

ESTUDIO DE LAS MÉTRICAS INDUCIDAS POR UN ANÁLISIS EN COMPONENTES PRINCIPALES¹

EDUARDO PIZA VOLIO²

Resumen

En este artículo se estudia una amplia gama de *métricas inducidas* obtenidas al realizar un Análisis en Componentes Principales con fines exploratorios, así como algunas de sus consecuencias. Se demuestra que la aplicación iterada del Análisis en Componentes Principales, utilizando sucesivamente las métricas inducidas, conduce de manera natural a la conocida métrica de Mahalanobis, $M = V^{-1}$, como único medio para garantizar la convergencia de la sucesión de métricas inducidas y obtener por lo tanto un proceso estable.

Palabras clave: métricas inducidas, Mahalanobis, Análisis en Componentes Principales.

1. Introducción

El Análisis en Componentes Principales (ACP) es una metodología ampliamente difundida del análisis multivariado de datos, útil a nivel de análisis descriptivo de los datos. Es muy empleado como una técnica eficiente para la condensación de la información multivariada, eliminando la redundancia de la información. También se emplea como técnica de representación visual de datos. Por estas razones es especialmente adecuado como estudio exploratorio y preliminar de la información en investigaciones que involucran Análisis de Datos [Pag76], [Tak82].

Mediante estudios de ACP es posible extraer una “radiografía” de los datos, dejando que éstos “hablen” por sí solos, y nos muestren su estructura interna que ocultan a simple vista. Por otra parte el ACP es la base para el estudio de muchas otras técnicas del análisis multivariado, en las cuales se imponen condiciones adicionales, o se cuenta con información adicional.

Una de las principales características del ACP es que nos permite visualizar la información en estudio representando simultáneamente tanto los “individuos” como las variables que los caracterizan. Esta visualización —en un plano o espacio tridimensional— es posible gracias al llamado *esquema de dualidad* [Pag76]. En la Figura 1 se incluye un ejemplo típico de estas representaciones mixtas de “individuos” y variables en un mismo plano.

Sin embargo al realizar un ACP con sus datos, el investigador debe escoger adecuadamente la *métrica* a emplear en el espacio euclídeo de los “individuos”. Dependiendo de la métrica escogida, así variarán sus resultados. Por ejemplo, en la Figura 1 se utilizó la métrica D_{1/σ^2} , obteniéndose esos resultados. Diferente panorama hubiésemos obtenido si realizamos el ACP con otra métrica.

¹Investigación patrocinada por el Deutscher Akademischer Austauschdienst (DAAD) y la Universidad de Costa Rica.

²PROGRAMA DE INVESTIGACIÓN EN MODELOS Y ANÁLISIS DE DATOS (PIMAD), ESCUELA DE MATEMÁTICA, UNIVERSIDAD DE COSTA RICA, 2060 SAN JOSÉ, COSTA RICA. E-mail: epiza@cariari.ucr.ac.cr

