Revista de Matemática: Teoría y Aplicaciones 2004 **11**(2) : 35–53
cimpa – ucr – ccss issn: 1409-2433

DEFINICIÓN DE SEMIPRODUCTOS ESCALARES ÚTILES EN ANÁLISIS DE DATOS

Javier Trejos–Zelaya*

Recibido/Received: 11 Sept 2003

Abstract

Se desarrolla la teoría necesaria para realizar el Análisis de Datos en presencia de semiproductos escalares, extendiendo los conceptos clásicos de productos escalares usualmente empleados. Para ello, retomamos las definiciones algebraicas básicas de las formas bilineales no degeneradas y vamos desarrollando todas las herramientas algebraicas necesarias. Se estudian los operadores más importantes en el espacio de individuos, como el operados VM y el operador MV. También se estudia el caso del semiproducto escalar de pesos en el espacio de variables, que en el caso de pesos nulos corresponde a la introducción de individuos suplementarios. Finalmente, llegamos a los conceptos usuales del Análisis en Componentes Principales.

Palabras clave: semiproductos escalares, formas bilineales no degeneradas, semimétricas, operador de proyección ortogonal, análisis en componentes principales.

Abstract

We develop the theory necessary for Data Analysis with inner semiproducts, extending tha classical concepts of inner products usually employed. For this, we use the basic algebraic definitions of non degenerated bilinear forms and develop all the algebraic tools needed. We study the most important operators on the individual space, such as the VM and the MV operators. We also study the case of the inner semiproduct of weights in the variable space, which corresponds to the introduction of supplementary individuals in the case of null weights. Finally, we arrive to the usual concepts of Principal Component Analysis.

Keywords: inner semiproducts, non degenerated bilinear forms, semimetrics, orthogonal projection operator, principal component analysis.

Mathematics Subject Classification: 11E39, 15A09, 15A63, 62H25

*CIMPA, Escuela de Matemática, Universidad de Costa Rica, 2060 San José, Costa Rica. E-Mail: `jtrejos@cariari.ucr.ac.cr`

1 Introducción

Em algunos métodos de Análisis de Datos, se encuentran dificultades para definir una distancia en el espacio de los individuos. Este es el caso en los análisis canónicos cuando las variables están muy correlacionadas, pues la matriz de covarianzas es entonces numéricamente no invertible, y por ello es difícil calcular la matriz de una distancia de Mahalanobis. Hemos encontrado dificultades similares para definir una distancia entre conjunciones de modalidades esplicativas [12, 14, 13]. Hemos por lo tanto estudiado, en una primera parte, el uso de semiproductos escalares en Análisis de Datos. En una segunda parte, estudiamos algunas propiedades —en este contexto de semiproductos escalares— de dos operadores útiles en Análisis de Datos. Finalmente, abordamos la definición de algunos conceptos clásicos del Análisis de Datos, como los relacionados con la dispersión de una nube y el Análisis en Componentes Principales. Hacemos notar que estos temas los hemos abordado en otras publicaciones de circulación restringida [11, 14, 13].

2 Algunos resultados sobre los semiproductos escalares

Sean L un espacio vectorial de dimensión finita, L^* su espacio dual, H un subespacio de L y T una forma bilineal simétrica sobre L. Se denota con la misma letra T la aplicación lineal de L en L^* associada a T. $T_{/H}$ es la restricción de la aplicación T a H.

Proposición 1 *Si T es una forma bilineal simétrica sobre L, T la aplicación lineal asociada y H un subespacio de L, se tiene:*

a) *Si T es no degenerada sobre H entonces la restricción $T_{/H}$ de T en H es inyectiva.*

b) *Si $T_{/H}$ es inyectiva y si* $\operatorname{rang} T = \dim H$ *entonces T es no degenerada sobre H.*

c) *Si $T_{/H}$ es inyectiva y si T es positiva entonces T es no degenerada sobre H.*

Demostración:

a) Sea $x \in \ker T_{/H}$. $\forall y \in H$ se tiene $T(x,y) = T(x)(y) = T_{/H}(x)(y) = 0$. Por hipótesis, T es no degenerada sobre H, de donde se tiene $\{z \in H / \forall y \in H, T(y,z) = 0\} = \{0\}$, luego $x = 0$ y $T_{/H}$ es inyectiva.

b) Se tiene $\operatorname{Im} T_{/H} \subset \operatorname{Im} T$. Como por hipótesis $T_{/H}$ es inyectiva y $\operatorname{rang} T = \dim H$, se tiene $\dim(\operatorname{Im} T_{/H}) = \dim H = \operatorname{rang} T$. Luego $\operatorname{Im} T_{/H} = \operatorname{Im} T$, es decir $\{T(y)/y \in H\} = \{T(y)/y \in L\}$.
Sea $x \in H$ tal que $\forall y \in H, T(x,y) = T(y)(x) = 0$.
Se deduce que $\forall y \in L, T(x)(y) = T(y)(x) = 0$, es decir, $T(x) = 0$ y entonces $x \in H \cap \ker T = \ker T_{/H}$.
Como $T_{/H}$ es inyectiva se tiene $x = 0$ y T es por tanto no degenerada sobre H.

c) Sea $x \in H$ tal que $\forall y \in H, T(x,y) = 0$. Como T es positiva por hipótesis, se tiene $\ker T = \{z \in L / T(z,z) = 0\}$, de donde $x \in H \cap \ker T = \ker T_{/H}$.
En vista de que $T_{/H}$ es inyectiva se tiene $x = 0$ y luego T es no degenerada sobre H.
∎

Nótese que la hipótesis rang $T = \dim H$ del inciso (b) de la proposición anterior podría ser reemplazada por rang $T_{/H} =$ rang T.

Proposición 2 *Si H y G son dos subespacios vectoriales de L, con ortogonales respectivos H° y G° en L^*, se tiene:*

$$L = H \oplus G \Leftrightarrow L^* = H^\circ \oplus G^\circ.$$

DEMOSTRACIÓN: Supńgase que $L = H \oplus G$. Sea $f \in H^\circ \cap G^\circ$, $\forall z \in L \;\; z = x + y$ donde $x \in H$ y $y \in G$, $f(z) = f(x+y) = f(x) + f(y) = 0$, de donde se deduce que $f = 0$ y $H^\circ \cap G^\circ = \{0\}$. Por consiguiente $L^* = H^\circ \oplus G^\circ$ puesto que $\dim H^\circ = \dim L - \dim H$, $\dim G^\circ = \dim L - \dim G$ y $\dim L = \dim L^*$.
El recíproco es immediato al identificar L^{**} y L, puesto que $H^{\circ\circ} = H, G^{\circ\circ} = G$. ∎

Denotando $\tilde{H}$ (respectivamente $\tilde{G}$) el espacio vectorial canónicamente isomorfo a H (resp. G), si $L = H \oplus G$ entonces G° y H° son respectivamente isomorfos a los espacios duales $\tilde{H}^*$ y $\tilde{G}^*$.

En efecto, se pueden definir dos aplicaciones ϕ y ψ sobre L^*, $\phi : L^* \to \tilde{H}^*$ y $\psi : L^* \to \tilde{G}^*$, cuyas restricciones respectivas a los subespacios G° y H° son tales que:

- $\forall y^\circ \in G^\circ$ y $\forall \tilde{x} \in \tilde{H}$, $\phi_{/G^\circ}(y^\circ)(\tilde{x}) = \phi(y^\circ)(\tilde{x}) = y^\circ(x)$, donde x es la imagen en H de $\tilde{x}$ por la inyección canónica de $\tilde{H}$ en H.
- $\forall x^\circ \in H^\circ$ y $\forall \tilde{y} \in \tilde{G}$, $\psi_{/H^\circ}(x^\circ)(\tilde{y}) = \psi(x^\circ)(\tilde{y}) = x^\circ(y)$, donde y es la imagen en G de $\tilde{y}$ por la inyección canónica de $\tilde{G}$ en G.

$\phi_{/G^\circ}$ y $\psi_{/H^\circ}$ son evidentemente lineales. Además, son biyectivas. En efecto, se tiene para $\phi_{/G^\circ}$:

- $\dim \tilde{H}^* = \dim \tilde{H} = \dim H = \dim G^\circ$;
- Sea $y^\circ \in \ker \phi_{/G^\circ} = G^\circ \cap \ker \phi$, $\forall x \in H \;\; y^\circ(x) = \phi_{/G^\circ}(y^\circ)(\tilde{x}) = 0$, luego $y^\circ \in H^\circ$. Como $H^\circ \cap G^\circ = \{0\}$, se tiene $y^\circ = 0$ y entonces $\ker \phi_{/G^\circ} = \{0\}$.

