Anales de la Real Academia de Doctores de España
Volumen 16, n.º 1, 2012, pp. 51-64

ELEMENTOS BÁSICOS EN EL ANÁLISIS EN COMPONENTES PRINCIPALES (ACP)

[Tercera parte: teoría (métricas distintas de las usuales en el ACP)]

Doctor D. Fco. Javier Díaz-Llanos Sainz-Calleja
Académico de Número de la Real Academia de Doctores de España
Sección Ingeniería. N.º 118
Medalla al Mérito Doctoral. Categoría de Plata

M. Yves Escoufier
Professeur d'Analyse des Données.
Ancien Président à l'Université de Montpellier II

Dra. Dña. M.ª del Carmen Cermeño Carrasco
Miembro de Número de la Sociedad Española de Genética Humana. N.º 467
Antigua profesora de Genética y Citogenética en las Universidades
Technische de München donde obtuvo «venia docendis»
y, Freie Angewandte, Berlin, Deutschland.
Referee de artículos desde dichas Universidades

D. Luis Felipe Grau Segura
Licenciado en Ciencias y Técnicas Estadísticas
en la Universidad Carlos III de Getafe (Madrid)

RESUMEN

El objetivo de este artículo es, no sólo mostrar —de la manera más didáctica posible— cómo se calcula: en primer lugar, la raíz cuadrada de una matriz simétrica real definida positiva de orden p y sus propiedades (1 (p. 33)), así como la simetrización de una matriz a diagonalizar (2 (pp. 102-103)) y, en segundo lugar, cómo se calculan los ejes principales, los factores principales y las componentes principales, fuera de la exclusividad de las dos **métricas** clásicas introducidas en R^p y la introducida en R^n en los paquetes de programas de análisis estadístico multidimensional, actualmente comercializados en los cuales está contenido el (ACP).

En esta ocasión, los investigadores científicos, conocedores de sus propios datos empíricos con el concurso de alguien que conozca el ACP à la française, podrán construir las métricas a introducir en R^p y en R^n, que más se adecuen a sus trabajos de investigación concretos.

A título —simplemente didáctico— realizaremos los cálculos matriciales, salvo el proceso de diagonalización de los operadores $\chi^{F}_{(p,n)} N_{(n,n)} \chi^{C}_{(n,p)} M_{(p,p)}$, $M_{(p,p)} \chi^{F}_{(p,n)} N_{(n,n)} \chi^{C}_{(n,p)}$ y $\chi^{C}_{(n,p)} M_{(p,p)} \chi^{F}_{(p,n)} N_{(n,n)}$, mediante una simple calculadora.

También observaremos dos puntos de interés:

1.º) En (3 (pp. 70-83)), los ejes principales son los mismos que los factores principales por el simple hecho de que, se verifica expresándolo mediante métricas que:

$$M_{(p,p)} = I_{(p,p)} \text{ y } N_{(n,n)} = 1/n\ I_{(n,n)},$$

En los casos que vamos a exponer en este artículo, no serán los mismos y, por supuesto, tampoco las componentes principales.

2.º) Así como, en los casos usuales —contenidos en el programa del ACP— el operador $\chi^{F}_{(p,n)} N_{(n,n)} \chi^{C}_{(n,p)} M_{(p,p)}$ es simétrico, en los casos que lo expondremos aquí, no lo es, por lo tanto, es aconsejable utilizar un procedimiento para la **simetrización de una matriz a diagonalizar.** Para lograr este objetivo nos hemos basado en (2 (pp. 102-103)) donde, la **matriz a diagonalizar** no es simétrica [análisis de correspondencias simples (**AFC**)].

Palabras claves: valores propios, vectores propios, raíz cuadrada de una matriz simétrica real definida positiva de orden p, simetrización de una matriz a diagonalizar, los ejes principales, los factores principales y las componentes principales.

PRE-INTRODUCCIÓN

En este artículo contemplamos una pre-introducción, no por simple capricho, sino porque, está íntimamente relacionada con el espíritu con el que se han elaborado —secuencialmente— los cuatro artículos que constituyen el contenido de lo que denominamos: «Elementos básicos en el análisis en componentes principales (ACP)».

A continuación, vamos hacer alusión a cinco frases claves emitidas por cinco grandes científicos que avalan nuestra forma de presentarlos.

1.º Friedrich Nietzsche (filósofo prusiano).

Nietzsche dijo: *«es mi ambición decir en 10 frases lo que todos los demás dicen en un libro, lo que todos los demás no dicen ni en un libro».*

2.º Immanuel Kant (conocido por sus estudios en Ciencias Naturales, Matemáticas y Filosofía).

Kant dijo: *«la teoría sin práctica es ciega, la práctica sin teoría es absurda».*

3.º Johann Wolfgang Goethe (aunque más conocido como literato, poeta y escritor, fue también conocido como naturalista).