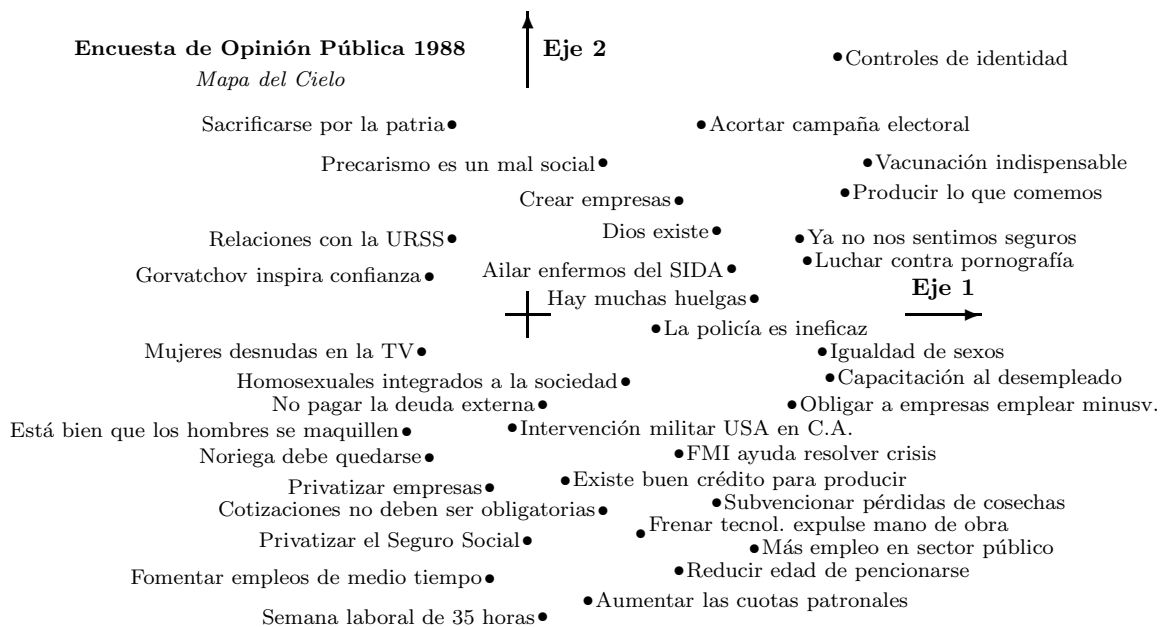
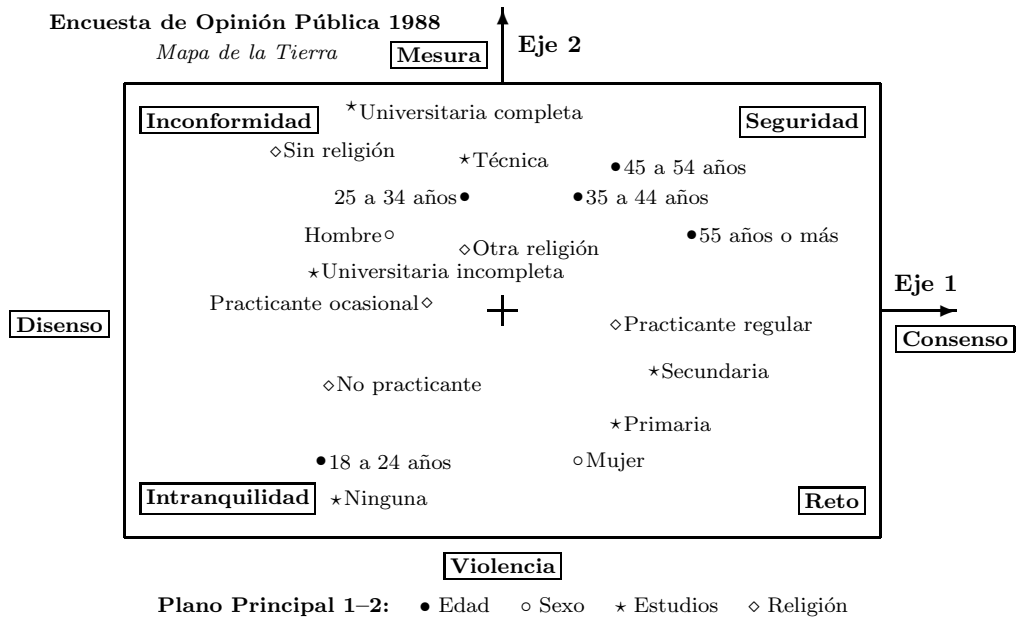


Figura 1: Ejemplo de una representación visual de un Análisis en Componentes Principales, tomada de [Pol89]. En el gráfico de arriba están representados algunos “individuos suplementarios”, mientras que en el de abajo están representadas las variables. Ambos gráficos son complementarios, gracias al esquema de dualidad.

De esta forma, “individuos” que antes parecían estar muy lejanos en el mapa, podrían estar mucho más cercanos de lo que se pensaba originalmente.