Es análogo para $\psi_{/H^\circ}$.

Para la descomposición $L = H \oplus G$, es entonces natural llamar H° (resp. G°) al **subespacio dual** de G (resp. H) y denotarlo G^* (resp. H^*). Diremos además que $L = H \oplus G$ y $L^* = H^* \oplus G^*$ son dos **descomposiciones duales** de L y L^*.

Observaciones:

1) Si $\{h_1, \ldots, h_p, g_1, \ldots, g_q\}$ es una base de L tal que $\forall i \;\; h_i \in H$ y $\forall j \;\; g_j \in G$, cuya base dual en L^* es $\{h_1^*, \ldots, h_p^*, g_1^*, \ldots, g_q^*\}$, se muestra fácilmente que $\{h_1^*, \ldots, h_p^*\}$ es una base de G° y que $\{g_1^*, \ldots, g_q^*\}$ es una base de H°.
2) Más generalmente, $L = \oplus H_s$ y $L^* = \oplus H_s^*$ son dos descomposiciones duales de L y L^* si y sólo si $H_s^* = (\bigoplus_{r \neq s} H_r)^\circ = \bigcap_{r \neq s} H_r^\circ$.

Proposición 3 *Las afirmaciones siguientes son equivalentes:*

a) *T es no degenerada sobre H y* $\text{rang } T = \dim H$.

b) *$T_{/H}$ es inyectiva y* $\text{rang } T = \dim H$.

c) $L = H \oplus \ker T$.

d) $L^* = \text{Im } T \oplus H^\circ$.

DEMOSTRACIÓN: La equivalencia entre (a) y (b) proviene inmediatamente de la propiedad 1.

(b) $\Leftrightarrow$ (c): Como $\dim L = \text{rang } T + \dim(\ker T)$, (b) es equivalente a: $H \cap \ker T = \{0\}$ y $\dim L = \dim H + \dim(\ker T)$, es decir equivalente a (c).

(c) $\Leftrightarrow$ (d): Identificando L^{**} y L, puesto que la forma bilineal T es simétrica, se tiene: $(\ker T)^\circ = \text{Im } {}^tT = \text{Im } T$. La equivalence se establece aplicando la propiedad 2. ∎

Se puede notar que las afirmaciones de la proposición 3 también son equivalentes a la afirmación: $T_{/H}$ es inyectiva y $\text{rang } T_{/H} = \text{rang } T$.

Corolario 1 *Si la forma bilineal simétrica T sobre L es no degenerada sobre un subespacio $H \subset L$, y si* $\text{rang } T = \dim H$ *entonces:*

a) *$L = H \oplus \ker T$ y $L^* = \text{Im } T \oplus H^\circ$ son dos descomposiciones duales de L y L^*, con $H^* = \text{Im } T$ y $(\ker T)^* = H^\circ$.*

b) *Se tiene además* $\ker T = H^\perp$ *y* $\text{Im } T = \{T(x)/x \in H\}$.

DEMOSTRACIÓN:

a) La afirmación (a) es una consecuencia directa de la proposición 3 y de la definición de las descomposiciones duales.

b) Se tiene $\ker T \subset H^\perp$. Ahora bien, como por hipótesis T es no degenerada sobre H, se tiene: $\dim L = \dim H + \dim H^\perp = \text{rang } T + \dim H^\perp$ puesto que $\text{rang } T = \dim H$, de donde $\dim H^\perp = \dim(\ker T)$ y luego $\ker T = H^\perp$.
Como $T_{/H}$ es inyectiva y $\text{rang } T = \dim H$, se tiene: $\text{Im } T = \text{Im } T_{/H} = \{T(x)/x \in H\}$. ∎

2.1 Aplicación al Análisis de Datos

En Análisis de Datos, se dispone de una tabla de datos que contiene la medida de p variables observadas sobre n individuos o unidades estadísticas. A esta tabla le corresponde una matriz con n filas y p columnas, denotada X. Al i-ésimo individuo (resp. j-'esima variable) se le asocia la fila i (resp. la columna j) de X, denotada x_i (resp. x^j). El elemento (i,j) de X, denotado x_i^j, es el valor de x^j observado sobre el individuo x_i. Así, x_i pertenece a un espacio vectorial real de p dimensiones denotado E y llamado **espacio de individuos**. Igualmente, x^j pertenece a un espacio vectorial real de n dimensiones F llamado **espacio de variables**.

La matriz X está asociada canónicamente a una aplicación lineal $X : E^* \longrightarrow F$ tal que $X(e_j^*) = x^j$, donde E^* es el dual de E y $\{e_j^*/j = 1, \ldots, p\}$ es la base dual de la base canónica $\{e_j/j = 1, \ldots, p\}$ de E. La transpuesta de X, denotada ${}^tX$, es una aplicación de F^* en $E = E^{**}$ tal que ${}^tX(f_i^*) = x_i$, donde F^* es el dual de F y $\{f_i^*\}$ es la base dual de la base canónica $\{f_i\}$ de F.

M es una forma bilineal simétrica, no degenerada sobre $\mathrm{Im}\ {}^tX$, que supondremos positiva y será utilizada posteriormente para definir una distancia entre individuos.

Como $(\mathrm{Im}\ {}^tX)^\circ = \ker X$, aplicando el corolario 1 se deduce:

Proposición 4 *Si la forma bilineal simétrica M, no degenerada sobre $\mathrm{Im}\ {}^tX$, es tal que $\mathrm{rang}\ M = \mathrm{rang}\ X$, entonces*

$$E = \mathrm{Im}\ {}^tX \oplus \ker M \ \ y \ \ E^* = \mathrm{Im}\ M \oplus \ker X$$

son dos descomposiciones duales de E y E^* tales que $(\mathrm{Im}\ {}^tX)^* = \mathrm{Im}\ M = \{M(x)/x \in \mathrm{Im}\ {}^tX\}$ y $(\ker M)^* = \ker X$.

2.2 La forma bilineal de covarianzas V

Denotamos con la misma letra D una forma bilineal simétrica sobre F y la aplicación lineal de F en F^* asociada.

Denotamos V una aplicación de $E^* \times E^*$ en $\mathbb{R}$ definida por

$$\forall (x^*, y^*) \in E^* \times E^* \quad V(x^*, y^*) = D(Xx^*, Xy^*).$$

Es fácil verificar que V es una forma bilineal simétrica sobre E^* y que ${}^tXDX$ es una expresión de la aplicación lineal de E^* en $E^{**} = E$, denotada V, asociada. Es más, V es positiva si D es positiva.

Se sabe que si la matriz asociada a D en la base canónica de F es la matriz diagonal de pesos de los individuos y si las variables son centradas con respecto a D, entonces la matriz asociada a V en la base canónica de E^* es la matriz de covarianzas de las variables. V es llamada la **forma bilineal de covarianzas**.

Recordamos a manera de ilustración el esquema de dualidad, introducido en [1], en la figura 1.

$$\begin{tikzcd} E \arrow[d, shift right, "M"'] & F^* \arrow[l, "{}^tX"'] \\ E^* \arrow[u, shift right, "V"'] \arrow[r, "X"'] & F \arrow[u, "D"'] \end{tikzcd}$$

Figura 1: *El esquema de dualidad.*

Proposición 5 **a)** *Si D es no degenerada sobre* $\mathrm{Im}\ X$*, entonces* $\ker V = \ker X$.

b) $\ker V = \ker X$ *es equivalente a* $\mathrm{Im}\ V = \mathrm{Im}\ {}^tX$.

c) *Si* $\ker V = \ker X$ *entonces la restricción de D a* $\mathrm{Im}\ X$*, denotada* $D_{/\mathrm{Im}\ X}$*, es inyectiva.*

d) *Si D es positiva o si* $\mathrm{rang}\ D = \mathrm{rang}\ X$*, entonces una condición necesaria y suficiente para que D sea no degenerada sobre* $\mathrm{Im}\ X$ *es que* $\ker V = \ker X$.

