) se les denomina vectores característicos.

Para las raíces características se utilizan también las denominaciones de valores propios o autovalores. Para los vectores característicos se utiliza alternativamente la expresión de vectores propios.

La ecuación $Au_j = \lambda_j u_j$ también se puede escribir como $(A - \lambda_j I)\, u_j = 0$

Las soluciones de la ecuación $(A - \lambda_j I)\, u_j = 0 \begin{cases} u_j = 0 \ \text{trivial} \\ \left|A - \lambda_j I\right| = 0 \ \text{ecuación característica de A} \end{cases}$

Resolviendo $\left|A - \lambda_j I\right| = 0$ se hallan n raíces características λ_j

A cada raíz característica (autovalor) λ_j va asociado un vector característico u_j

Cada vector característico (vector propio) u_j puede multiplicarse arbitrariamente por una constante sin afectar al resultado, dado que la matriz $(A - \lambda_j I)$ es singular.

En muchas aplicaciones, para soslayar la arbitrariedad del resultado, se procede a normalizar cada vector característico imponiendo la condición: $u_j\, u_j^t = 1$

De todas formas, aun después de normalizar subsiste una arbitrariedad en el signo, de forma que si u_j es una solución, $(-u_j)$ también lo es.

Es conveniente en muchas aplicaciones definir una matriz U en la que cada columna es un vector característico u_j. Por tanto:

$$U = \begin{bmatrix} u_1 & u_2 & \cdots & u_j & \cdots & u_n \end{bmatrix} = \begin{bmatrix} u_{11} & u_{21} & \cdots & u_{j1} & \cdots & u_{n1} \\ u_{12} & u_{22} & \cdots & u_{j2} & \cdots & u_{n2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ u_{1n} & u_{2n} & \cdots & u_{jn} & \cdots & u_{nn} \end{bmatrix}$$

Propiedades de las raíces y vectores característicos

(a) Las raíces características de una matriz diagonal son los elementos de la diagonal.

(b) Las matrices A y A' tienen las mismas raíces características, pero no necesariamente los mismos vectores característicos.

(c) Si λ es una raíz característica de A, entonces $1/\lambda$ es una raíz característica de A^{-1}

(d) Designando a las raíces características de A por $\lambda_1, \lambda_2, \cdots, \lambda_n$, se verifica:

$$\mathrm{tr}(A) = \sum_{j=1}^{n} \lambda_j \qquad |A| = \prod_{j=1}^{n} \lambda_j$$

(e) Una matriz A real y simétrica de orden n, da lugar a n vectores que son ortogonales entre sí, propiedad relevante en el Análisis Multivariante.

Se dice que los vectores u_j y u_h son ortogonales si $u_j u_h = 0$

Un conjunto de vectores se dice que son ortonormales, si además de ser ortogonales están normalizados, es decir, $u_j\, u_j^t = 1$

La matriz U formada por vectores característicos normalizados de una matriz simétrica, es decir, por vectores normalizados, se denomina ortonormal y verifica $U\,U' = I$

📖 1 Ejemplo: Sea la matriz $A = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$

Aplicando la ecuación característica $(A - \lambda_j I)\, u_j = 0$: $\left(\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} - \lambda_j \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} U_{1j} \\ U_{2j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Raíces características o autovalores: $|A - \lambda_j I| = 0$: $\left| \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} - \lambda_j \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\begin{vmatrix} 4-\lambda_j & 3 \\ 1 & 2-\lambda_j \end{vmatrix} = 0 \;\rightarrow\; \lambda_j^2 - 6\lambda_j + 5 = 0 \;\rightarrow\; \begin{cases} \lambda_1 = 5 \\ \lambda_2 = 1 \end{cases}$$

Vectores característicos o vectores propios:

- $\lambda_1 = 5 \quad u_1 = \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix}$

$$\left(\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} - 5 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \;\rightarrow\; \begin{bmatrix} -1 & 3 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \;\Rightarrow\; \begin{cases} -U_{11} + 3\,U_{21} = 0 \\ U_{11} - 3\,U_{21} = 0 \end{cases}$$

Las dos ecuaciones que se obtienen son proporcionales debido al carácter singular de la matriz $(A - \lambda_j I)$.