Lo anterior nos lleva a preguntarnos, en forma precisa, ¿cuál es la métrica que realmente nos interesa utilizar, mediante la cual podamos explicar satisfactoriamente nuestro universo de datos en estudio? Además, una vez escogida una métrica particular M y realizado un ACP con dicha métrica, ¿cuál es la métrica que induce ese ACP? ¿Qué sucede si ponderamos adecuadamente los ejes principales, empleando convenientemente el porcentaje de inercia total que explica tal eje? ¿Cuál es la “mejor” métrica a utilizar, si estamos interesados en observar las distintas agrupaciones de los individuos, y bajo cuál criterio es la “mejor”?

La respuesta a tales preguntas es quizás una de las cosas más difíciles del Análisis de Datos, y no hay un criterio unificado. Sin embargo veremos en este artículo las consecuencias de una u otra escogencia de una métrica particular, desde un punto de vista de estabilidad del análisis.

2. Resumen de resultados de un ACP

En esta sección haremos un breve resumen teórico de los resultados clásicos del Análisis en Componentes Principales. Una exposición detallada puede encontrarse por ejemplo en [Pag76].

Disponemos de una colección finita $\Omega = \{\underline{x}_1, \dots, \underline{x}_n\}$ de n individuos en estudio. Cada individuo \underline{x}_i tiene asignado un peso ω_i , donde $\sum_{i=1}^n \omega_i = 1$.

Disponemos además de p variables cuantitativas o atributos que caracterizan a nuestros individuos, de manera que el i -ésimo individuo \underline{x}_i se representa en el espacio euclídeo $E = \mathbf{R}^p$ mediante $\underline{x}_i = (x_i^1, x_i^2, \dots, x_i^p)^t \in E$, para $i = 1, 2, \dots, n$. La matriz de datos es entonces

$$X_{p \times n} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) = \begin{pmatrix} x_1^1 & \cdots & x_n^1 \\ \vdots & \ddots & \vdots \\ x_1^p & \cdots & x_n^p \end{pmatrix}.$$

Supondremos además que $n > p$ y todos las variables han sido previamente centradas, de manera que X es una matriz centrada por filas: la suma de cada fila, ponderada por los pesos de los individuos, es nula.

Al espacio euclídeo $E = \mathbf{R}^p$ se le conoce como el *espacio de los individuos*, mientras que al espacio euclídeo $F = \mathbf{R}^n$ se le llama el *espacio de las variables*.

En $E = \mathbf{R}^p$ podemos representar a los individuos mediante los vectores columna de la tabla de datos centrada X :

$$X = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) \quad ; \quad \text{donde } \underline{x}_i \in E \text{ es la representación del } i\text{-ésimo individuo.}$$

En $F = \mathbf{R}^n$ podemos representar a las variables mediante los vectores fila de la tabla de datos centrada X :

$$X = \begin{pmatrix} \underline{x}^1 \\ \underline{x}^2 \\ \vdots \\ \underline{x}^p \end{pmatrix} \quad ; \quad \text{donde } \underline{x}^j \in F \text{ es la representación de la } j\text{-ésima variable centrada.}$$

Sea M una métrica definida sobre E , esto es, una matriz simétrica y definida positiva que nos permite definir un producto interno y una norma en E , de la manera usual: $\|x\|_M = \sqrt{\langle x, x \rangle_M} =$

$\sqrt{x^t M x}$. Luego se verifica el conocido *esquema de dualidad*, que permite organizar matemáticamente los objetos “individuos” y “variables” mediante un enfoque unificado, por dualidad:

$$\begin{array}{ccc} E = \mathbf{R}^p & \xleftarrow{X} & F^* \\ M \downarrow \uparrow V & & W \downarrow \uparrow D_\omega \\ E^* & \xrightarrow{X^t} & F = \mathbf{R}^n \end{array}$$

donde:

- D_ω es la métrica de los pesos de los individuos en el espacio F de las variables: $D_\omega = \text{diag} [\omega_1, \omega_2, \dots, \omega_n]$.
- E^* y F^* son los espacios duales algebraicos de E y F respectivamente.
- X^t es la traspuesta de X .
- $V = X \circ D_\omega \circ X^t$ es la forma cuadrática de inercia o matriz de variancias y covariancias de las p variables. Es simétrica y definida positiva.³
- $W = X^t \circ M \circ X$.