Demostración:

a) Se tiene $\ker V \supset \ker X$. Mostremos que $\ker V \subset \ker X$: Sea $x^* \in \ker V$, $\forall y^* \in E^*$ se tiene: $D(Xx^*, Xy^*) = \langle {}^tXDXx^*, y^*\rangle = \langle Vx^*, y^*\rangle = 0$.
Como por hipótesis D es no degenerada sobre $\mathrm{Im}\, X$, se deduce que $X(x^*) = 0$, es decir $x^* \in \ker X$ y entonces $\ker V \subset \ker X$.

b) Se deduce la equivalencia propuesta de:
- $\mathrm{Im}\ V \subset \mathrm{Im}\ {}^tX$ y $\ker V \supset \ker X$, y
- $\dim E^* = \mathrm{rang}\ V + \dim(\ker V) = \mathrm{rang}\ {}^tX + \dim(\ker X)$.

c) Sea $x \in \ker D_{/\mathrm{Im}\ X} = \mathrm{Im}\ X \cap \ker D$. $\exists y^* \in E^*$ tal que $x = X(y^*)$. Se tiene entonces $V(y^*) ={}^t XDX(y^*) ={}^t XD(x) = 0$, de donde $y^* \in \ker V = \ker X$. Por consiguiente, $x = X(y^*) = 0$ y $\ker D_{/\mathrm{Im}\ X} = \{0\}$.

d) La condición necesaria ha sido establecida en (a), mostremos la condición suficiente. Si se supone que $\ker V = \ker X$ entonces $D_{/\mathrm{Im}\ X}$ es inyectiva. Como D es positiva o $\mathrm{rang}\ D = \mathrm{rang}\ X = \dim(\mathrm{Im}\ X)$, de la proposición 1 se deduce que D es no degenerada sobre $\mathrm{Im}\ X$. ■

Proposición 6 *Si D es no degenerada sobre* $\mathrm{Im}\ X$*, entonces V es no degenerada sobre todo subespacio suplementario de* $\ker X$.

Demostración: Sea S un subespacio suplementario de $\ker X$ en E^*. Se quiere probar que V es no degenerada sobre S, es decir que $\{x^* \in S / \forall y^* \in S, V(x^*, y^*)\} = 0$.
Sea $x^* \in S$ tal que $\forall y^* \in S \quad V(x^*, y^*) = 0$. Como D es no degenerada sobre $\mathrm{Im}\ X$ se tiene $\mathrm{Im}\ V = \mathrm{Im}\ {}^tX$, de donde se deduce que $\forall z^* \in F^*, \exists y^* \in S$ tal que ${}^tX(z^*) = V(y^*)$. Se tiene entonces $\forall z^* \in F^*$:

$$\langle Xx^*, z^*\rangle = \langle x^*, {}^tXz^*\rangle = \langle x^*, Vy^*\rangle = V(x^*, y^*) = 0,$$

lo que implica que $X(x^*) = 0$ o sea $x^* \in \ker X$. Como $S \cap \ker X = \{0\}$, se tiene $x^* = 0$ y V es entonces no degenerada sobre S. ■

En lo que sigue, diremos que T es un **semiproducto escalar no degenerado sobre** H si T es una forma bilineal sobre L, simétrica, positiva y no degenerada sobre H. Evidentemente, se trata de un abuso de lenguaje, pero que tiene el fin de simplificar los términos usados.

Corolario 2 *Si M y D son semiproductos escalares no degenerados respectivamente sobre los subespacios* $\mathrm{Im}\ {}^tX$ *y* $\mathrm{Im}\ X$*, y si* $\mathrm{rang}\ M = \mathrm{rang}\ V$*, entonces*

a) *V es un semiproducto escalar no degenerado sobre* $\mathrm{Im}\ M$.

b) $E = \mathrm{Im}\ V \oplus \ker M$ *y* $E^* = \mathrm{Im}\ M \oplus \ker V$ *son dos descomposiciones duales de E y E^*, donde* $(\mathrm{Im}\ V)^* = \mathrm{Im}\ M$ *y* $(\ker M)^* = \ker V$.

Demostración:

a) Hemos visto que por definición V es bilineal, simétrica y positiva. Mostremos que V es no degenerada sobre $\mathrm{Im}\ M$. Como por hipótesis D es no degenerada sobre $\mathrm{Im}\ X$ se tiene (c.f. proposición 5) $\ker V = \ker X$ y luego $\mathrm{rang}\ M = \mathrm{rang}\ V = \mathrm{rang}\ X$. Por la proposición 4, $\mathrm{Im}\ M$ es un suplementario de $\ker X$ en E^*, y entonces V es no degenerada sobre $\mathrm{Im}\ M$, según la proposición anterior.

b) Por la proposición 5 se tiene $\ker V = \ker X$ y $\mathrm{Im}\ V = \mathrm{Im}\ {}^tX$, el resultado se establece entonces aplicando la proposición 4. ∎

2.3 Inversas generalizadas de V y del semiproducto escalar M del espacio de individuos

SZe supone que M (resp.D) es un semiproducto escalar no degenerado sobre $\mathrm{Im}\ {}^tX$ (resp. $\mathrm{Im}\ X$) y que $\mathrm{rang}\ M = \mathrm{rang}\ X$.

Recordamos que una inversa generalizada algebraica, denotada B^-, de una aplicación lineal $B : L_1 \longrightarrow L_2$, con L_1 y L_2 espacios vectoriales, es una aplicación lineal de L_2 en L_1 tal que $B^-BB^- = B^-$ y $BB^-B = B$.
B^- está caracterizada por $\mathrm{Im}\ B^-$ y $\ker B^-$, donde estos subespacios son respectivamente los suplementarios de $\ker B$ en L_1 y $\mathrm{Im}\ B$ en L_2.

Por el corolario 2, $E = \mathrm{Im}\ V \oplus \ker M$ y $E^* = \mathrm{Im}\ M \oplus kerV$ son dos descomposiciones duales de E y E^*. Una inversa generalizada algebraica M^- de M puede entonces ser caracterizada por $\mathrm{Im}\ M^- = \mathrm{Im}\ V$ y $\ker M^- = \ker V$. Notando que la restricción de M a $\mathrm{Im}\ V$, denotada $M_{/\mathrm{Im}\ V}$, es una biyección de $\mathrm{Im}\ V$ en $\mathrm{Im}\ M$, se muestra [6] fácilmente (idempotencia de MM^- o de M^-M) la propiedad siguiente:

Proposición 7 *La inversa generalizada algebraica de M, especificada por* $\mathrm{Im}\ M^- = \mathrm{Im}\ V$ *y* $\ker M^- = \ker V$*, es tal que*

$$\forall x^* \in \mathrm{Im}\ M \quad M^-(x^*) = M^{-1}_{/\mathrm{Im}\ V}(x^*). \quad \blacksquare$$

Entonces $E = \mathrm{Im}\ M^- \oplus \ker M$ y $E^* = \mathrm{Im}\ M \oplus \ker M^-$ son dos descomposiciones duales de E y E^* tales que $(\mathrm{Im}\ M^-)^* = \mathrm{Im}\ M$ y $(\ker M)^* = \ker M^-$.

Notemos que:

$$\mathrm{Im}\ M^- = \{M^-(x^*)/x^* \in \mathrm{Im}\ M\} = \mathrm{Im}\ V = \{V(x^*)/x^* \in \mathrm{Im}\ M\}.$$

Proposición 8
La forma bilineal, denotada M^-*, definida por*

$$\forall (x^*, y^*) \in E^* \times E^* \quad M^-(x^*, y^*) = \langle x^*, M^- y^* \rangle$$

es un semiproducto escalar no degenerado sobre Im M*.*

Demostración: $\forall (x^*, y^*) \in E^* \times E^*, x^* = x_1^* + x_2^*$ y $y^* = y_1^* + y_2^*$, donde $x_1^*, y_1^* \in \text{Im } M$ y $x_2^*, y_2^* \in \ker M^-$, se tiene:

$$\begin{aligned} M^-(x^*, y^*) &= \langle x^*, M^- y^* \rangle = \langle x_1^* + x_2^*, M^- y_1^* \rangle \\ &= \langle M M^- x_1^*, M^- y_1^* \rangle + \langle x_2^*, M^- y_1^* \rangle. \end{aligned}$$

Como $x_2^* \in \ker M^- = \ker V = (\text{Im } V)^\circ$ y $\text{Im } M^- = \text{Im } V$, se tiene $\langle x_2^*, M^- y_1^* \rangle = 0$ y entonces:

$$M^-(x^*, y^*) = M(M^- x_1^*, M^- y_1^*).$$

Como M es simétrica y positiva, también lo es M^-.
Siendo M^- positiva y la restricción de M^- a Im M inyectiva (c.f. proposición 7), por la proposición 1 se obtiene que M^- es no degenerada sobre Im M. ■

Así mismo, la inversa generalizada algebraica V^- de V, caracterizada por $\ker V^- = \ker M$ y $\text{Im } V^- = \text{Im } M$, es tal que

$$\forall x \in \text{Im } V \quad V^-(x) = V^{-1}_{/\text{Im } M}(x).$$

$E = \text{Im } V \oplus \ker V^-$ y $E^* = \text{Im } V^- \oplus \ker V$ son entonces dos descomposiciones duales de E y E^* (c.f. corolario 2) tales que $(\text{Im } V)^* = \text{Im } V^-$ y $(\ker V^-)^* = \ker V$.
Es más, se tiene:

1) $\text{Im } V^- = \{V^-(x)/x \in \text{Im } V\} = \text{Im } M = \{M(x)/x \in \text{Im } V\}$.

