

# GENERACIÓN DE MÉTRICAS EN ANÁLISIS DE DATOS

FABRICIO BOLAÑOS\*      Javier Trejos\*\*

*Recibido: 20 Mar 2002*

---

## Resumen

En el presente trabajo se presentan los generadores primarios de métricas, que consisten en combinaciones lineales de métricas conocidas. Se establecen algunas propiedades teóricas de estos generadores, referentes al cálculo de componentes principales, correlaciones principales e inercia. Se aplican los generadores primarios al Análisis en Componentes Principales usando diversas combinaciones de métricas clásicas.

**Palabras clave:** Análisis en Componentes Principales, métricas Euclídeas, correlación.

## Abstract

In this work we introduce primary generators of metrics, that consist of linear combinations of known metrics. We establish some theoretical properties of these generators, related to the computation of principal components, principal correlations and inertia (generalized variance). We apply primary generators in Principal Component Analysis using some combinations of classic metrics.

**Keywords:** Principal Component Analysis, Euclidean metrics, correlation.

**Mathematics Subject Classification:** 62H25

## 1. Introducción

En este trabajo se aborda el problema de la escogencia de la métrica en el (A.C.P.); en este método se analiza una tabla de datos donde son medidas sobre  $p$  variables cuantitativas  $x^1, \dots, x^p$ ,  $n$  individuos. Supondremos que las variables están centradas y que los individuos tienen pesos  $p_i > 0$  tales que  $\sum_{i=1}^p p_i = 1$ . Esta tabla de datos forma una matriz  $X_{n \times p}$ . Se define el espacio de individuos  $E = \mathbb{R}^p$  y se dota de una métrica  $M$ , análogamente se define el espacio de variables  $F = \mathbb{R}^n$ , que tiene la métrica usual de

---

\*Sede Regional del Pacífico, Universidad de Costa Rica, Puntarenas, Costa Rica. E-Mail: fbolanos@srp.ucr.ac.cr

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

pesos  $D = \text{diag}(p_i)$ . Como es bien sabido [2] los resultados de este análisis dependen en gran medida de la escogencia de la métrica y algunos autores han estudiado diferentes propiedades de las métricas en Análisis de Datos, como por ejemplo M. Maurin [6, 7] y E. Piza [8] entre otros.

En 1983 Y. Schektman propuso los generadores de métricas que se describen en la siguiente sección; estos generadores fueron usados por Labrèche, y algunos fueron implementados por Bolaños para la obtención de algunos de los resultados que aquí se presentan.

## 2. Generadores primarios de métricas.

Los generadores primarios de métricas se pueden definir por:  $M_r = w_1(r)\bar{M}_1 + w_2(r)\bar{M}_2$ , donde  $w_1(r) = \frac{m-r}{m} \in [0, 1]$ ,  $w_2(r) = 1 - w_1(r) \in [0, 1]$  y  $r \in \{0, \dots, m\}$ ,  $m \in \mathbb{N}$ .

Usualmente  $\bar{M}_1$  y  $\bar{M}_2$  son escogidos en el conjunto  $\{V, V^{-1}, I, D_{1/\sigma^2}\}$ . Existen otros tipos de combinación [1] que aquí no presentamos. Citamos a continuación algunas propiedades que hemos obtenido en [1]

**Proposición 1** *Si  $\bar{M}_1, \bar{M}_2 \in \{V, V^{-1}, I_p\}$ , entonces los vectores propios de  $M_r$  son idénticos a los de  $V$  y los valores propios correspondientes se deducen de los de  $V$ .*

**Proposición 2** *Si  $\bar{M}_1 = D_{1/\sigma^2}$ ,  $\bar{M}_2 = V^{-1}$ , entonces los vectores propios de  $VM_r$  son iguales a los de  $VD_{1/\sigma^2}$ .*

**Proposición 3** *Si  $\lambda_1, \dots, \lambda_p$  son los valores propios de  $V$ , asociados a los vectores propios  $u_1, \dots, u_p$ , entonces tenemos:*