Goethe dijo: *«las ideas atrevidas son como piezas de ajedrez llevadas al ataque; pueden caer vencidas, pero también pueden iniciar un juego vencedor».*

4.º Pierre Dagnelie [professeur d'Analyse Statistique à la Faculté des Sciences Agronomiques de Gembloux (Belgique)].

Dagnelie dijo: *«L'honnêteté et l'esprit critique sont deux éléments essentiels du code de déontologie du statisticien, professionnel ou amateur».*

5.º Jean-Paul Benzécri [professeur d'Analyse des Données à l'Université de Pierre-et-*Marie*-Curie de Paris (France) et Président du Comité Scientifique de Les Cahiers de l´Analyse des Données (30 professeurs)].

Benzécri dijo: *«Statistique n'est pas probabilité. Sous le nom de Statistique mathématique, des auteurs (qui, je vous le dis en français, n'écrivent guère dans notre langue…) ont édifié une pompeuse discipline, riche en hypothèses qui ne sont jamais satisfaites dans la pratique. Ce n'est pas de ces auteurs qu'il faut attendre la solution de nos problèmes typologiques»* (4 (pp. 3-6)).

A parte de estas cinco frases dichas por estos cinco grandes investigadores científicos, no está de más apuntar las aportaciones de la investigadora M.ª del Carmen Cermeño Carrasco sobre el **método científico** en un capítulo del libro titulado: «El Análisis de Datos en el cierre de ventas» (5 (pp. 15-19)), y la propia definición de **método científico** elaborada por Luis Martínez de Velasco (profesor de Filosofía y Ética en el IES «Silverio Lanza» de Getafe) y Fco. Javier Díaz-Llanos, once años antes del lanzamiento del libro —ya aludido con anterioridad— al mercado.

Martínez de Velasco, en colaboración con Díaz-Llanos, dijeron:

*«El **método científico** consiste en la articulación de diversos momentos empíricos y abstractos en el marco **metateórico** de una metodología comprometida con la realidad»* (6 (p. 16)).

Metateoría: *Reflexión sobre la propia teoría. Parte de **hipótesis técnico-normativas** y desemboca en una reflexión puramente filosófica acerca del comportamiento, su utilidad, su finalidad, su legitimidad o ilegitimidad moral, etc.*

Hipótesis técnico-normativa: *La que dice qué se debe hacer para lograr un buen conocimiento de la realidad. Por ejemplo, ¿Cuál debería ser el **criterio matemático** que más se aproxime a la realidad para **segmentar un mercado**? ¿Cuál debería ser el criterio matemático que más se aproxime a la realidad para la **elección de variables**, dentro de un conjunto de variables predeterminadas «a priori» por los conocedores del tema que se esté estudiando? Ni que decir tiene, tanto la **segmentación de un mercado** como el problema de la **elección de variables** debe hacerse **fuera de hipótesis distribucionales «a priori»** tal como lo aconseja Jean-Paul Benzécri (4 (pp. 3-6)). En (7), entre otros análisis de datos que se proponen para la obtención de una clasificación espacio-temporal, hemos aplicado un **criterio matemático** para la **elección de variables** propuesto por Yves Escoufier **fuera de hipótesis distribucionales «a priori»** consistente en tres métodos: **ACP** à la française, **ACPVI** (análisis en componentes principales con respecto a variables instrumentales y Elección de variables **RV** (coeficiente de correlación vectorial de Yves Escoufier).*

*En última instancia, hemos de recordar que los que tienen que decidir cómo se debe cortar un **dendrograma** para la obtención de **clases lo más homogéneas posibles** (línea horizontal o línea sinuosa) son los investigadores científicos, conocedores de sus propios datos empíricos, tal como lo apunta F. Benzécri (uno de los 30 investigadores que constituyen el Comité Científico ya aludido) (8 (pp. 49-72)).*

Estos cuatro artículos han sido elaborados amparándonos en las cinco frases lanzadas por los cinco investigadores científicos y en los comentarios de la investigadora científica M.ª del Carmen Cermeño Carrasco y el profesor Luis Martínez de Velasco junto la aportación de Fco. Javier Díaz-Llanos en temas de **filosofía de la ciencia,** tan vivos en otra época pasada como frecuentemente olvidados en el siglo XXI (2011) por una parte notable de los trabajos científicos.

Esperamos que nuestra honestidad y espíritu crítico serán de utilidad para los investigadores científicos, y éstos apliquen el **análisis en componentes principales (ACP)** a sus propios datos empíricos y, de esta manera, según la frase ya aludida de Johann Wolfgang Goethe, consigan un «juego vencedor» en el cual, la credibilidad de sus resultados se aproxime lo máximo posible a la realidad (6).