2) La forma bilineal, denotada V^-, definida por:

$$\forall (x, y) \in E \times E \quad V^-(x, y) = \langle x, V^- y \rangle$$

es un semiproducto escalar no degenerado sobre Im V.

Definición 1 *Cuando* D *es el producto escalar de pesos se dice que* V^- *es un* **semiproducto escalar de Mahalanobis**.

3 Operadores útiles en Análisis de Datos

En las dos partes de esta sección se supone que M y D son semiproductos escalares no degenerados respectivamente sobre Im ${}^tX$ y Im X, y que rang M = rang X. Se tiene entonces también rang M = rang V (c.f. proposición 5).

3.1 El operador VM

En esta subsección estudiamos algunas propiedades del operador VM sobre E. Estas propiedades serán importantes más adelante donde se recordarán las definiciones de la inercia y de los ejes principales de dispersión de una nube de puntos. Así mismo, este operador será utilizado en expresiones de los productos escalares relacionales.

Recordemos algunas definiciones y propiedades generales en el caso en que L es un espacio vectorial de dimensión finita y T un semiproducto escalar no degenerado sobre el subespacio $H \subset L$. Diremos que el operador A sobre L es

- T-*simétrico* si $\forall (x,y) \in L \times L \;\; T(Ax,y) = T(x,Ay)$
- T-*positivo* si $\forall (x,x) \in L \times L \;\; T(Ax,x) \geq 0$.

Observación: En presencia de semiproductos escalares, no se garantiza la existencia y la unicidad de la adjunta de un operador A. Denotando T a la aplicación lineal de L en L^* asociada a T, una condición suficiente de existencia es que Im $({}^tAT) \subset$ Im T. En efecto, en este caso si T^- es una inversa generalizada interna de T, entonces:

$$\forall (x,y) \in L \times L \;\; \text{se tiene} \;\; T(Ax,y) = \langle x, {}^t ATy\rangle = \langle x, TT^- \, {}^tATy\rangle = T(x, T^- \, {}^tATy)$$

pues TT^- es un proyector sobre Im T, lo que prueba la existencia de la adjunta $A^* = T^- \, {}^tAT$ de A.
Notemos que VM satisface esta condición, pues Im ${}^t(VM)M =$ Im $MVM \subset$ Im M. ∎

Sea $P : L \longrightarrow$ Im $P \subset H \subset L$ una aplicación tal que $\forall x \in$ Im $P \;\; P(x) = x$ y $\ker P = (\text{Im } P)^\perp$. Notemos que $L = \text{Im } P \oplus (\text{Im } P)^\perp$ puesto que, $T_{/H}$ y por tanto $T_{/\text{Im } P}$ siendo inyectivas (c.f. proposición 1(a)) y T positiva, T es no degenerada sobre Im P (c.f. proposición 1(c)). Por lo tanto es natural llamar P el **operador de proyección ortogonal** sobre Im P respecto a T.

Proposición 9 *Como T es un semiproducto escalar no degenerado sobre H, las siguientes afirmaciones son equivalentes:*

a) *P es un operador de proyección ortogonal sobre* Im P *respecto a T*

b) Im $P \subset H$, *P es idempotente y T-simétrico.*

DEMOSTRACIÓN:

a) *Rightarrow* b)): Supongamos que P es un operador de proyección ortogonal. Se tiene Im $P \subset H$ por definición de P. P es idempotent pues $\forall x \in L, x = x_1 + x_2$ donde $x_1 \in$ Im P y $x_2 \in \ker P$, se tiene: $P^2(x) = P^2(x_1) = P(x_1) = P(x)$.
Mostremos que P es T-simétrico:
$\forall (x,y) \in L \times L, x = x_1 + x_2$ y $y = y_1 + y_2$ donde $x_1, y_1 \in$ Im P y $x_2, y_2 \in \ker P$, se tiene:

$$\begin{aligned} T(Px,y) &= T[P(x_1+x_2), y_1+y_2] \\ &= T(x_1, y_1+y_2) \\ &= T(x_1,y_1) \;\; \text{pues } \ker P = (\text{Im } P)^\perp \end{aligned}$$

e igualmente $T(x,Py) = T(x_1,y_1)$.
Se tiene entonces $T(Px,y) = T(x,Py)$.

b) *Rightarrow* a): $\forall x_1 \in \text{Im } P, \exists y \in L$ tal que $P(y) = x_1$. Como por hipótesis P es idempotente, se tiene $P(x_1) = PP(y) = P(y) = x_1$. Es más, $\forall x_2 \in \ker P, \forall z \in \text{Im } P$ se tiene $T(x_2, z) = T(x_2, Pz) = T(Px_2, z) = 0$ pues P es T-simétrico, entonces $\ker P = (\text{Im } P)^{\perp}$.
Como $\text{Im } P \subset H$ por hipótesis, P es, por definición, un operador de proyección ortogonal. ∎

Observación: Se puede mostrar que
$\forall x \in L \;\; T(x - Px, x - Px) = \min_{y \in \text{Im } P} T(x - y, x - y)$. ∎

Observación: En el caso en que M es un semiproducto escalar no degenerado sobre $\text{Im } {}^tX$ tal que rang M = rang X, M^-M es un operador de proyección ortogonal sobre $\text{Im } {}^tX$ respecto a M y MM^- lo es sobre $\text{Im } M$ respecto a M^-.

Proposición 10 *Como M y D son semiproductos escalares no degenerados respectivamente sobre* $\text{Im } {}^tX$ *y* $\text{Im } X$*, y como* rang M = rang X*, se tiene:*

a) *La restricción de VM a* $\text{Im } {}^tX$*, denotada $VM_{/\text{Im } {}^tX}$, es inyectiva.*

b) $\text{Im } VM = \text{Im } {}^tX$.

c) $\ker VM = \ker M$.

d) *VM es M-simétrico y M-positivo.*

Demostración:

a) Como por hipótesis M es no degenerada sobre $\text{Im } {}^tX$ y V es no degenerada sobre $\text{Im } M$ (c.f. corolario 2), $M_{/\text{Im } {}^tX}$ y $V_{/\text{Im } M}$ son inyectivas. Por lo tanto, $VM_{/\text{Im } {}^tX}$ es inyectiva.

b) Se tiene $\text{Im } VM \subset \text{Im } V = \text{Im } {}^tX$ y rang $(VM_{/\text{Im } {}^tX})$ = rang ${}^tX$. Puesto que rang $VM \geq$ rang $(VM_{/\text{Im } {}^tX})$ se tiene rang $VM \geq$ rang ${}^tX$ y luego $\text{Im } VM = \text{Im } {}^tX$.

c) Se tiene $\ker VM \supset \ker M$. Además rang VM = rang V = rang M, entonces $\dim(\ker VM) = \dim(\ker M)$ y $\ker VM = \ker M$.

d) Como V es un semiproducto escalar se tiene ${}^tV = V$ al identificar E^{**} y E, de donde $\forall (x, y) \in E \times E$ se tiene:

$$M(VMx, y) = \langle VMx, My \rangle = \langle Mx, VMy \rangle = M(x, VMy) = V(Mx, My).$$

Por lo tanto VM es M-simétrico y M-positivo. ∎

Según la igualdad $E = \text{Im } {}^tX \oplus \ker M$ y los incisos (b) y (c) de la proposición anterior, podemos escribir $E = \text{Im } VM \oplus \ker VM$.

Denotamos $\widetilde{\text{Im}^tX}$ el espacio vectorial canónicamente isomorfo a $\text{Im } {}^tX$. Sean $\widetilde{pr_{^tX}}$: $\text{Im } {}^tX \subset E \longrightarrow \widetilde{\text{Im}^tX}$ la proyección canónica de $\text{Im } {}^tX$ sobre $\widetilde{\text{Im}^tX}$ y $\widetilde{in_{^tX}} : \widetilde{\text{Im}^tX} \longrightarrow \text{Im } {}^tX \subset E$ la inyección canónica de $\widetilde{\text{Im}^tX}$ en $\text{Im } {}^tX$.