El Análisis en Componentes Principales sobre el espacio de los “individuos” básicamente es el estudio de los valores y vectores propios del operador $V \circ M$:

$$E \xrightarrow{M} E^* \xrightarrow{V} E$$

Sean $\lambda_1, \lambda_2, \dots, \lambda_p$ los p valores propios de $V \circ M$ ordenados en orden decreciente contando posibles repeticiones (todos los λ_i son reales y positivos pues $V \circ M$ es M -simétrica y definida positiva):

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$$

Sean $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_p$ los vectores propios correspondientes, escogidos de manera que $\{\underline{u}_i : i = 1, \dots, p\}$ sea una base M -ortonormal de E , esto es:

$$\forall i, \forall j : \underline{u}_i^t M \underline{u}_j = \delta_{ij}.$$

Luego, los ejes principales del ACP están dados por $\Delta_{u_1}, \Delta_{u_2}, \dots, \Delta_{u_p}$, donde Δ_{u_i} denota la recta engendrada por \underline{u}_i . Cada eje aporta un porcentaje de explicación de la inercia, el cual es igual a $(100 \lambda_i / \sum_{k=1}^p \lambda_k) \%$ para el eje Δ_{u_i} . Los individuos se proyectan en los diversos ejes principales, obteniéndose

$$\underline{x}_i = \sum_{k=1}^p \hat{x}_i^k \underline{u}_k \quad ; \quad \hat{x}_i^k = \underline{x}_i^t M \underline{u}_k$$

donde \hat{x}_i^k es la coordenada de \underline{x}_i en el eje Δ_{u_k} .

El primer plano principal es $\langle \Delta_{u_1}, \Delta_{u_2} \rangle$ y dentro de todos los posibles planos de E es el que mejor representa a los individuos, con un porcentaje de inercia explicada igual a $100 [(\lambda_1 + \lambda_2) / \sum_{k=1}^p \lambda_k] \%$.

En síntesis, el Análisis en Componentes Principales es la búsqueda de los mejores sub-espacios $\Delta_{u_1}, \langle \Delta_{u_1}, \Delta_{u_2} \rangle, \langle \Delta_{u_1}, \Delta_{u_2}, \Delta_{u_3} \rangle, \dots$, en donde se alcanzan las mejores proyecciones M -ortogonales de la nube inicial de individuos Ω .

³Sin pérdida de generalidad vamos a suponer que las variables son linealmente independientes, desde el punto de vista algebraico. En caso contrario, se puede prescindir de aquellas variables redundantes.

3. Métricas clásicas y métricas inducidas

La escogencia de la métrica M es de central importancia cuando se realiza un Análisis en Componentes Principales. Dentro del espacio de los individuos $F = \mathbf{R}^n$ siempre se escoge la métrica de pesos D_ω , llamada también la *métrica de la democracia*, pues respeta la asignación de pesos de los individuos y tiene una interpretación estadística natural [Pag76].

Algunas escogencias usuales para la métrica M son:

- $M = I_p$: la matriz identidad, o métrica euclídea clásica. Es utilizada cuando se presupone que las variables son independientes, y todas están medidas en la misma escala.
- $M = \text{diag}[1/\sigma_1^2, \dots, 1/\sigma_p^2]$: la matriz diagonal cuyas entradas diagonales son las inversas de las variancias de las variables. Se utiliza cuando hay razones para presuponer que las variables son independientes, pero medidas con diferentes escalas.

Una vez escogida una métrica M para el ACP, que puede ser tan general como se desee, podemos echar una mirada a los primeros planos principales, para encontrar dibujos como el mostrado en la Figura 1. Allí inmediatamente nos ponemos a hacer comparaciones: ciertos individuos se encuentran concentrados en algún sector del plano, muy cercanos a algunas variables, y muy alejados de otros grupos de individuos, etc.

Esa manera de medir cercanías y lejanías entre los individuos evidentemente la sopesamos tomando además en consideración la importancia relativa del eje o componente que estemos mirando. El primer eje tiene mayor importancia que los restantes, de forma que es de suponer que las diferencias de los individuos en el primer eje son posiblemente más importantes que las correspondientes diferencias en los otros ejes. Esto al menos es cierto si lo que nos interesa es visualizar agrupaciones de individuos, para intentar formular algunas hipótesis posteriormente.