INTRODUCCIÓN

Advertimos a todos los profesionales de la importancia de la previa asimilación de los dos artículos anteriores (9, 10), antes de la lectura de este tercer artículo ya que, estos tres, junto con un cuarto, forman parte del mismo cuerpo titulado: **«Elementos básicos en el análisis en componentes principales».**

En esta ocasión, antes de proceder al cálculo de **los ejes principales, los factores principales y las componentes principales fuera de la exclusividad de las métricas usuales en** R^p y R^n, consideramos aconsejable recordar no sólo cómo se calcula **la raíz cuadrada de una matriz simétrica real definida positiva de orden p,** así como sus propiedades (1 (p. 33)), sino también, **la simetrización de una matriz de orden p a diagonalizar** (2 (pp. 102-103)).

El hecho de que consideremos, no sólo el primer punto (1 (p. 33)) sino también, la adaptación del segundo (2 (pp. 102-103)) a nuestro caso concreto, como partes integrantes del apartado **método,** no es por puro capricho sino, porque en los libros de Álgebra Lineal, al menos los escritos en castellano, se omiten estos puntos básicos para llevar con éxito la interpretación de los datos empíricos, principal objetivo de nuestro trabajo.

El conocimiento de estos dos puntos permitirá —sin duda— **aprender y comprender cómo se calculan los ejes principales, los factores principales y las componentes principales fuera de las métricas usuales en** R^p y R^n.

MATERIAL

El material básico que necesitamos para alcanzar nuestro objetivo es el conocimiento —en profundidad— de los elementos que constituyen el triplete estadístico:

$$\left(\chi^{c}_{(n,p)}, M_{(p,p)}, N_{(n,n)}\right)$$

Métricas introducidas en R^p

Por el simple hecho de que las métricas introducidas en no sean las dos usuales en el (ACP):

$$M_{(p,p)} = I_{(p,p)} \qquad M_{(p,p)} = D_{\frac{1}{s^{2j}}} \qquad j = 1, \ldots, p$$

$$\text{y en } R^n\text{: } N_{(n,n)} = \frac{1}{n} I_{(n,,n)}$$

el operador $\chi^F_{(p,n)} N_{(n,n)} \chi^C_{(n,p)} M_{(p,p)}$, ya no será simétrico.

Métricas propuestas

Vamos a proponer dos tipos de **métricas simétricas reales definidas positivas de orden p.** Mientras que las **métricas tipo I** son diagonales, las **métricas tipo II** no lo son.

Aquellos lectores que deseen conocer con más precisión el significado de término **definida no negativa y definida positiva**, les aconsejamos que consulten el libro de Bernard Guerrien (11 (pp. 211-245)).

Métricas tipo I

Este tipo de **métricas** adopta la siguiente forma:

$$M_{(p,p)} = \operatorname{diag}\left(m^j_{j'}\right) \qquad j = j' \qquad j = 1, \dots, p$$

Los $m^j_{j'}$, representan los pesos asignados consigo mismo de cada una de las p variables cuantitativas que hayamos elegido «a priori» para la realización del trabajo de investigación concreto.

Métricas tipo II

Este tipo de **métricas** adopta la siguiente forma:

$$M_{(p,p)} = \begin{pmatrix} m^1_1 & m^2_1 & \cdots & m^p_1 \\ m^1_2 & m^2_2 & \cdots & m^p_2 \\ \vdots & \vdots & \ddots & \vdots \\ m^1_p & m^2_p & \cdots & m^p_p \end{pmatrix}$$

Si M es de orden p tendremos que calcular

$$\frac{[p(p-1)]}{2} m^j_{j'}, \qquad j < j', \quad j = 1, \dots, p$$

más los $m^j_{j'}$, elementos contenidos en la diagonal principal de $M_{(p,p)}$.

Los m^j_j representan el peso de las variables j frente a las j´ y los m^j_j representan el peso de las variables j´ frente a las j.

Métricas introducidas en R^n

Si en R^n introducimos una métrica que adopte la siguiente estructura:

$$N_{(n,n)} = \mathrm{diag}(p_i) \qquad \sum_{i=1}^{i=n} p_i = 1$$

donde, los p_i representan los pesos de cada individuo con respecto a los individuos que se consideran en el trabajo de investigación concreto. No está demás recordar que, el producto de las matrices $\chi^F_{(p,n)} N_{(n,n)} \chi^C_{(n,p)}$ sea cual sea la matriz $N_{(n,n)}$ seguirá siendo una matriz simétrica.

MÉTODO

Tal como ya hemos indicado en la introducción, aunque el cálculo de la **raíz cuadrada de una matriz simétrica real definida positiva de orden p y el de la simetrización de una matriz a diagonalizar** son dos conceptos básicos a tener en cuenta para el cálculo de **los ejes principales, los factores principales y las componentes principales**, los incluiremos en este apartado dada su ausencia en los libros de Álgebra Lineal, al menos, escritos en la lengua española (castellano).