Definimos la forma bilineal $\widetilde{M}$ sobre $\widetilde{\mathrm{Im}^t X}$ por:

$$\forall (\tilde{x}, \tilde{y}) \in \widetilde{\mathrm{Im}^t X} \times \widetilde{\mathrm{Im}^t X} \quad \widetilde{M}(\tilde{x}, \tilde{y}) = M(x, y)$$

donde $x = in_{^tX}(\tilde{x})$ y $y = in_{^tX}(\tilde{y})$.

Proposición 11 *$\widetilde{M}$ es un producto escalar sobre $\widetilde{\mathrm{Im}^t X}$ y una expresión de la aplicación lineal asociada, también denotada $\widetilde{M}$, es $\widetilde{M} =^t in_{^tX} M in_{^tX}$.*

DEMOSTRACIÓN: La bilinealidad de $\widetilde{M}$ se deduce de la de M y de la linealidad de $in_{^tX}$. La simetría y la positividad se deduicen de las de M. Mostremos que $\widetilde{M}$ es no degenerada sobre el espacio $\widetilde{\mathrm{Im}\ ^t X}$.
Sea $\tilde{x} \in \widetilde{\mathrm{Im}^t X}$ tal que $\forall \tilde{y} \in \widetilde{\mathrm{Im}^t X}$ $M(x, y) = \widetilde{M}(\tilde{x}, \tilde{y}) = 0$, donde $x = in_{^tX}(\tilde{x})$ y $y = in_{^tX}(\tilde{y})$. Como M es no degenerada sobre $\mathrm{Im}\ ^t X$, se tiene $x = 0$ y por consiguiente $\tilde{x} = pr_{^tX}(x) = 0$. Por lo tanto $\widetilde{M}$ es no degenerado.
Mostremos ahora que $\widetilde{M} =^t in_{^tX} M in_{^tX}$ es una expresión de la aplicación lineal asociada a $\widetilde{M}$.
$\forall (\tilde{x}, \tilde{y}) \in \widetilde{\mathrm{Im}^t X} \times \widetilde{\mathrm{Im}^t X}$, $x = in_{^tX}(\tilde{x})$ y $y = in_{^tX}(\tilde{y})$, se tiene:

$$\begin{aligned} \widetilde{M}(\tilde{x}, \tilde{y}) = M(x, y) &= M(in_{^tX}\tilde{x}, in_{^tX}\tilde{y}) \\ &= \langle in_{^tX}\tilde{x}, M in_{^tX}\tilde{y} \rangle \\ &= \langle \tilde{x}, ^t in_{^tX} M in_{^tX}\tilde{y} \rangle \quad \blacksquare \end{aligned}$$

Proposición 12 *El operador $\widetilde{VM}$ sobre $\widetilde{\mathrm{Im}^t X}$ definido por $\widetilde{VM} = pr_{^tX} VM in_{^tX}$ es $\widetilde{M}$-simétrico y $\widetilde{M}$-positivo.*

DEMOSTRACIÓN: $\forall (\tilde{x}, \tilde{y}) \in \widetilde{\mathrm{Im}^t X} \times \widetilde{\mathrm{Im}^t X}$, $x = in_{^tX}(\tilde{x})$ y $y = in_{^tX}(\tilde{y})$, se tiene:

$$\begin{aligned} \widetilde{M}(\widetilde{VM}\tilde{x}, \tilde{y}) &= \widetilde{M}(pr_{^tX} VM in_{^tX}\tilde{x}, \tilde{y}) \\ &= M(in_{^tX} pr_{^tX} VMx, y) \\ &= M(VMx, y). \end{aligned}$$

Como VM es M-simétrico y M-positivo, entonces $\widetilde{VM}$ es $\widetilde{M}$-simétrico y $\widetilde{M}$-positivo. $\blacksquare$

Proposición 13 *El operador VM tiene* $\mathrm{rang}\, X$ *valores propios positivos no nulos y* $\mathrm{Im}\, VM$ *es suma directa de los subespacios propios M-ortogonales asociados a estos valores propios. El subespacio propio associado al valor propio nulo, si existe, es* $\ker VM$.

DEMOSTRACIÓN: Como $\widetilde{M}$ es un producto escalar sobre $\widetilde{\mathrm{Im}\ ^t X}$ y $\widetilde{VM}$ un operador $\widetilde{M}$-simétrico y $\widetilde{M}$-positivo, $\widetilde{VM}$ tiene $\mathrm{rang}\ \widetilde{VM} = \mathrm{rang}\ VM = \mathrm{rang}\ X$ valores propios estrictamente positivos.
Sea $\{\lambda_j / j = 1, \ldots, \mathrm{rang}\ X\}$ el conjunto de valores propios no nulos de $\widetilde{VM}$ y

$\{\tilde{u}_j / j = 1, \ldots, \text{rang } X\}$ el conjunto de vectores propios asociados. Para todo $j \in \{1, \ldots, \text{rang } X\}$ se tiene

$$VM in_{^tX} \tilde{u}_j = in_{^tX} pr_{^tX} VM in_{^tX} \tilde{u}_j = in_{^tX} \widetilde{VM} \tilde{u}_j = \lambda_j in_{^tX} \tilde{u}_j.$$

Si ponemos $u_j = in_{^tX}(\tilde{u}_j)$ se obtiene $VM(u_j) = \lambda_j u_j$. Los u_j sont por lo tanto vectores propios de VM asociados a los valores propios positivos λ_j. Como los $\tilde{u}_j$ asociados a valores propios diferentes son $\widetilde{M}$-ortogonales, los u_j correspondientes son M-ortogonales pues $M(u_j, u_k) = \widetilde{M}(\tilde{u}_j, \tilde{u}_k)$.
Debido a la propiedad 10, los u_j pertenecen a Im VM (= Im ${}^tX$) y por tanto Im VM es suma directa de los subespacios propios M-ortogonales asociados a los valores propios no nulos de VM.
Comme $E = \text{Im } VM \oplus \ker VM$, $\ker VM$ (= $\ker M$) es el subespacio propio asociado al valor propio nulo, si éste existe. ∎

Por la proposición 12, se tiene la descomposición espectral $\widetilde{VM} = \sum_{j=1}^{\text{rang } X} \lambda_j \widetilde{P}_j$ de $\widetilde{VM}$, donde $\widetilde{P}_j$ es el operador de proyección $\widetilde{M}$-ortogonal sobre el subespacio propio asociado al valor propio λ_j.

Proposición 14 *Si* $\forall x \in \text{Im } {}^tX \quad P_j(x) = in_{^tX} \widetilde{P}_j pr_{^tX}(x)$
$\forall x \in (\text{Im } {}^tX)^\perp \quad P_j(x) = 0$
entonces P_j es el operador de proyección M-ortogonal sobre el subespacio propio de VM asociado a λ_j y $VM = \sum_{j=1}^{\text{rang } X} \lambda_j P_j$ es la descomposición espectral de VM respecto a M.

Demostración: Por definición de P_j se tiene Im $P_j \subset$ Im ${}^tX$.
P_j es idempotente pues $\forall x \in E \;\; x = x_1 + x_2$ con $x_1 \in \text{Im } {}^tX$ y $x_2 \in (\text{Im } {}^tX)^\perp$, se tiene

$$\begin{aligned} P_j^2(x) = P_j^2(x_1) &= in_{^tX} \widetilde{P}_j pr_{^tX} in_{^tX} \widetilde{P}_j pr_{^tX}(x_1) \\ &= in_{^tX} \widetilde{P}_j^2 pr_{^tX}(x_1) \\ &= in_{^tX} \widetilde{P}_j pr_{^tX}(x_1) \;=\; P_j(x). \end{aligned}$$

Mostremos que P_j es M-simétrico: $\forall (x, y) \in E \times E$ tal que $x = x_1 + x_2$ y $y = y_1 + y_2$, donde $x_1, y_1 \in \text{Im } {}^tX$ y $x_2, y_2 \in \ker M = (\text{Im } {}^tX)^\perp$, se tiene $M(P_j x, y) = M(P_j x_1, y_1)$.
Además, por definición de $\widetilde{M}$ y en denotando $\tilde{x}_1 = pr_{^tX}(x_1)$ y $\tilde{y}_1 = pr_{^tX}(y_1)$, se tiene:

$$\begin{aligned} M(P_j x_1, y_1) &= \langle M in_{^tX} pr_{^tX} P_j x_1, in_{^tX} pr_{^tX} y_1 \rangle \\ &= \widetilde{M}(pr_{^tX} P_j x_1, pr_{^tX} y_1) \\ &= \widetilde{M}(\widetilde{P}_j \tilde{x}_1, \tilde{y}_1). \end{aligned}$$