1. *Si  $\bar{M}_1 = V$ ,  $\bar{M}_2 = I$ , entonces los vectores propios de  $VM_r$  son iguales a los de  $V$  y los valores propios son:  $(\frac{m-r}{m})\lambda_i^2 + \frac{r}{m}\lambda_i$ .*
2. *Si  $\bar{M}_1 = I$ ,  $\bar{M}_2 = V^{-1}$  y  $r < m$ , entonces los vectores propios de  $VM_r$  son iguales a los de  $V$  y los valores propios son:  $(\frac{m-r}{m})\lambda_i + \frac{r}{m}$ .*

DEMOSTRACIÓN:

Veamos la demostración del primer apartado.

$$\begin{aligned} VM_r u_i &= V\left[\frac{m-r}{m}V + \frac{r}{m}I\right]u_i = V\left[\frac{m-r}{m}\lambda_i u_i + \frac{r}{m}u_i\right] \\ &= \frac{m-r}{m}\lambda_i^2 u_i + \frac{r}{m}\lambda_i u_i = \left[\frac{m-r}{m}\lambda_i^2 + \frac{r}{m}\lambda_i\right]u_i. \end{aligned}$$

La demostración del segundo apartado es análoga. ■

### Proposición 4

1. *Si  $\bar{M}_1 = V$ ,  $\bar{M}_2 = I$  entonces las componentes principales se pueden calcular por:  $C_\alpha^k = \sqrt{[(1-\alpha)\lambda_k + \alpha]} C_1^k$ .*
2. *Si  $\bar{M}_1 = I$ ,  $\bar{M}_2 = V^{-1}$  y  $r < m$ , entonces las componentes principales restantes se pueden calcular por:  $C_\alpha^k = \sqrt{[(1-\alpha) + \frac{\alpha}{\lambda_k}]} C_1^k$ .*

**Proposición 5**

1. Si  $\bar{M}_1 = V$  y  $\bar{M}_2 = I$  entonces las correlaciones entre las variables y las componentes principales de los respectivos A.C.P. son las mismas para todo  $r$ .
2. Si  $\bar{M}_1 = I$ ,  $\bar{M}_2 = V^{-1}$  y si  $r < m$ , entonces las correlaciones entre las variables y las componentes principales de los respectivos A.C.P. son las mismas que las anteriores.

DEMOSTRACIÓN:

Sean  $\lambda_1, \dots, \lambda_p$  los valores propios de  $V$  asociados a los vectores propios  $u_1, \dots, u_p$ .

Para la demostración de la primera parte, en el caso sea  $C_1^k = XIu_k$ . Veamos que:

$$C_\alpha^k = X[(1 - \alpha)V + \alpha I]u_k = [(1 - \alpha)\lambda_k + \alpha]Xu_k = [(1 - \alpha)\lambda_k + \alpha]C_1^k.$$

Con base en lo anterior escribimos:

$$\rho(x^j, C_\alpha^k) = \frac{\text{cov}(x^j, C_\alpha^k)}{\|x^j\|_D \|C_\alpha^k\|_D} = \frac{[(1 - \alpha)\lambda_k + \alpha]\text{cov}(x^j, C_1^k)}{\|x^j\|_D [(1 - \alpha)\lambda_k + \alpha] \|C_1^k\|_D} = \rho(x^j, C_1^k).$$

La demostración de la segunda parte es análoga. ■

La proposición anterior es muy importante para el análisis, ya que de ella se deduce que cuando se efectúa una generación primaria de  $V$  a  $I$  o de  $I$  a  $V^{-1}$  (sin llegar a esta última), las variables no se pueden interpretar de otra forma, es decir la elección de la métrica (primaria) es independiente de los resultados, para las variables.

**Proposición 6** Si  $\alpha = \frac{r}{m}$  entonces la inercia de la nube  $\mathcal{N}_r = (X, M_r, D)$  es:

1. Si  $\bar{M}_1 = V$  y  $\bar{M}_2 = I_p$  entonces  $I(\mathcal{N}_r) = (1 - \alpha) \sum_{k=1}^p \sum_{j=1}^p \text{cov}^2(x^j, x^k) + \alpha \sum_{j=1}^p \text{var}(x^j)$ .
2. Si  $\bar{M}_1 = I_p$  y  $\bar{M}_2 = V^{-1}$  entonces  $I(\mathcal{N}_r) = (1 - \alpha) \sum_{j=1}^p \text{var}(x^j) + \alpha p$ .
3. Si  $\bar{M}_1 = I_p$  y  $\bar{M}_2 = D_{1/\sigma^2}$  entonces  $I(\mathcal{N}_r) = (1 - \alpha) \sum_{j=1}^p \text{var}(x^j) + \alpha p$ .