Por este motivo, con estas ecuaciones sólo se puede determinar la relación entre U_{11} y U_{21}, pero no valores concretos. La relación es $U_{11} - 3\,U_{21} = 0$

Con la ecuación de normalización $UU' = I$, es decir, $\begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} \begin{bmatrix} U_{11} & U_{21} \end{bmatrix} = U_{11}^2 + U_{21}^2 = 1$

Se tiene el sistema: $\begin{cases} U_{11} - 3U_{21} = 0 \\ U_{11}^2 + U_{21}^2 = 1 \end{cases} \rightarrow \begin{array}{c} (3U_{21})^2 + U_{21}^2 = 1 \\ U_{21} = 1/\sqrt{10} \end{array} \quad U_{11} = 3/\sqrt{10}$

Un vector característico (vector propio) será $u_1 = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$

Obsérvese que $(-u_1)$ también es solución del sistema anterior.

- $\lambda_2 = 1 \quad u_2 = \begin{bmatrix} U_{12} \\ U_{22} \end{bmatrix}$

$$\left(\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} - 1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} U_{12} \\ U_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} U_{12} \\ U_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} 3U_{12} + 3U_{22} = 0 \\ U_{12} + U_{22} = 0 \end{cases}$$

Con la ecuación de normalización $U_{12}^2 + U_{22}^2 = 1$

Se forma el sistema: $\begin{cases} U_{12} + U_{22} = 0 \\ U_{12}^2 + U_{22}^2 = 1 \end{cases} \rightarrow \begin{array}{c} (U_{12})^2 + (-U_{12})^2 = 1 \\ U_{12} = 1/\sqrt{2} \end{array} \quad U_{22} = -1/\sqrt{2}$

Un vector característico (vector propio) será $u_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

- La matriz U (cada columna es un vector característico) será: $U = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{10}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$

📖 2 Cálculo de raíces y vectores característicos de una matriz simétrica $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

Aplicando la ecuación característica $(A - \lambda_j I)\, u_j = 0$: $\left(\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} - \lambda_j \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} U_{1j} \\ U_{2j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Cálculo de las raíces características o autovalores: $|A - \lambda_j I| = 0$: $\left| \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} - \lambda_j \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\begin{vmatrix} 1-\lambda_j & 2 \\ 2 & 1-\lambda_j \end{vmatrix} = 0 \rightarrow \lambda_j^2 - 2\lambda_j - 3 = 0 \rightarrow \begin{cases} \lambda_1 = 3 \\ \lambda_2 = -1 \end{cases}$$

Cálculo de los vectores característicos o vectores propios:

- $\lambda_1 = 3 \quad u_1 = \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix}$

$$\left(\begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix} - 3\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix}U_{11}\\ U_{21}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \;\rightarrow\; \begin{bmatrix}-2 & 2\\ 2 & -2\end{bmatrix}\begin{bmatrix}U_{11}\\ U_{21}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \;\Rightarrow\; \begin{cases}-2U_{11} + 2U_{21} = 0\\ \;\;2U_{11} - 2U_{21} = 0\end{cases}$$

Con la ecuación de normalización $UU' = I$, es decir, $\begin{bmatrix}U_{11}\\ U_{21}\end{bmatrix}\begin{bmatrix}U_{11} & U_{21}\end{bmatrix} = U_{11}^2 + U_{21}^2 = 1$

Se tiene el sistema: $\begin{cases}2U_{11} - 2U_{21} = 0\\ U_{11}^2 + U_{21}^2 = 1\end{cases} \;\rightarrow\; \begin{matrix}U_{11}^2 + U_{11}^2 = 1\\ U_{11} = 1/\sqrt{2}\end{matrix} \qquad U_{21} = 1/\sqrt{2}$

Un vector característico (vector propio) será $u_1 = \begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$

- $\lambda_2 = -1 \quad u_2 = \begin{bmatrix}U_{12}\\ U_{22}\end{bmatrix}$

$$\left(\begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix} + 1\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix}U_{12}\\ U_{22}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \;\rightarrow\; \begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix}\begin{bmatrix}U_{12}\\ U_{22}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \;\Rightarrow\; \begin{cases}2U_{12} + 2U_{22} = 0\\ 2U_{12} + 2U_{22} = 0\end{cases}$$