Así mismo, $M(x, P_j y) = \widetilde{M}(\tilde{x}_1, \widetilde{P}_j \tilde{y}_1)$ de donde el resultado puesto que $\widetilde{P}_j$ es $\widetilde{M}$-simétrico.
Ahora bien, se tiene

$$VM = in_{^tX} \widetilde{VM} pr_{^tX}$$

$$
\begin{aligned}
&= in_{{}^tX}\left(\sum_{j=1}^{\text{rang }X} \lambda_j \widetilde{P}_j\right) pr_{{}^tX} \\
&= \sum_{j=1}^{\text{rang }X} \lambda_j \left(in_{{}^tX} \widetilde{P}_j pr_{{}^tX}\right. .
\end{aligned}
$$

Se tiene entonces $VM = \sum_{j=1}^{\text{rang }X} \lambda_j P_j$. ∎

Observación: Los P_j son M-positivos: $\forall x \in E$, como los P_j son M-simétricas e idempotentes, se tiene: $M(P_j x, x) = M(P_j^2 x, x) = M(P_j x, P_j x) \geq 0$. ∎

Proposición 15 *Si MVM es la forma bilineal definida por*

$$\forall (x,y) \in E \times E \quad MVM(x,y) = \langle x, MVMy \rangle$$

entonces MVM es un semiproducto escalar no degenerado sobre Im ${}^tX$.

Demostración: MVM es simétrica (resp. positiva) pues M es simétrica (resp. positiva) y VM es M-simétrica (resp. M-positiva). Como M es no degenerada sobre Im ${}^tX$, la restricción de M a Im ${}^tX$ es inyectiva (c.f. proposición 1). De la inyectividad de $VM_{/\text{Im }{}^tX}$ y de Im $(VM) =$ Im ${}^tX$ (c.f. proposición 10), se deduce que la restricción de la aplicación compuesta MVM a Im ${}^tX$ es inyectiva.
La forma bilineal MVM, siendo positiva, por la proposición 1 es no degenerada sobre Im ${}^tX$. ∎

Es fácil ver, como para la proposición 7, que la inversa generalizada algebraica $(VM)^-$ de VM, caracterizada por $\ker(VM)^- = \ker VM (= \ker M)$ y Im $(VM)^- =$ Im VM ($=$ Im ${}^tX$), es tal que: $\forall x \in$ Im ${}^tX \quad (VM)^-(x) = VM^{-1}_{/\text{Im }{}^tX}(x)$.

Observación: Por abuso de languaje, se puede decir que $(VM)^-$ es una inversa generalizada de VM ponderada por el semiproducto escalar M. En efecto, como $(VM)^- VM$ y $VM(VM)^-$ son proyectores, entonces son operadores idempotentes; además, se tiene Im $(VM)^- VM =$ Im $VM(VM)^- =$ Im ${}^tX$ y $\ker(VM)^- VM = \ker VM(VM)^- = \ker M =$ (Im ${}^tX)^\perp$. Luego, $(VM)^- VM$ y $VM(VM)^-$ son operadores de proyección ortogonal sobre Im ${}^tX$ respecto a M. Esta propiedad es característica de las inverses generalizadas ponderadas por un producto escalar, en el sentido de Chipman [6]. ∎

Proposición 16 *El operador $(VM)^-$ es M-simétrico y M-positivo.*

Demostración: Mostremos que $(VM)^-$ es M-simétrico. $\forall (x,y) \in E \times E, x = x_1 + x_2$ y $y = y_1 + y_2$ donde $x_1, y_1 \in$ Im ${}^tX$ y $x_2, y_2 \in \ker M$, como VM es M-simétrico y $VM(VM)^-$ es un operador de proyección sobre Im ${}^tX$ se tiene:

$$
\begin{aligned}
M[(VM)^- x, y] &= M[(VM)^- x_1, y_1] = M[(VM)^- x_1, VM(VM)^- y_1] \\
&= M[VM(VM)^- x_1, (VM)^- y_1] = M[x_1, (VM)^- y_1] \\
&= M[x, (VM)^- y]
\end{aligned}
$$

de donde $(VM)^-$ es M-simétrico.
La M-positividad de $(VM)^-$ es inducida en la tercera igualdad por la M-positividad de VM. ∎

Podemos notar que, si u_j es un vector propio de VM asociado a una valor propio no nulo λ_j, entonces u_j también es vector propio de $(VM)^-$ asociado al valor propio $1/\lambda_j$: en efecto, puesto que $u_j \in \mathrm{Im}\ {}^tX$ se tiene: $u_j = (VM)^- VM(u_j) = \lambda_j (VM)^-(u_j)$. De allí se deduce:

Proposición 17 *El operador* $(VM)^-$ *tiene* $\mathrm{rang}\ X$ *valores propios positivos no nulos, inversos de los de* VM*. Los subespacios propios ortogonales asociados son idénticos a los de* VM *y* $\ker(VM)^-$ $(= \ker M = \ker VM)$ *es el subespacio propio asociado al valor propio nulo, si éste existe.*

3.2 El operador MV

Proposición 18 *Siendo* M *y* D *semiproductos escalares no degenerados, respectivamente sobre* $\mathrm{Im}\ {}^tX$ *y* $\mathrm{Im}\ X$*, y como* $\mathrm{rang}\ M = \mathrm{rang}\ X$*, se tiene:*

a) *La restricción de* MV *a* $\mathrm{Im}\ M$ *es inyectiva.*

b) $\mathrm{Im}\ MV = \{M(x)/x \in \mathrm{Im}\ {}^tX\} = \mathrm{Im}\ M$.

c) $\ker MV = \ker X$.

d) MV *es* M^-*-simétrico y* M^-*-positivo.*

Demostración:

a) Como M y V son no degeneradas sobre $\mathrm{Im}\ {}^tX = \mathrm{Im}\ V$ y $\mathrm{Im}\ M$, respecticamente (c.f. corolario 2 pues $\mathrm{rang}\ V = \mathrm{rang}\ X$ según la proposición 5), entonces $M_{/\mathrm{Im}\ V}$ y $V_{/\mathrm{Im}\ M}$ son inyectivas. Se deduce luego la inyectividad de la restricción de MV a $\mathrm{Im}\ M$.

b) En efecto, se tiene $\mathrm{Im}\ V = \mathrm{Im}\ {}^tX$ y $\mathrm{Im}\ M = \{M(x)/x \in \mathrm{Im}\ {}^tX\}$ (c.f. proposición 4).

c) Se tiene $\ker MV \supset \ker X$, pues $\ker V = \ker X$ (proposición 5). Además se deduce de (b) que $\mathrm{rang}\ MV = \mathrm{rang}\ M$ y por hipótesis $\mathrm{rang}\ M = \mathrm{rang}\ X$, de donde $\dim(\ker MV) = \dim(\ker X)$ y por tanto $\ker MV = \ker X$.

d) $\forall (x^*, y^*) \in E^* \times E^*$, puesto que M^- y V son formas bilineales simétricas, M^-M un operador de proyección sobre $\mathrm{Im}\ M^- = \mathrm{Im}\ V = \mathrm{Im}\ {}^tX$, se tiene:

$$\begin{aligned} M^-(MVx^*, y^*) &= \langle M^-MVx^*, y^*\rangle = \langle Vx^*, y^*\rangle \\ &= V(x^*, y^*) = \langle x^*, Vy^*\rangle \\ &= \langle x^*, M^-MVy^*\rangle = M^-(x^*, MVy^*). \end{aligned}$$

Por lo tanto, MV es M^--simétrica y M^--positiva puesto que V es positiva. ∎

Si u_j es un vector propio de VM asociado a $\lambda_j > 0$, se tiene: $MV[M(u_j)] = \lambda_j M(u_j)$. Puesto que $u_j \in \operatorname{Im} {}^tX$ y que M es inyectiva sobre $\operatorname{Im} {}^tX$, $M(u_j)$ es vector propio de MV asociado a λ_j. Los subespacios propios de MV asociados a los $\lambda_j > 0$ son entonces las imágenes por M de los subespacios propios de VM y $\operatorname{Im} M$ en es la suma directa. Es más, puesto que $E^* = \operatorname{Im} M \oplus \ker X$, $\operatorname{Im} MV = \operatorname{Im} M$, $\ker MV = \ker X$ y $\forall\ x, y \in \operatorname{Im} {}^tX \quad M(x, y) = M^-(Mx, My)$, se deduce:

Proposición 19 *El operador MV tiene* rang X *valores propios positivos no nulos, idénticos a los de VM, y los vectores propios asociados son las imágenes por M de los vectores propios de VM. Los subespacios propios de MV son M^--ortogonales y* $\operatorname{Im} MV$ *en es la suma directa. El subespacio propio asociado al valor propio nulo, si existe, es* $\ker MV$.