La demostración anterior se pueden consultar en [1]

También se ha propuesto otro tipo de generación que se llama generación espectral [9], para la que Labrèche ha establecido propiedades que ligán a ésta con algunas propiedades de las métricas que estableció M. Maurin [5]. Nosotros trabajamos con un tipo de generador espectral llamado proporcional y obtuvimos algunas propiedades interesantes que se presentaron en [13]

### 3. Resultados numéricos

Por limitaciones de espacio, nos restringiremos a presentar solamente los resultados obtenidos sobre una tabla de datos, si el lector desea más información puede consultar [1]. La tabla estudiada es la de de notas escolares donde para 10 estudiantes se obtienen las notas en 5 materias, ver tabla 1.

Puede observarse que para algunas escogencias de  $\bar{M}_1$  y  $\bar{M}_2$  se presenta un único plano principal y un círculo de correlaciones gracias a las proposiciones 4 y 5.

Estudiante	Matemáticas	Ciencias	Español	Historia	Ed. Física
Lucía	7.0	6.5	9.2	8.6	8.0
Pedro	7.5	9.4	7.3	7.0	7.0
Carmen	7.6	9.2	8.0	8.0	7.5
Luis	5.0	6.5	6.5	7.0	9.0
Andrés	6.0	6.0	7.8	8.9	7.3
Ana	7.8	9.6	7.7	8.0	6.5
Carlos	6.3	6.4	8.2	9.0	7.2
José	7.9	9.7	7.5	8.0	6.0
Sonia	6.0	6.0	6.5	5.5	8.7
María	6.8	7.2	8.7	9.0	7.0

Cuadro 1: Tabla de datos de las notas escolares

### 3.1. Generación $V \rightarrow I_5$

En la tabla 2 se muestra un resumen conteniendo los valores propios y los porcentajes de inercia explicada para 5 de las 13 métricas generadas.

Así mismo, la tabla 3 contiene las componentes principales correspondientes y la tabla 4 las correlaciones entre las componentes principales y las variables originales.

$i =$	$\alpha = 0/12(M = V)$		$\alpha = 3/12$		$\alpha = 6/12$		$\alpha = 9/12$		$\alpha = 12/12(M = I)$	
	$\lambda_i$	% inercia	$\lambda_i$	% inercia	$\lambda_i$	% inercia	$\lambda_i$	% inercia	$\lambda_i$	% inercia
1	12,2382	78,6983	10,0532	76,7806	7,8682	73,9769	5,6833	69,4884	3,4983	61,1442
2	3,2163	20,6828	2,8606	21,8477	2,5049	23,5507	2,1491	26,2772	1,7934	31,3458
3	0,0733	0,4716	0,1227	0,9372	0,1721	1,6179	0,2214	2,7076	0,2708	4,7334
4	0,0228	0,1469	0,0549	0,4195	0,0870	0,8180	0,1191	1,4559	0,1512	2,6418
5	0,0001	0,0004	0,0020	0,0151	0,0039	0,0365	0,0058	0,0709	0,0077	0,1348

Cuadro 2: Los valores propios y porcentajes de inercia explicada, generador de  $V$  a  $I$ .