1. Raíz cuadrada de una matriz simétrica real definida positiva de orden p

Aunque para desarrollar este apartado nos hemos basado en el trabajo realizado en (1 (p. 33)) en el cual, se parte de una matriz simétrica definida no-negativa, las fórmulas expuestas aquí, sin embargo, son igualmente válidas para las definidas positivas.

En al ACP, la métrica $M_{(p,p)}$ es una matriz simétrica definida positiva (12 (p. 50)); es decir, todos los valores propios de $M_{(p,p)}$ son positivos y, ninguno es cero.

1.1. *Primera opción:*

La **métrica** $M_{(p,p)}$ adopta la siguiente forma:

$$M_{(p,p)} = \mathrm{diag}\left(m^j_{j'}\right) \qquad j = j' \qquad j = 1, \dots, p$$

En esta opción el cálculo de la raíz cuadrada es inmediato:

$$M^{\frac{1}{2}}_{(p,p)} = \mathrm{diag}\left(\sqrt{m^j_{j'}}\right) \qquad j = j' \qquad j = 1, \dots, p$$

1.2. *Segunda opción:*

La **métrica** $M_{(p,p)}$ adopta la siguiente forma:

$$M_{(p,p)} = \begin{pmatrix} m_1^1 & m_1^2 & \cdots & m_1^p \\ m_2^1 & m_2^2 & \cdots & m_2^p \\ \vdots & \vdots & \ddots & \vdots \\ m_p^1 & m_p^2 & \cdots & m_p^p \end{pmatrix}$$

El cálculo de esta segunda opción ya no es inmediato, por lo tanto, tendremos que definir el concepto de la raíz cuadrada de una matriz simétrica real definida positiva de p y estudiar sus propiedades.

1.3. *Definición*

Si $M_{(p,p)}$ es una matriz simétrica real definida positiva de orden p, su raíz cuadrada se define de la siguiente manera:

$$M_{(p,p)}^{\frac{1}{2}} = U_{(p,p)} L_{(p,p)}^{\frac{1}{2}} U_{(p,p)}^{T} \qquad (1)$$

donde,

$U_{(p,p)}$: es una matriz de orden p cuyas columnas son los vectores propios normalizados de $M_{(p,p)}$

$$L_{(p,p)}^{\frac{1}{2}} = \mathrm{diag}\left(\sqrt{\lambda_j}\right) \qquad j = 1, \dots, p$$

λ_j: son los valores propios de $M_{(p,p)}$

$U^T_{(p,p)}$: es la matriz transpuesta de $U_{(p,p)}$

1.4. *Propiedades*

1.4.1. *Primera:*

$$M_{(p,p)}^{\frac{1}{2}} M_{(p,p)}^{\frac{1}{2}} = M_{(p,p)}$$

Demostración:

Para llevar a cabo dicha demostración tendremos en cuenta la definición de $M_{(p,p)}^{\frac{1}{2}}$ (1),

$$U^{T}_{(p,p)}U_{(p,p)} = I_{(p,p)} \quad (2) \quad \text{y} \quad L^{\frac{1}{2}}_{(p,p)}I_{(p,p)}L^{\frac{1}{2}}_{(p,p)} = L_{(p,p)} \quad (3)$$

$$M^{\frac{1}{2}}_{(p,p)}M^{\frac{1}{2}}_{(p,p)} \overset{1}{=} U_{(p,p)}L^{\frac{1}{2}}_{(p,p)}U^{T}_{(p,p)}U_{(p,p)}L^{\frac{1}{2}}_{(p,p)}U^{T}_{(p,p)} \overset{2}{=}$$

$$= U_{(p,p)}L^{\frac{1}{2}}_{(p,p)}\, I_{(p,p)}L^{\frac{1}{2}}_{(p,p)}U^{T}_{(p,p)} \overset{3}{=} U_{(p,p)}\, L_{(p,p)}\, U^{T}_{(p,p)}$$

De lo que se desprende que,

$$M_{(p,p)} = U_{(p,p)}\, L_{(p,p)}\, U^{T}{}_{(p,p)}$$

1.4.2. *Segunda:*

$$\left(M^{\frac{1}{2}}_{(p,p)}\right)^{-1} = M^{-\frac{1}{2}}_{(p,p)}$$

Demostración:

Para llevar a cabo dicha demostración tendremos en cuenta la definición de $M^{\frac{1}{2}}_{(p,p)}$ (1) y además, las siguientes propiedades:

$$U^{T}_{(p,p)} = U^{-1}_{(p,p)} \;\; (2), \qquad \left(U^{T}_{(p,p)}\right)^{-1} = U_{(p,p)} \;\; (3), \;\; \text{y}$$

$$\left(U_{(p,p)}\, L^{\frac{1}{2}}_{(p,p)}U^{T}_{(p,p)}\right)^{-1} = \left(U^{T}_{(p,p)}\right)^{-1}\left(L^{\frac{1}{2}}_{(p,p)}\right)^{-1} U^{-1}_{(p,p)} \quad (4)$$

$$\left(M^{\frac{1}{2}}_{(p,p)}\right)^{-1} \overset{1}{=} \left(U_{(p,p)}L^{-\frac{1}{2}}_{(p,p)}U^{T}_{(p,p)}\right)^{-1} \overset{4}{=}$$

$$= \left(U^{T}_{(p,p)}\right)^{-1}\left(L^{\frac{1}{2}}_{(p,p)}\right)^{-1} U^{-1}_{(p,p)} \overset{(2,3)}{=} U_{(p,p)}L^{-\frac{1}{2}}_{(p,p)}U^{T}_{(p,p)}$$

De lo que se desprende que,

$$M^{-\frac{1}{2}}_{(p,p)} = U_{(p,p)}\, L^{-\frac{1}{2}}_{(p,p)}U^{T}_{(p,p)}$$

donde,

$$L^{-\frac{1}{2}}_{(p,p)} = \mathrm{diag}\left(\frac{1}{\sqrt{\lambda_j}}\right) \qquad\qquad j = 1, \dots, p$$

2. Simetrización del operador $\chi^F_{(p,n)}N_{(n,n)}\chi^c_{(n,p)}M_{(p,p)}$ a diagonalizar:

Dado que si introducimos en R^p una **métrica distinta de** las dos usuales en el ACP, el operador $\chi^F_{(p,n)}N_{(n,n)}\chi^c_{(n,p)}M_{(p,p)}$ ya no será simétrico y, por lo tanto, si deseamos encontrar, en principio los **ejes principales**, es aconsejable proceder a la simetrización de dicho operador.

2.1. *Proceso de simetrización del operador* $\chi^F_{(p,n)}N_{(n,n)}\chi^c_{(n,p)}M_{(p,p)}$

Si llamamos $R_{(p,p)} = \chi^F_{(p,n)}N_{(n,n)}\chi^c_{(n,p)}$ (1)

y post-multiplicamos ambos miembros de (1) por $M_{(p,p)}$ tendremos,

$$\chi^F_{(p,n)}N_{(n,n)}\chi^c_{(n,p)}M_{(p,p)} = R_{(p,p)}M_{(p,p)}$$

De (1) observamos que $R_{(p,p)}$ es una matriz simétrica real, pero el operador $R_{(p,p)}M_{(p,p)}$ no lo es, por lo tanto, procederemos a simetrizar $R_{(p,p)}M_{(p,p)}$.

Para lograr dicho objetivo, utilizaremos la primera propiedad de la **raíz cuadrada de una matriz simétrica real definida positiva de orden p.**

$$M^{\frac{1}{2}}_{(p,p)}M^{\frac{1}{2}}_{(p,p)} = M_{(p,p)}$$

Así pues,

$$\chi^F_{(p,n)}N_{(n,n)}\chi^c_{(n,p)}M_{(p,p)} = R_{(p,p)}\, M^{\frac{1}{2}}_{(p,p)}M^{\frac{1}{2}}_{(p,p)} \qquad (2)$$

Partiendo de la relación:

$$\chi^F_{(p,n)}N_{(n,n)}\chi^c_{(n,p)}M_{(p,p)}\vec{\mathbf{u}}^j = \lambda_j\, \vec{\mathbf{u}}^j \qquad (3)$$

y, teniendo en cuenta el segundo miembro de (2), llegamos a:

$$R_{(p,p)}\, M^{\frac{1}{2}}_{(p,p)}M^{\frac{1}{2}}_{(p,p)}\vec{\mathbf{u}}^j = \lambda_j\, \vec{\mathbf{u}}^j$$

Finalmente, pre-multiplicando ambos miembros de esta igualdad por $M^{\frac{1}{2}}_{(p,p)}$ y teniendo en cuenta

$$M^{\frac{1}{2}}_{(p,p)}\vec{\mathbf{u}}^j = \vec{\mathbf{s}}^j,$$

llegamos al siguiente resultado:

$$M_{(p,p)}^{\frac{1}{2}} R_{(p,p)} M_{(p,p)} \vec{\mathbf{s}}^{j} = \lambda_j \ \vec{\mathbf{s}}^{j}$$