Cuando hacemos este tipo de visualizaciones —cosa común en todos los análisis de ACP— estamos en realidad introduciendo una nueva métrica, o manera de medir las distancias entre los individuos. En general, cualquier escogencia de la métrica M nos lleva, luego de realizar el ACP, a la definición de una *métrica inducida* correspondiente, d_M , la cual expresa la cercanía o lejanía entre los individuos, al ser mirados en una representación del ACP, tomando alguna consideración de la importancia relativa de cada eje que estemos mirando. Esta métrica inducida usualmente se define de la manera siguiente:

$$d_M(\underline{x}_i, \underline{x}_j) = \sum_{k=1}^p \lambda_k (\hat{x}_i^k - \hat{x}_j^k)^2$$

donde:

- M es la métrica empleada en el ACP.
- λ_k es la inercia explicada por el k -ésimo eje principal Δ_{u_k} del ACP.
- \hat{x}_i^k y \hat{x}_j^k son la coordenadas de los individuos \underline{x}_i y \underline{x}_j proyectados sobre el k -ésimo eje principal Δ_{u_k} .

En la definición anterior empleamos las inercias explicadas por cada eje, λ_k , como una forma de ponderar la importancia relativa del eje k -ésimo. Sin embargo, surgen ahora varias otras posibilidades de definición de métricas inducidas, utilizando potencias o radicales de λ_k , en vez de λ_k .

Por ejemplo:

$$d_M(\underline{x}_i, \underline{x}_j) = \sum_{k=1}^p \lambda_k^r (\hat{x}_i^k - \hat{x}_j^k)^2,$$

donde $r > -2$ es un número real. Todavía más general, podemos definir la métrica inducida por un Análisis en Componentes Principales, $d_{(M,f)}$, mediante:

$$d_{(M,f)}(\underline{x}_i, \underline{x}_j) = \sum_{k=1}^p f(\lambda_k) (\hat{x}_i^k - \hat{x}_j^k)^2, \quad (1)$$

donde f es una función no negativa de variable real. En [Piz89] se estudió el caso particular en que f es la función identidad. Ahora abordaremos un caso más general.

4. Algunos resultados obtenidos

Un resultado elemental de la teoría de matrices establece que si f es una función analítica definida en un dominio abierto que contenga a todos los valores propios *distintos* $\lambda_1, \lambda_2, \dots, \lambda_s$ de una matriz cualquiera A , con multiplicidades m_1, m_2, \dots, m_s , entonces tiene sentido la expresión $f(A)$, matriz que se calcula como

$$f(A) = p(A),$$

donde p es cualquier polinomio que satisface $p^{(l)}(\lambda_k) = f^{(l)}(\lambda_k)$, para $k = 1, 2, \dots, s$, y $l = 0, 1, \dots, m_k - 1$. El valor de $f(A)$ no depende del polinomio específico escogido, siempre que éste cumpla con las condiciones descritas. Una manera rápida de hallar el polinomio p es mediante interpolación de Lagrange [Gol89, Zad63].

En nuestro contexto la matriz A será además definida positiva y M -simétrica,⁴ lo cual simplifica las cosas, pues todos sus valores propios son reales y positivos. Más aún, en tal caso A admite una descomposición espectral del tipo

$$A = R \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_p] R^t M, \quad (2)$$

donde R es una matriz cuyas columnas son los vectores propios M -normalizados correspondientes a los valores propios $\lambda_1, \dots, \lambda_p$. La matriz R es M -ortonormal.

Fácilmente se demuestra [Gol89, Zad63] que, para cualquier función analítica f definida en un dominio abierto que contenga a todos los valores propios λ_k de una matriz A como en (2), se cumple que

$$f(A) = R \text{diag}[f(\lambda_1), f(\lambda_2), \dots, f(\lambda_p)] R^t M, \quad (3)$$

esto es, $f(\lambda_k)$ es valor propio de $f(A)$, con los mismos vectores propios asociados que A .