Ind	$\alpha = 0/12$		$\alpha = 3/12$		$\alpha = 6/12$		$\alpha = 9/12$		$\alpha = 12/12$	
	$C^1$	$C^2$	$C^1$	$C^2$	$C^1$	$C^2$	$C^1$	$C^2$	$C^1$	$C^2$
$x_1$	-1,4303	2,1183	-1,2964	1,9977	-1,1469	1,8694	-0,9747	1,7315	-0,7647	1,5818
$x_2$	3,1214	-1,8641	2,8291	-1,7580	2,5028	-1,6451	2,1271	-1,5238	1,6689	-1,3920
$x_3$	2,9519	-0,4011	2,6754	-0,3783	2,3669	-0,3540	2,0116	-0,3279	1,5782	-0,2995
$x_4$	-4,8761	-1,7680	-4,4194	-1,6674	-3,9098	-1,5602	-3,3229	-1,4452	-2,6070	-1,3202
$x_5$	-2,6911	1,7887	-2,4390	1,6869	-2,1578	1,5785	-1,8338	1,4621	-1,4388	1,3357
$x_6$	4,3915	-0,5197	3,9802	-0,4901	3,5212	-0,4586	2,9926	-0,4248	2,3479	-0,3881
$x_7$	-1,6716	2,0341	-1,5151	1,9183	-1,3403	1,7951	-1,1391	1,6627	-0,8937	1,5189
$x_8$	4,9562	-0,5698	4,4920	-0,5373	3,9740	-0,5028	3,3775	-0,4658	2,6498	-0,4255
$x_9$	-4,9183	-2,9240	-4,4577	-2,7575	-3,9436	-2,5804	-3,3516	-2,3901	-2,6296	-2,1834
$x_{10}$	0,1664	2,1056	0,1508	1,9857	0,1334	1,8582	0,1134	1,7212	0,0890	1,5723

Cuadro 3: Las componentes principales, generador de  $V$  a  $I$ .

Variable	$\alpha = 0/12$	
	$r(C^1, x^j)$	$r(C^2, x^j)$
$x^1$	0,9438	0,0448
$x^2$	0,9468	-0,2952
$x^3$	0,2561	0,8776
$x^4$	0,2489	0,9414
$x^5$	-0,8556	-0,3049

Cuadro 4: Correlaciones entre variables y componentes, generador de  $V$  a  $I$ .

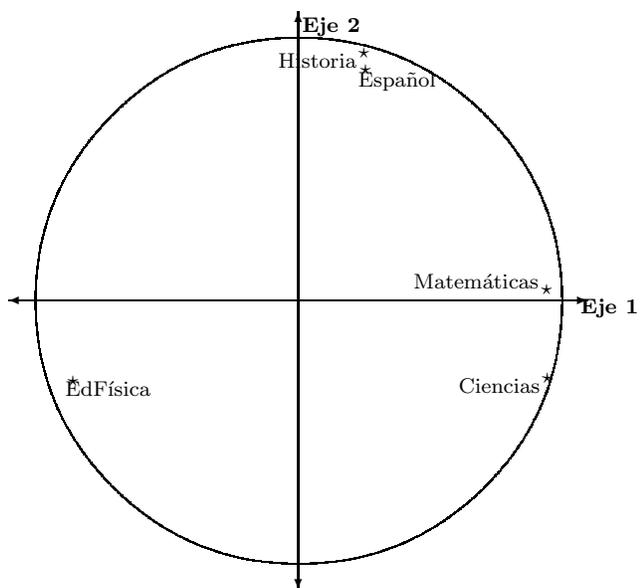
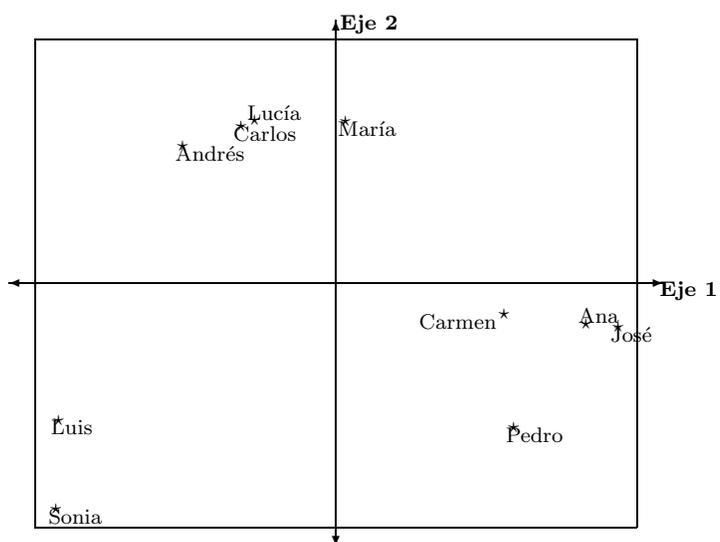


Figura 1: Plano principal y círculo de correlaciones cuando  $\alpha = 0$ , de  $V$  a  $I$ .