Así como $R_{(p,p)}$ es una matriz simétrica real, la matriz

$$M_{(p,p)}^{\frac{1}{2}} R_{(p,p)} M_{(p,p)}^{\frac{1}{2}}$$

también lo es y, además

$$M_{(p,p)}^{\frac{1}{2}} R_{(p,p)} \ M_{(p,p)}^{\frac{1}{2}}$$

tiene los mismos **valores propios** que $R_{(p,p)}M_{(p,p)}$

Los **vectores propios** de $\chi^{F}{}_{(p,n)}N_{(n,n)}\,\chi^{c}{}_{(n,p)}M_{(p,p)}$ se calcularán a partir de los **vectores propios** de $M_{(p,p)}^{\frac{1}{2}} R_{(p,p)} M_{(p,p)}^{\frac{1}{2}}$ haciendo uso de la siguiente relación:

$$\vec{\mathbf{u}}^{j} = M_{(p,p)}^{-\frac{1}{2}} \vec{\mathbf{s}}^{j}$$

Para el cálculo de $M_{(p,p)}$ haremos uso de la segunda propiedad de la raíz cuadrada de una matriz simétrica real definida positiva de orden p.

3. El cuadrado de la norma de los vectores $\vec{\mathbf{u}}^{j}$ en el sentido de la métrica $M_{(p,p)}$ es igual al cuadrado de la norma de los vectores $\vec{\mathbf{s}}^{j}$ en el sentido de la métrica $I_{(p,p)}$; es decir,

$$\left\|\vec{\mathbf{u}}^{j}\right\|_{M_{(p,p)}}^{2} = \left\|\vec{\mathbf{s}}^{j}\right\|_{I_{(p,p)}}^{2}$$

Demostración:

Para la realización de esta demostración partiremos de la definición de $\left\|\vec{\mathbf{u}}^{j}\right\|_{M_{(p,p)}}^{2}$ y de la relación: $\vec{\mathbf{u}}^{j} = M_{(p,p)}^{-\frac{1}{2}} \vec{\mathbf{s}}^{j}$ (1)

$$\left\|\vec{\mathbf{u}}^{j}\right\|_{M}^{2} = \vec{\mathbf{u}}^{j\,T}\, M\, \vec{\mathbf{u}}^{j} \overset{1}{\downarrow}{=} \vec{\mathbf{s}}^{j\,T}\, M_{(p,p)}^{-\frac{1}{2}} M_{(p,p)} M_{(p,p)}^{-\frac{1}{2}} \vec{\mathbf{s}}^{j} =$$

$$= \vec{\mathbf{s}}^{j\,T} \vec{\mathbf{s}}^{j} = \left\|\vec{\mathbf{s}}^{j}\right\|_{I_{(p,p)}}^{2}$$

Por lo tanto,

$$\left\|\vec{\mathbf{u}}^{j}\right\|^{2}_{M_{(p,p)}} = \left\|\vec{\mathbf{s}}^{j}\right\|^{2}_{I_{(p,p)}}$$

De lo que se desprende que los vectores propios $\vec{\mathbf{s}}^{j}$ de

$M^{\frac{1}{2}}_{(p,p)} R_{(p,p)} M^{\frac{1}{2}}_{(p,p)}$ son $I_{(p,p)}$-**ortonormados.**

4. Los ejes principales

Los **ejes principales** son los vectores propios de $\chi^{F}_{(p,n)} N_{(n,n)} \chi^{C}_{(n,p)} M_{(p,p)}$ siendo estos $M_{(p,p)}$-**ortonormados**.

$\chi^{F}_{(p,n)} N_{(n,n)} \chi^{c}_{(n,p)}\ M_{(p,p)} \vec{\mathbf{u}}^{j} = \lambda_j \vec{\mathbf{u}}^{j}$ $\quad M_{(p,p)}$ -**ortonormados**; es decir, para aquellos que no conozcan esta nomenclatura indicamos que, los vectores propios de $\chi_{(p,n)} N_{(n,n)} \chi^{c}_{(n,p)}\ M_{(p,p)}$ tendrán que cumplir las siguientes condiciones:

$$\begin{aligned} {\vec{\mathbf{u}}^{j}}^{T} M_{(p,p)} \vec{\mathbf{u}}^{j} &= 1 \qquad j = 1, \dots, p \\ {\vec{\mathbf{u}}^{j}}^{T} M_{(p,p)} \vec{\mathbf{u}}^{j'} &= 0 \qquad j \neq j' \end{aligned}$$

5. El cuadrado de la norma de los vectores $\vec{\mathbf{v}}^{j}$ en el sentido de la métrica $M^{-1}_{(p,p)}$ es igual al cuadrado de la norma de los vectores $\vec{\mathbf{s}}^{j}$ en el sentido de la métrica $I_{(p,p)}$

$$\left\|\vec{\mathbf{v}}^{j}\right\|^{2}_{M^{-1}_{(p,p)}} = \left\|\vec{\mathbf{s}}^{j}\right\|^{2}_{I_{(p,p)}}$$