Volviendo a nuestro problema de métricas inducidas, tenemos el siguiente resultado.

Teorema 1 *Sea f una función analítica no negativa definida en un dominio abierto que contenga a todos los valores propios de VM . Entonces la métrica inducida $d_{(M,f)}$ definida en (1) tiene matriz asociada $Mf(VM)$.*

⁴Una matriz A es M -simétrica si cumple con la condición $(MA)^t = MA$, o bien, $A^t M = MA$. Las principales propiedades de las matrices M -simétricas pueden consultarse en [Pag76], páginas 152-153.

Demostración: Se utilizará simplemente el hecho de que $f(\lambda_k)$ es valor propio de $f(VM)$, con vector propio asociado \underline{u}_k , como se explicó anteriormente en (3). En efecto, de la definición (1) se obtiene directamente,

$$\begin{aligned}
d_{(M,f)}(\underline{x}_i, \underline{x}_j) &= \sum_{k=1}^p f(\lambda_k) (\hat{x}_i^k - \hat{x}_j^k)^2 \\
&= \sum_{k=1}^p f(\lambda_k) (\underline{x}_i^t M \underline{u}_k - \underline{x}_j^t M \underline{u}_k)^2 \\
&= \sum_{k=1}^p f(\lambda_k) (\underline{x}_i - \underline{x}_j)^t M \underline{u}_k \underline{u}_k^t M (\underline{x}_i - \underline{x}_j) \\
&= (\underline{x}_i - \underline{x}_j)^t M \left[\sum_{k=1}^p f(\lambda_k) \underline{u}_k \underline{u}_k^t \right] M (\underline{x}_i - \underline{x}_j) \\
&= (\underline{x}_i - \underline{x}_j)^t M \left[\sum_{k=1}^p f(VM) \underline{u}_k \underline{u}_k^t \right] M (\underline{x}_i - \underline{x}_j) \\
&= (\underline{x}_i - \underline{x}_j)^t M f(VM) U U^t M (\underline{x}_i - \underline{x}_j),
\end{aligned}$$

donde $U = (\underline{u}_1, \underline{u}_2, \dots, \underline{u}_p)$. Como los vectores propios \underline{u}_k son M -ortonormados, entonces

$$\forall i, \forall j : \quad \underline{u}_i^t M \underline{u}_j = \delta_{ij}$$

de donde $U^t M U = I_p$. Luego se obtiene que $M = (U^t)^{-1} U^{-1} = (U U^t)^{-1}$, lo cual lleva a la simplificación $M f(VM) U U^t M = M f(VM)$, de donde finalmente se obtiene

$$d_M(\underline{x}_i, \underline{x}_j) = (\underline{x}_i - \underline{x}_j)^t M f(VM) (\underline{x}_i - \underline{x}_j). \quad \blacksquare$$

Del teorema anterior se deduce que el único caso en que la métrica inducida $d_{(M,f)}$ coincide con la métrica inicial M es cuando f es la función constante 1, esto es, del todo no ponderamos los ejes principales por ningún factor. En todos los demás casos, se induce una métrica distinta $M f(VM)$, de la cual se obtendrán conclusiones interesantes.

Surge la inquietud de estudiar qué sucede si ahora realizamos el ACP utilizando la métrica inducida $Q_1 := M f(VM)$ en vez de $Q_0 := M$, debido a que Q_1 es en realidad la que empleamos en las interpretaciones de los planos principales. Obtenemos una nueva métrica inducida $Q_2 := M f(VM) f(VM f(VM))$. Y si continuamos aplicando esta métrica para realizar otro nuevo ACP, obtendríamos ahora la nueva métrica inducida

$$Q_3 := M f(VM) f(VM f(VM)) f(VM f(VM) f(VM f(VM))),$$

cuya fórmula es cada vez más complicada, pero la métrica es cada vez más adecuada para los fines de clasificación de los individuos. De inmediato nos preguntamos si esta sucesión de ACP's es convergente hacia alguna *métrica estable*, cuya interpretación fuese sin duda la "mejor" métrica posible para efectos de clasificación de objetos.