4 Aplicación al Análisis de Datos

4.1 Inercia, momento de inercia y producto de inercia

El triplete (X, M, D) caracteriza la nube de puntos individuos $\{x_1, \ldots, x_n\}$, denotada $\mathcal{N}$. Como las variables se suponen centradas, el centro de gravedad de la nube $\mathcal{N}$ es el origen de coordenadas de E. Puesto que la restricción de M a $\operatorname{Im} {}^tX$ es un producto escalar y como $\forall i \quad x_i \in \operatorname{Im} {}^tX$, las expresiones clásicas [1] de la inercia, del momento de inercia y del producto de inercia de $\mathcal{N}$ siguen siendo válidas, pero tomando encuenta algunas precauciones. Por ejemplo, si H es un subespacio de $\operatorname{Im} {}^tX$, no se podrá hablar de la proyección ortogonal sobre $H^\perp$ pues $H^\perp$ no es necesariamente no isotrópico.

Recordemos las definiciones y expresiones de estos índices:

a) **La inercia** de la nube $\mathcal{N}$ respecto a su centro de gravedad es:

$$I[\mathcal{N}] = \sum_{i=1}^{n} p_i \|x_i\|_M^2 = \text{traza}(VM)$$

donde V es la matriz des covarianzas de las variables.

b) Si H es un subespacio vectorial de $\operatorname{Im} {}^tX$ y P_H es el operador de proyección ortogonal sobre H respecto a M, el **momento de inercia de** $\mathcal{N}$ respecto a H es (c.f. ver nota en página 44):

$$I_H[\mathcal{N}] = \sum_{i=1}^{n} p_i \|x_i - P_H(x_i)\|_M^2$$

c) Sean u y v dos vectores de $\operatorname{Im} {}^tX$, Δu y Δv las rectas que generan, si se denota $(\Delta u)^\perp$ y $(\Delta v)^\perp$ los subespacios ortogonales respectivos de Δu y Δv, entonces el **producto de inercia** de $\mathcal{N}$ respecto a $(\Delta u)^\perp$ y $(\Delta v)^\perp$ es:

$$PI[\mathcal{N}/(\Delta u)^\perp \times (\Delta v)^\perp] = \sum_{i=1}^{n} p_i \frac{M(x_i, u)}{\|u\|_M} \frac{M(x_i, v)}{\|v\|_M} = \frac{MVM(u, v)}{\|u\|_M \|v\|_M}.$$

Obsérvese que, como $P_{\Delta u}$ es el operador de proyección ortogonal sobre Δu, se tiene:

$$I[P_{\Delta u}(\mathcal{N})] = PI[\mathcal{N}/(\Delta u)^{\perp} \times (\Delta u)^{\perp}] = \frac{MVM(u,u)}{\|u\|_M^2}.$$

4.2 Análisis en Componentes Principales

El espacio E de individuos está provisto de un semiproducto escalar M, que suponemos no degenerado sobre Im ${}^tX$ y tal que rang M = rang X. El espacio F de las variables está provisto del semiproducto escalar de pesos D, definido por $D(f_j, f_k) = p_j \delta_{jk}$, donde $p_j \geq 0$ y $\sum_{j=1}^n p_j = 1$, y lo suponemos no degenerado sobre Im X. Es más, en lo que sigue supondremos que las variables están centradas.

El hecho de dotar de pesos nulos a algunos individuos, corresponde generalmente a la técnica de definir individuos suplementarios en Análisis de Datos (c.f. por ejemplo [2]).

Proposición 20 a) *Una condición necesaria y suficiente para que D sea no degenerado sobre* Im X *es que* Im $X \cap \ker D = \{0\}$.

b) *Notemos K el subconjunto de $\{1, \ldots, n\}$ tal que $\forall k \in K \;\; p_k = 0$ y $\forall \ell \notin K \;\; p_\ell > 0$, y $\langle f_k / k \in K \rangle$ el subespacio de F generado por $\{f_k / k \in K\}$. Se tiene* Im $X \cap \ker D = \{0\}$ *si y sólo si* Im $X \cap \langle f_k / k \in K \rangle = \{0\}$.

Demostración:

a) Puesto que la forma bilineal simétrica D es positiva, el semiproducto escalar D es no degenerado sobre Im X si y sólo si la restricción de D a Im X es inyectiva (c.f. proposición 1(a) y (c)), lo que es equivalente a decir que Im $X \cap \ker D = \{0\}$.

b) Mostremos que $\ker D = \langle f_k / k \in K \rangle$: $\forall x = \sum_{i=1}^n x_i f_i \in \ker D$ se tiene $\sum_{\ell \notin K} p_\ell (x_\ell)^2 = \sum_{i=1}^n p_i (x_i)^2 = D(x,x) = 0$, ahora bien $\forall \ell \notin K \;\; p_\ell > 0$, luego $\forall \ell \notin K \;\; x_\ell = 0$ lo que implica $x \in \langle f_k / k \in K \rangle$.
Inversamente, $\forall x = \sum_{k \in K} x_k f_k \in \langle f_k / k \in K \rangle$ se tiene $D(x,x) = \sum_{k \in K} p_k (x_k)^2 = 0$ pues $\forall k \in K \;\; p_k = 0$; x es entonces isotrópico y como D es positiva, $x \in \ker D$.
Se tiene por lo tanto Im $X \cap \ker D = \{0\}$ si y sólo si Im $X \cap \langle f_k / k \in K \rangle = \{0\}$. ∎

En presencia de pesos nulos, las condiciones necesarias y suficientes establecidas en la propiedad 20, muestran que la hipótesis de no degeneración de D sobre Im X es generalmente satisfecha. En efecto, si $y \in$ Im $X \cap \langle f_k / k \in K \rangle, y \neq 0$ entonces existen $\alpha_1, \ldots, \alpha_p$ no todos nulos tales que

$$y = \sum_{j=1}^p \alpha_j x^j = \sum_{j=1}^p \alpha_j \Big(\sum_{i=1}^n x_i^j f_i \Big) = \sum_{i=1}^n \Big(\sum_{j=1}^p \alpha_j x_i^j \Big) f_i.$$

Se tiene entonces $\sum_{j=1}^p \alpha_j x_k^j = 0$ para todo $k \notin K$, es decir tal que $p_k \neq 0$. Ahora bien, en general p es pequeño respecto al número d'individuos con un peso no nulo, por lo que el sistema de $n - |K|$ ecuaciones con p incógnitas no tiene, en general, solución.

Puesto que la restricción de M a Im ${}^tX$ es un producto escalar y teniendo en cuenta las propiedades establecidas para VM, los resultados clásicos siguientes [1] sobre los elementos principales del Análisis en Componentes Principales (ACP) del triplete (X, M, D) se enuncian en términos idéntico:

a) Los **vectores axiales principales**, denotados $c_1, \ldots, c_{\text{rang } X}$, forman una base de Im ${}^tX$ constituida por rang X vectores propios M-ortonormados de VM asociados a los valores propios no nulos. Para $\lambda_j > 0$, los vectores $C^j = XMc_j$ son las **componentes principales**. Para $\lambda_j > 0$, se dice que (λ_j, c_j, C^j) es un **triplete principal** del ACP de (X, M, D).

b) Si (λ_j, c_j, C^j) es un triplete principal del ACP de (X, M, D), entonces C^j es vector propio de $W = XM^tX$ asociado al valor propio λ_j.

c) Las componentes principales constituyen una base ortogonal, respecto a D, de Im X; es más, $\|C^j\|_D = MVM(c_j, c_j) = \sqrt{\lambda_j}$.

d) Las C^j tienen media nula y varianza igual a λ_j.

Como VM es un operador M-positivo (c.f. proposición 10(d)), se tiene que $VM = \sum_{j=1}^{\text{rang } X} \lambda_j P_j$ (c.f. proposición 14); denotando $(VM)^{1/2}$ al operador sobre E tal que $(VM)^{1/2} = \sum_{j=1}^{\text{rang } X} \sqrt{\lambda_j} P_j$, se deduce:

- Im $(VM)^{1/2} = \text{Im } [\sum_{j=1}^{\text{rang } X} \sqrt{\lambda_j} P_j] = \text{Im } [\sum_{j=1}^{\text{rang } X} \lambda_j P_j] = \text{Im } VM = \text{Im } {}^tX$ (c.f. proposición 10(b))
- $\ker (VM)^{1/2} = (\text{Im } {}^tX)^\perp = \ker M$ (c.f. §2.1).