### 3.2. Generación $I_5 \rightarrow V^{-1}$

En la tabla 5 se muestra un resumen conteniendo los valores propios y los porcentajes de inercia explicada para 5 de las 13 métricas generadas.

Así mismo, la tabla 6 contiene las componentes principales correspondientes y la tabla 7 las correlaciones entre las componentes principales y las variables originales.

$i =$	$\alpha = 3/12$		$\alpha = 6/12$		$\alpha = 9/12$		$\alpha = 11/12$		$\alpha = 12/12(M = V^{-1})$	
	$\lambda_i$	% inercia	$\lambda_i$	% inercia	$\lambda_i$	% inercia	$\lambda_i$	% inercia	$\lambda_i$	% inercia
1	2,8737	51,8626	2,2492	41,9564	1,6246	31,3604	1,2082	23,8768	1,0000	20,0001
2	1,5951	28,7863	1,3967	26,0546	1,1984	23,1327	1,0661	21,0690	1,0000	20,0000
3	0,4531	8,1774	0,6354	11,8531	0,8177	15,7847	0,9392	18,5615	1,0000	20,0000
4	0,3634	6,5577	0,5756	10,7369	0,7878	15,2072	0,9293	18,3645	1,0000	20,0000
5	0,2558	4,6161	0,5039	9,3990	0,7519	14,5150	0,9173	18,1282	1,0000	20,0000

Cuadro 5: Los valores propios y porcentajes de inercia explicada, de  $I$  a  $V^{-1}$ .

Ind	$\alpha = 3/12$		$\alpha = 6/12$		$\alpha = 9/12$		$\alpha = 11/12$		$\alpha = 12/12(M = V^{-1})$	
	$C^1$	$C^2$	$C^1$	$C^2$	$C^1$	$C^2$	$C^1$	$C^2$	$C^1$	$C^2$
$x_1$	-0,6931	1,4917	-0,6132	1,3959	-0,5211	1,2930	-0,4494	1,2196	1,0797	0,6323
$x_2$	1,5126	-1,3127	1,3381	-1,2284	1,1373	-1,1378	0,9807	-1,0732	-0,8918	0,7435
$x_3$	1,4304	-0,2824	1,2655	-0,2643	1,0755	-0,2448	0,9275	-0,2309	0,4227	-2,0030
$x_4$	-2,3629	-1,2451	-2,0904	-1,1651	-1,7766	-1,0792	-1,5321	-1,0179	-2,1524	-0,6222
$x_5$	-1,3040	1,2596	-1,1536	1,1787	-0,9805	1,0918	-0,8455	1,0298	0,9040	-0,6443
$x_6$	2,1280	-0,3660	1,8826	-0,3425	1,6000	-0,3172	1,3798	-0,2992	0,2863	-0,5645
$x_7$	-0,8100	1,4324	-0,7166	1,3404	-0,6090	1,2416	-0,5252	1,1711	0,7659	0,2943
$x_8$	2,4017	-0,4012	2,1247	-0,3755	1,8058	-0,3478	1,5573	-0,3281	0,4120	-0,2169
$x_9$	-2,3833	-2,0591	-2,1085	-1,9268	-1,7920	-1,7848	-1,5454	-1,6834	-1,1975	0,5023
$x_{10}$	0,0806	1,4828	0,0713	1,3875	0,0606	1,2852	0,0523	1,2122	0,3712	1,8784

Cuadro 6: Las componentes principales, generador de  $I$  a  $V^{-1}$ .

Variable	$\alpha = i/12 \ i = 0, \dots, 11$		$\alpha = 12/12$	
	$r(C^1, x^j)$	$r(C^2, x^j)$	$r(C^1, x^j)$	$r(C^2, x^j)$
$x^1$	0,9438	0,0448	0,4869	-0,0745
$x^2$	0,9468	-0,2952	0,0714	-0,2841
$x^3$	0,2561	0,8776	0,8434	0,2693
$x^4$	0,2489	0,9414	0,8117	0,0559
$x^5$	-0,8556	-0,3049	-0,5882	-0,0448

Cuadro 7: Correlaciones entre variables y componentes principales, de  $I$  a  $V^{-1}$ .