Hay una respuesta por lo menos para el caso sencillo en que f tiene una forma simple: $f(z) = z^r$, donde $r > -2$ es un número real. Para esta función simple, la fórmula (2) es aplicable aún en los casos en que r no sea entero. La solución a este problema está asociada a la famosa métrica de Mahalanobis, $M = V^{-1}$, la cual tiene varias propiedades interesantes en Análisis de Datos [Pag76].

$$\begin{array}{ccccc}
E & \xleftarrow{X} & F^* & \xrightarrow{X} & E \\
M=Q_0 \downarrow \uparrow V & & \uparrow D_\omega & & V \uparrow \downarrow Q_1 \\
E^* & \xrightarrow{X^t} & F & \xleftarrow{X^t} & E^* \\
Q_2 \uparrow \downarrow V & & \downarrow D_\omega & & V \downarrow \uparrow Q_3 \\
E & \xleftarrow{X} & F^* & \xrightarrow{X} & E \\
Q_4 \downarrow \uparrow V & & \uparrow D_\omega & & V \uparrow \downarrow Q_5 \\
E^* & \xrightarrow{X^t} & F & \xleftarrow{X^t} & E^* \\
\vdots & & \vdots & & \vdots
\end{array}$$

Figura 2: Sucesión de esquemas de dualidad de ACP's, correspondientes a la métrica inicial $Q_0 = M$ y una función f que pondera la importancia relativa de cada eje.

Teorema 2 Sea $f(z) = z^r$, donde $r > -2$ es un número real. Considérese la sucesión de aplicaciones del ACP como en la Figura 2, a partir de una métrica inicial M . Entonces las métricas inducidas Q_1, Q_2, Q_3, \dots , tienen la forma matricial

$$Q_m = M (VM)^{(r+1)^m - 1}, \quad m \in \mathbf{N}.$$

Además, pueden distinguirse tres casos:

- (a) Si $-2 < r < 0$, el proceso siempre es convergente hacia la métrica de Mahalanobis, V^{-1} , independientemente de M .
- (b) Si $r = 0$ (caso trivial) el proceso siempre produce la métrica constante M .
- (c) Si $r > 0$ una condición necesaria y suficiente para que el proceso sea convergente a una métrica es que M sea la métrica de Mahalanobis.

Demostración: Para establecer la primera parte, calculamos los primeros términos:

$$\begin{aligned}
Q_1 &= M f(VM) = M (VM)^r \\
Q_2 &= d_{(Q_1, f)} = Q_1 (VQ_1)^r = M (VM)^{r^2 + 2r} = M (VM)^{(r+1)^2 - 1} \\
Q_3 &= d_{(Q_2, f)} = Q_2 (VQ_2)^r = M (VM)^{r^3 + 3r^2 + 3r} = M (VM)^{(r+1)^3 - 1}
\end{aligned}$$

y fácilmente se demuestra por inducción matemática la fórmula general para Q_m , con $m \in \mathbf{N}$:

$$\begin{aligned}
Q_m &= M (VM)^{(r+1)^m - 1} \\
&= M R \operatorname{diag}[\lambda_1^{(r+1)^m - 1}, \dots, \lambda_p^{(r+1)^m - 1}] R^t M,
\end{aligned}$$

en donde R es la matriz M -ortonormal de la descomposición espectral de VM , esto es, $VM = R \operatorname{diag}[\lambda_1, \dots, \lambda_p] R^t M$, como en (2), y estamos haciendo uso directo de (3).