Puesto que $E = \text{Im } {}^tX \oplus \ker M$ (c.f. proposición 4) y $\ker(VM)^{1/2} = \ker M$, $(VM)^{1/2}_{/\text{Im } {}^tX}$ es inyectiva. La inversa generalizada algebraica de $(VM)^{1/2}$, denotada $(VM)^{1/2-}$, de núcleo $(\text{Im } {}^tX)^\perp$ y de imagen Im ${}^tX$, es tal que

$$\forall x \in \text{Im } {}^tX \quad (VM)^{1/2-}(x) = (VM)^{-1/2}_{/\text{Im } {}^tX}(x) \tag{1}$$

donde $(VM)^{-1/2}_{/\text{Im } {}^tX}$ es la aplicación recíproca de $(VM)^{1/2}_{/\text{Im } {}^tX}$ (c.f. proposición 7).
De $(VM)^{1/2} = \sum_{j=1}^{\text{rang } X} \sqrt{\lambda_j} P_j$, se deduce $(VM)^{-1/2}_{/\text{Im } {}^tX} = \sum_{j=1}^{\text{rang } X} \frac{1}{\sqrt{\lambda_j}} P_j$.
Puesto que $\ker(VM)^{1/2-} = (\text{Im } {}^tX)^\perp$ y que $\forall x \in (\text{Im } {}^tX)^\perp \;\; P_j(x) = O$ (c.f. proposición 14) se deduce de 1: $(VM)^{1/2-} = \sum_{j=1}^{\text{rang } X} \frac{1}{\sqrt{\lambda_j}} P_j$.

Como los P_j son M-simétricas (proposición 9) y los λ_j positivos, se tiene el siguiente resultado:

Proposición 21 a) *El operador $(VM)^{1/2-}$ es M-simétrico y M-positivo.*

b) *Si (λ_j, c_j, C^j) es un triplete principal del ACP de (X, M, D), entonces c_j es vector propio de $(VM)^{1/2-}$ asociado al valor propio $1/\sqrt{\lambda_j}$.*

Demostración:

a) Mostremos que $(VM)^{1/2}$ es M-simétrico y M-positivo.
Se tiene $\forall(x,y) \in \text{Im }{}^tX \times \text{Im }{}^tX$:
$M[(VM)^{1/2}x,y] = \sum_j \sqrt{\lambda_j} M(P_jx,y) = \sum_j \sqrt{\lambda_j} M(x,P_jy) = M[x,(VM)^{1/2}y]$.
Por lo tanto $(VM)^{1/2}$ es M-simétrico.
Como $\sqrt{\lambda_j} > 0$ y los P_j son M-positivos, se deduce la M positividad de $(VM)^{1/2}$.
El final de la demostración es análogo al de la proposición 16.

b) Se tiene $(VM)^{1/2}c_j = \sum_k \sqrt{\lambda_k} P_k c_j = \sqrt{\lambda_j} c_j$,
es más, $c_j \in \text{Im }{}^tX$ y $(VM)^{1/2-}(VM)^{1/2}$ es un proyector sobre $\text{Im }{}^tX$, luego $c_j = (VM)^{1/2-}(VM)^{1/2}(c_j) = \sqrt{\lambda_j}(VM)^{1/2-}(c_j)$ lo que implica $(VM)^{1/2-}c_j = \frac{1}{\sqrt{\lambda_j}}(c_j)$.
■

En la proposición que sigue se establecen ciertas propiedades de la aplicación $XM(VM)^{1/2-}$, de E en F, que serán utilizadas en un próximo artículo para dar expresiones algebraicas a los semiproductos escalares relacionales.

Proposición 22 a) *Se tiene* $XM(VM)^{1/2-}(c_j) = \frac{C^j}{\|C^j\|_D}$, *para todo triplete principal* (λ_j, c_j, C^j) *del ACP de* (X,M,D).

b) $\text{Im }[XM(VM)^{1/2-}] = \text{Im } X$.

c) $\ker[XM(VM)^{1/2-}] = (\text{Im }{}^tX)^\perp$.

Demostración:

a) Se tiene $(VM)^{1/2-}(c_j) = \frac{1}{\sqrt{\lambda_j}} c_j$ (c.f. proposición 21(b)). Ahora bien, $\sqrt{\lambda_j} = \|C^j\|_D$, de donde $(VM)^{1/2-}c_j = \frac{c_j}{\|C^j\|_D}$. Por lo tanto, se obtiene $XM(VM)^{1/2-}(c_j) = XM\left(\frac{c_j}{\|C^j\|_D}\right) = \frac{C^j}{\|C^j\|_D}$.

b) Se deduce inmediatamente de (a).

c) Se tiene $\ker XM(VM)^{1/2-} \supset \ker(VM)^{1/2-} = \ker M$. Es más, $XM(VM)^{1/2-}(c_j) = \frac{C^j}{\|C^j\|_D}$, luego $\dim[\text{Im } XM(VM)^{1/2-}] = \dim(\text{Im } X) = \text{rang } X$ (c.f. (c) de la page 51). Ahora bien, $E = \text{Im }{}^tX \oplus \ker M$, entonces $\dim(\ker M) = \dim E - \text{rang } X = \dim[\ker XM(VM)^{1/2-}]$. Por lo tanto $\ker XM(VM)^{1/2-} = \ker M = (\text{Im }{}^tX)^\perp$. ■

Obsérvese que $XM(VM)^{1/2-}$ es una isometría de $\text{Im }{}^tX$ en $\text{Im } X$.

5 Conclusiones y perspectivas

Se han establecido propiedades útiles de los semiproductos escalares, que han servido para trabajar en una definición de las semimétricas útiles en Análisis de Datos. Tanto en el espacio de individuos como en el espacio de variables, el uso de estas semimétricas ha permitido trabajar de manera que se puede desarrollar una teoría coherente.

En una próxima publicación, extenderemos estas ideas al caso de semimétricas relacionales, es decir, definidas por bloques, las cuales serán útiles para el caso de trabajar con tablas múltiples o bien para extender una serie de técnicas de Análisis de Datos.

Referencias

[1] Cailliez, F.; Pagès, J.P. (1976) *Introduction à l'Analyse des Données*. SMASH, Paris

[2] Escofier, B.; Pagès, J. (1988) *Analyses Factorielles Simples et Multiples*. Dunod, Paris.

[3] Glazman, I.; Liubitch, Y. (1972) *Analyse Linéaire dans les Espaces de Dimensions Finies*. Editions MIR, Moscou.

[4] Grifone, J. (1990) *Algèbre Linéaire*. Cépaduès-Editions, Toulouse.

[5] Halmos, P.R. (1958) *Finite Dimensional Vector-Spaces*. Van Nostrand, Princeton, New Jersey.

[6] Nashed, Z. (ed.) (1976) *Generalized Inverses and its Applications*. Academic Press Inc., New York.

[7] Pontier, J.; Dufour, A.-B; Normand, M. (1991) *Le Modèle Euclidien en Analyse des Données*. S.M.A., Editions Ellipses, Bruxelles.

[8] Queysanne, A. (1964) *Algèbre*. Armand Colin, Ed., Paris.

[9] Schektman, Y. (1978) *Contribution à la Mesure en Facteurs dans les Sciences Expérimentales y à la mise en Œuvre Automatique dans les Calculs Statistiques*. Thèse d'Etat, Toulouse.

[10] Schektman, Y. (1988) *Analyse et Traitement Informatique de Données*. Cours de D.E.A., Université Paul Sabatier, Toulouse.

[11] Schektman, Y. (1991) "Eléments mathématiques de base pour la définition de semi-produits scalaires utiles en analyse de données. Quelques applications", Note interne, Laboratoire MLAD, Université Paul Sabatier, Toulouse.

[12] Schektman, Y.; Troupé, M.; Trejos, J. (1992) "Un générateur de règles floues pàrtir de bases de données volumineuses", In: E. Diday & Y. Kodratoff (Eds.) *III Journées Symboliques–Numériques*, Université Paris IX–Dauphine, pp. 121–130.

[13] Troupé, M. (1994) *Contribution à la Protection de la Régression Multiple MMultidimensionnelle et l̀a Génération de Règles Prévisionelles.* Thèse de Doctorat, Université Paul Sabatier, Toulouse.

[14] Trejos, J. (1994) *Contribution à l'Acquisition Automatique de Connaissances à Partir de Données Qualitatives.* Thèse de Doctorat, Université Paul Sabatier, Toulouse.