### 3.3. Generación $I_5 \rightarrow D_{1/\sigma^2}$

En la tabla 8 se muestra un resumen conteniendo los valores propios y los porcentajes de inercia explicada para 5 de las 13 métricas generadas. Así mismo, la tabla 9 contiene las componentes principales correspondientes y la tabla 10 las correlaciones entre las componentes principales y las variables originales.

### 3.4. Comentarios

Puede notarse que en la generación de  $V$  a  $I_5$  los resultados son prácticamente los mismos para todas las métricas de la familia generada, tal como lo establecen las proposiciones 5 y 4. Esta estabilidad se sigue observando cuando se pasa de  $I_5$  hacia  $V^{-1}$ , justo antes de llegar a esta última.

En general se observa lo siguiente:

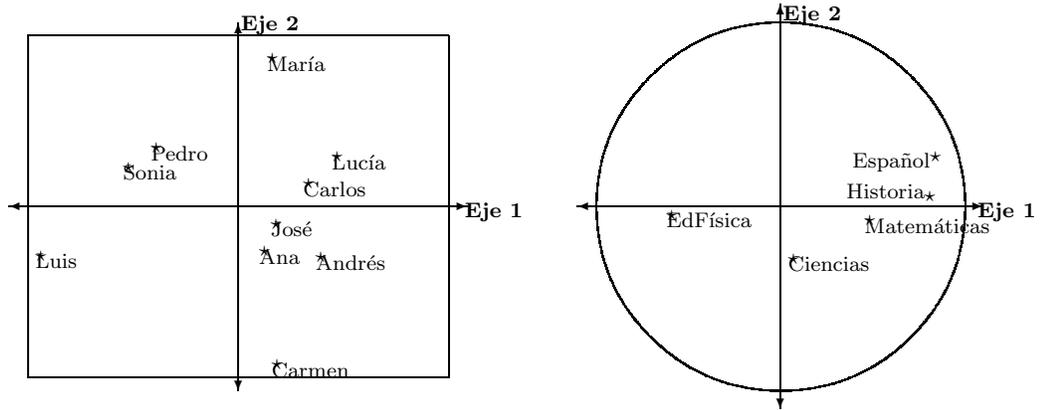


Figura 2: Plano principal y círculo de correlaciones con  $\alpha = 12$ , de  $I$  a  $V^{-1}$ .

$i =$	$\alpha = 3/12$		$\alpha = 6/12$		$\alpha = 9/12$		$\alpha = 12/12$	
	$\lambda_i$	% inercia	$\lambda_i$	% inercia	$\lambda_i$	% inercia	$\lambda_i$	% inercia
1	3,2965	59,4921	3,1227	58,2517	2,9862	57,6447	2,8932	57,8650
2	1,7955	32,4035	1,7737	33,0873	1,7197	33,1968	1,6287	32,5730
3	0,2912	5,2552	0,3107	5,7956	0,3292	6,3548	0,3466	6,9319
4	0,1497	2,7023	0,1451	2,7069	0,1365	2,6344	0,1226	2,4523
5	0,0081	0,1469	0,0085	0,1586	0,0088	0,1693	0,0089	0,1778

Cuadro 8: Los valores propios y porcentajes de inercia explicada de  $I$  a  $D_{1/\sigma^2}$ .

- En el círculo de correlaciones, se nota que la primera componente principal hace una oposición entre la variable “Educación Física” y las materias científicas “Matemáticas” y “Ciencias”, mientras que la segunda componente principal sigue la tendencia de las humanísticas “Historia” y “Español”.
- En el plano principal se aprecian tres grupos de estudiantes: Luis y Sonia (los fuertes en Ed. Física); Pedro, Carmen, Ana y José (los buenos en materias científicas) y Andrés, Carlos, Lucía y María (los buenos en materias humanísticas).