Demostración:

Para la realización de esta demostración partiremos de la definición de $\left\|\vec{\mathbf{v}}^{j}\right\|^{2}_{M^{-1}_{(p,p)}}$ y de la relación: $\vec{\mathbf{v}}^{j} = M^{\frac{1}{2}}_{(p,p)} \vec{\mathbf{s}}^{j}$ (1)

$$\left\|\vec{\mathbf{v}}^{j}\right\|^{2}_{M^{-1}_{(p,p)}} = {\vec{\mathbf{v}}^{j}}^{T}\ M^{-1}_{(p,p)} \vec{\mathbf{v}}^{j} \overset{1}{=} {\vec{\mathbf{s}}^{j}}^{T} M^{\frac{1}{2}}_{(p,p)} M_{(p,p)}\ M^{\frac{1}{2}}_{(p,p)} \vec{\mathbf{s}}^{j} =$$
$$= {\vec{\mathbf{s}}^{j}}^{T} \vec{\mathbf{s}}^{j} = \left\|\vec{\mathbf{s}}^{j}\right\|^{2}_{I_{(p,p)}}$$

Por lo tanto,

$$\left\|\vec{\mathbf{v}}^{j}\right\|^{2}_{M^{-1}_{(p,p)}} = \left\|\vec{\mathbf{s}}^{j}\right\|^{2}_{I_{(p,p)}}$$

De lo que se desprende que, la norma al cuadrado de los vectores $\vec{\mathbf{s}}^{j}$ de

$$M_{(p,p)}^{\frac{1}{2}} R_{(p,p)}\, M_{(p,p)}^{\frac{1}{2}} \text{ son } I_{(p,p)}\textbf{-ortonormados.}$$

6. Los factores principales

Los **factores principales** son los **vectores propios de**

$M_{(p,p)}\chi_{(p,n)}^{F} N_{(n,n)}\chi_{(n,p)}^{c} \qquad M_{(p,p)}^{-1}$ **-ortonormados.**

$$M_{(p,p)}\chi_{(p,n)}^{F} N_{(n,n)}\chi_{(n,p)}^{c}\vec{\mathbf{v}}^{j} = \lambda_j \vec{\mathbf{v}}^{j} \qquad M_{(p,p)}^{-1}\textbf{-ortonormados.}$$

Igualmente indicamos a aquellos no familiarizados con esta nomenclatura que, los vectores propios de

$$M_{(p,p)}\ \chi_{(p,n)}^{F} N_{(n,n)}\chi_{(n,p)}^{c}$$

tendrán que cumplir las siguientes condiciones:

$$\vec{\mathbf{v}}^{j\,T} M_{(p,p)}^{-1}\vec{\mathbf{v}}^{j} = 1 \qquad j = 1, \dots, p$$
$$\vec{\mathbf{v}}^{j\,T} M_{(p,p)}^{-1}\vec{\mathbf{v}}^{j\,\prime} = 0 \qquad j \neq j'$$

7. El cuadrado de la norma de los vectores $\vec{\mathbf{c}}^{j}$ en el sentido de la métrica $N_{(n,n)}$ es igual al cuadrado de la norma de los vectores $\vec{\mathbf{s}}^{j}$ en el sentido de la métrica

$$M_{(p,p)}^{\frac{1}{2}}\ \chi_{(p,n)}^{F} N_{(n,n)}\chi_{(n,p)}^{c}\ M_{(p,p)}^{\frac{1}{2}}$$

$$\left\|\vec{\mathbf{c}}^{j}\right\|_{N_{(n,n)}}^{2} = \left\|\vec{\mathbf{s}}^{j}\right\|_{M_{(p,p)}^{\frac{1}{2}}\ \chi_{(p,n)}^{F} N_{(n,n)}\chi_{(n,p)}^{c}\ M_{(p,p)}^{\frac{1}{2}}}^{2}$$

Demostración:

Para la realización de esta demostración partiremos de la definición de $\left\|\vec{\mathbf{c}}^{j}\right\|_{N_{(n,n)}}^{2}$ y de la relación: $\vec{\mathbf{c}}^{j} = \chi_{(n,p)}^{c}\ M_{(p,p)}^{\frac{1}{2}}\vec{\boldsymbol{s}}^{j}(1)$

$$\left\|\vec{\mathbf{c}}^{j}\right\|_{N_{(n,n)}}^{2} = \vec{\mathbf{c}}^{j\,T} N_{(n,n)}\vec{\mathbf{c}}^{j} \overset{1}{=}$$

$$= \vec{\mathbf{s}}^{j\,T} M_{(p,p)}^{\frac{1}{2}}\ \chi_{(p,n)}^{F} N_{(n,n)}\chi_{(n,p)}^{c}\ M_{(p,p)}^{\frac{1}{2}}\vec{\mathbf{s}}^{j} =$$

$$= \left\|\vec{\mathbf{s}}^{j}\right\|_{M_{(p,p)}^{\frac{1}{2}} R_{(p,p)}\ M_{(p,p)}^{\frac{1}{2}}}^{2}$$