Con la ecuación de normalización $UU' = I$, es decir, $\begin{bmatrix}U_{11}\\ U_{21}\end{bmatrix}\begin{bmatrix}U_{11} & U_{21}\end{bmatrix} = U_{11}^2 + U_{21}^2 = 1$

Se tiene el sistema: $\begin{cases}2U_{12} + 2U_{22} = 0\\ U_{12}^2 + U_{22}^2 = 1\end{cases} \;\rightarrow\; \begin{matrix}(-U_{22})^2 + U_{22}^2 = 1\\ U_{22} = 1/\sqrt{2}\end{matrix} \qquad U_{12} = -1/\sqrt{2}$

Un vector característico (vector propio) será $u_2 = \begin{bmatrix}-1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$

- La matriz U (cada columna es un vector característico) será: $U = \begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}$

Se verifica que U es una matriz ortonormal, es decir, $UU' = I$

$$UU' = \begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = I$$

CÁLCULO DE DERIVADAS DE UN ESCALAR RESPECTO A UN VECTOR

- Derivada de una forma lineal respecto a un vector

Sean $a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ y $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, entonces $\dfrac{\partial a'x}{\partial x} = a$

$$a'x = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

Derivando la expresión anterior respecto de cada uno de los elementos de x se obtiene:

$$\begin{cases} \dfrac{\partial a'x}{\partial x_1} = a_1 \\ \dfrac{\partial a'x}{\partial x_2} = a_2 \\ \quad \vdots \\ \dfrac{\partial a'x}{\partial x_n} = a_n \end{cases}$$

Reuniendo en un vector las derivadas del escalar $a'x$ con respecto a cada elemento de x, se tiene la derivada de dicho escalar con respecto al vector x. En consecuencia:

$$\frac{\partial a'x}{\partial x} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a$$

- Derivada de una forma cuadrática respecto a un vector

Sean $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$ y $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, entonces $\dfrac{\partial x'Ax}{\partial x} = (A + A')\,x$

$$\mathbf{x'Ax} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} =$$

$$= \left[a_{11}x_1 + a_{21}x_2 + \dots + a_{n1}x_n \quad a_{12}x_1 + a_{22}x_2 + \dots + a_{n2}x_n \quad \dots \quad a_{1n}x_1 + a_{2n}x_2 + \dots + a_{nn}x_n\right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} =$$

$$= \left[(a_{11}x_1 + a_{21}x_2 + \dots + a_{n1}x_n)\,x_1 \quad (a_{12}x_1 + a_{22}x_2 + \dots + a_{n2}x_n)\,x_2 \quad \dots \quad (a_{1n}x_1 + a_{2n}x_2 + \dots + a_{nn}x_n)\,x_n\right]$$

Derivando la expresión anterior respecto a cada uno de los elementos de x se tiene:

$$\begin{cases} \dfrac{\partial x'Ax}{\partial x_1} = 2a_{11}x_1 + (a_{12} + a_{21})x_2 + \dots + (a_{1n} + a_{n1})x_n \\ \dfrac{\partial x'Ax}{\partial x_2} = (a_{12} + a_{21})x_1 + 2a_{22}x_2 + \dots + (a_{2n} + a_{n2})\,x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dfrac{\partial x'Ax}{\partial x_n} = (a_{n1} + a_{1n})x_1 + (a_{n2} + a_{2n})x_2 + \dots + 2a_{nn}x_n \end{cases}$$

Reuniendo en un vector las derivadas del escalar $x'Ax$ con respecto a cada elemento de x, se tiene la derivada de dicho escalar con respecto al vector x. En consecuencia:

$$\frac{\partial x'Ax}{\partial x} = \left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = (A + A')x$$

Si la matriz A es simétrica, se tiene: $\dfrac{\partial x'Ax}{\partial x} = 2Ax$

Instrumentos Estadísticos Avanzados
Facultad Ciencias Económicas y Empresariales
Departamento de Economía Aplicada
Profesor: Santiago de la Fuente Fernández