Ahora, en el caso (a) tenemos que $-2 < r < 0$, de donde $|r + 1| < 1$ y entonces,

$$\begin{aligned}
\lim_{m \rightarrow \infty} Q_m &= \lim_{m \rightarrow \infty} M R \operatorname{diag}[\lambda_1^{(r+1)^m - 1}, \dots, \lambda_p^{(r+1)^m - 1}] R^t M \\
&= M R \operatorname{diag}[\lambda_1^{-1}, \dots, \lambda_p^{-1}] R^t M \\
&= M (VM)^{-1} = V^{-1},
\end{aligned}$$

obteniéndose la métrica de Mahalanobis. El caso (b) es evidente. En el caso (c), en el cual $r > 0$, el exponente $(r + 1)^m - 1$ tiende a $+\infty$, de manera que la única posibilidad de convergencia de la sucesión de métricas (Q_m) ocurre cuando $0 < \lambda_k \leq 1$, para $k = 1, 2, \dots, p$. En tal caso obtenemos

$$\begin{aligned} \lim_{m \rightarrow \infty} Q_m &= \lim_{m \rightarrow \infty} M R \text{diag}[\lambda_1^{(r+1)^m - 1}, \dots, \lambda_p^{(r+1)^m - 1}] R^t M \\ &= M R \text{diag}[a_1, \dots, a_p] R^t M, \end{aligned}$$

donde los números $a_k \in \{0, 1\}$. Específicamente,

$$a_k = \begin{cases} 1 & , \text{ si } \lambda_k = 1 \\ 0 & , \text{ si } 0 < \lambda_k < 1. \end{cases}$$

Ahora, si alguno de los a_k fuese nulo, el proceso no converge a una métrica, pues la matriz límite no sería definida positiva. Luego es necesario que todos los valores propios λ_k de VM sean iguales a 1, de donde se concluye que $M = V^{-1}$. Es evidente que esta condición es además suficiente para garantizar la convergencia de (Q_m) . ■

5. Conclusiones

Uno de los problemas más difíciles y hasta cierto punto subjetivo que se presenta en la disciplina del Análisis de Datos es la escogencia de la métrica a utilizar en los estudios particulares.

Para cada tipo de análisis hay algunas métricas consideradas clásicas, cuyo empleo se recomienda generalmente a falta de otra opción específica que obedezca a alguna razón de importancia en el estudio realizado.

Se demostró en este trabajo que para una amplia gama de métricas inducidas mediante ACP's, la única métrica inicial M que conduce a un proceso estable de métricas es la métrica de Mahalanobis $M = V^{-1}$. Para otra amplia gama de métricas inducidas el proceso en sí converge hacia la métrica de Mahalanobis $M = V^{-1}$, independientemente de la escogencia inicial de M .

Los resultados obtenidos en este trabajo ponen en evidencia que, para aquellos estudios exploratorios con fines de búsqueda de tipologías en los individuos, a partir de una tabla de datos de individuos contra variables, es innecesario realizar primero un ACP, pues es mejor opción trabajar directamente con la métrica de Mahalanobis, $M = V^{-1}$.

En efecto, la métrica de Mahalanobis $M = V^{-1}$ conduce a procesos estables y permite una mejor diferenciación de los individuos, poniendo en evidencia las relaciones de cercanía y lejanía entre ellos. Con esta métrica es innecesario realizar un ACP, pues en este caso VM sería igual a la matriz identidad.

Esto simplifica al menos parcialmente una vieja costumbre arraigada en los investigadores que buscan clasificar sus datos, cual es la de realizar primero un Análisis en Componentes Principales y luego emplear técnicas de Clasificación Automática.

Referencias

- [Gol89] Golub, G.; Van Loan, C. (1989) *Matrix Computations*, Second Edition. The John Hopkins University Press, Baltimore.
- [Pag76] Cailliez, F.; Pages, J. P. (1976). *Introduction à l'Analyse des Données*. Smash, Paris.
- [Piz89] Piza, E. (1989). "La métrica inducida por un Análisis en Componentes Principales", *Revista Ciencia y Tecnología*, **XIII** (1-2): 167-179.

- [Pol89] Poltronieri, J.; Piza, E. (1989) *Estructuras de la Opinión Pública en Costa Rica*. Editorial de la Universidad de Costa Rica, San José.
- [Tak82] Takeuchi, K.; Yanai, H.; Mukherjee, B.Ñ. (1982) *The Foundations of Multivariate Analysis*. Wiley Eastern Limited, New Delhi.
- [Zad63] Zadeh, L.; Desoer, C. (1963) *Linear System Theory: the State Space Approach*. McGraw-Hill, New York.