Por lo tanto,

$$\left\|\vec{\mathbf{c}}^{j}\right\|^{2}_{N_{(n,n)}} = \left\|\vec{\mathbf{s}}^{j}\right\|^{2}_{M^{\frac{1}{2}}_{(p,p)} R_{(p,p)} M^{\frac{1}{2}}_{(p,p)}}$$

De lo que se desprende que los vectores $\vec{\mathbf{s}}^{j}$ son

$M^{\frac{1}{2}}_{(p,p)}\ \chi^{F}_{(p,n)} N_{(n,n)} \chi^{c}_{(n,p)} M^{\frac{1}{2}}_{(p,p)}$ **-ortonormados**.

8. Las componentes principales

Las componentes principales son los vectores propios de:

$\chi^{c}_{(n,p)} M_{(p,p)} \chi^{F}_{(p,n)} N_{(n,n)} \qquad N_{(n,n)}$**-ortogonales**

$$\chi^{c}_{(n,p)} M_{(p,p)} \chi^{F}_{(p,n)} N_{(n,n)} \vec{\mathbf{c}}^{j} = \lambda_j\ \vec{\mathbf{c}}^{j} \qquad N_{(n,n)}\textbf{-ortogonales}$$

Es decir, los vectores propios $\vec{\mathbf{c}}^{j}$, además de que verifiquen:

$$\vec{\mathbf{c}}^{j\,T} N_{(n,n)} \vec{\mathbf{c}}^{j} = \lambda_j$$

tendrán que cumplir la siguiente condición:

$$\vec{\mathbf{c}}^{j\,T} N_{(n,n)} \vec{\mathbf{c}}^{j'} = 0 \qquad j \neq j'$$

CONCLUSIÓN

Una vez presentada la parte teórica (métricas no usuales en el ACP), en el cuarto y último artículo presentaremos dos ejercicios de aplicación que servirán únicamente para que el lector pueda entender y comprender la parte teórica asociada a las métricas no usuales en el ACP.

BIBLIOGRAFÍA

1. Denis, J. by colaborador (1976): *Apuntes redactados con ocasión del cursillo impartido en la Sección de Proceso de Datos sobre el Tratamiento de Datos.* INIA. MAPA.
2. Lebart, L.; Morineau, A.; Piron, M. (1995): *Statistique exploratoire multidimensionnelle.* Dunod.
3. Bon, J.; Grégory, P. (1986): *Techniques marketing.* Vuibert Gestion. Paris.
4. Benzécri, J.-P. et collaborateurs (1973): *L'Analyse des Données.* 2. *L'Analyse des Correspondances.* Dunod.
5. Díaz-Llanos, Fco. J. (2002): *El análisis de datos en el cierre de ventas.* La Muralla, S. A. Hespérides, S. L.
6. Díaz-Llanos, Fco. J. (1991): *Criterios de segmentación de la realidadempírica: Juicios de valor y mediaciones técnicas.*

Discurso de ingreso en la RADE. Depósito Legal: M-40.021-1991.

7. Díaz-Llanos, Fco. J. (1985): *Técnicas multidimensionales para el estudio de la evolución del sector agrario y afines a nivel provincial.* Tesis dirigida por el Profesor Sr. D. José Luis de Miguel Arenal. Catedrático de Matemática I en la ETSI Agrónomos de Madrid.

8. Benzécri, J-P.; Benzécri, F., et collaborateurs (1986): *Pratique de l'Analyse des Données en Économie.* Dunod.

9. Díaz-Llanos, Fco. J.; Escoufier, Yves.; Cermeño, M.ª C., y Grau L. F. (2011): «Elementos básicos en el análisis en componentes principales [Primera parte: teoría]». *Anales de la Real Academia de Doctores de España.* Volumen 15, n.º 1, pp. 51-75.

10. Díaz-Llanos, Fco. J.; Escoufier, Yves.; Cermeño, M.ª C., y Grau L. F. (2011): «Elementos básicos en el análisis en componentes principales [Segunda parte: práctica (métricas usuales en el ACP)]». *Anales de la Real Academia de Doctores de España.* Volumen 15, n.º 2, pp. 71-90.

11. Guerrien, B. (1980): *Álgèbre Linéaire pour economistes.* Rappels de cours et exercices corrigés. Economica.

12. Escoufier, Yves (1979): *Cours d'Analyse des Données.* RT 7901 CRIG. Av. d'Occitanie 34075 Montpellier. Cedex.