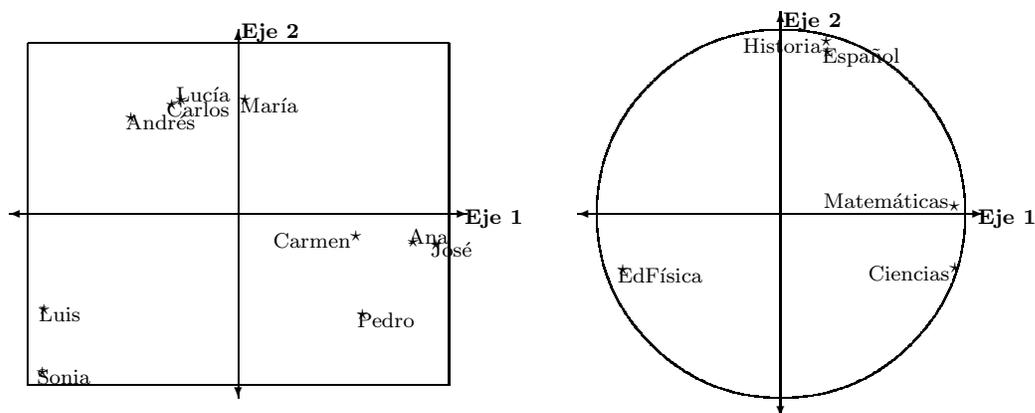
Cuando  $M = V^{-1}$  se observan algunos cambios importantes:

- En el círculo de correlaciones, se nota que deja de tener importancia la oposición “Educación Física vs Materias científicas”. Recuérdese que en este caso cualquier

Individuo	$\alpha = 3/12$		$\alpha = 6/12$		$\alpha = 9/12$		$\alpha = 12/12$	
	$C^1$	$C^2$	$C^1$	$C^2$	$C^1$	$C^2$	$C^1$	$C^2$
$x_1$	-0,5521	1,6748	-0,2926	1,7486	0,0081	1,7855	0,3231	1,7725
$x_2$	1,4668	-1,4743	1,2261	-1,5567	0,9512	-1,6196	0,6654	-1,6387
$x_3$	1,4534	-0,3651	1,3135	-0,4344	1,1600	-0,4924	1,0025	-0,5157
$x_4$	-2,7469	-1,1704	-2,9020	-0,9394	-3,0522	-0,6265	-3,1721	-0,2628
$x_5$	-1,2377	1,3726	-1,0049	1,4028	-0,7474	1,4076	-0,4889	1,3654
$x_6$	2,2264	-0,5301	2,0801	-0,6971	1,9047	-0,8718	1,7086	-1,0217
$x_7$	-0,6932	1,5449	-0,4577	1,5550	-0,1955	1,5323	0,0676	1,4623
$x_8$	2,5385	-0,6061	2,3976	-0,8226	2,2197	-1,0579	2,0119	-1,2759
$x_9$	-2,7148	-2,0197	-2,8249	-1,8019	-2,9430	-1,5354	-3,0420	-1,2549
$x_{10}$	0,2596	1,5733	0,4647	1,5457	0,6945	1,4782	0,9239	1,3694

Cuadro 9: Las componentes principales, generador de  $I$  a  $D_{1/\sigma^2}$ .

Var	$\alpha = 3/12$		$\alpha = 6/12$		$\alpha = 9/12$		$\alpha = 12/12$	
	$r(C^1, x^j)$	$r(C^2, x^j)$	$r(C^1, x^j)$	$r(C^2, x^j)$	$r(C^1, x^j)$	$r(C^2, x^j)$	$r(C^1, x^j)$	$r(C^2, x^j)$
$x^1$	0,948	-0,027	0,945	-0,121	0,928	-0,231	0,895	-0,345
$x^2$	0,917	-0,370	0,871	-0,459	0,805	-0,557	0,722	-0,648
$x^3$	0,327	0,861	0,415	0,830	0,514	0,781	0,610	0,717
$x^4$	0,320	0,915	0,407	0,876	0,505	0,819	0,599	0,748
$x^5$	-0,881	-0,228	-0,904	-0,128	-0,917	-0,008	-0,913	0,119

Cuadro 10: Las correlaciones entre las variables y las componentes de  $I$  a  $D_{1/\sigma^2}$ .Figura 3: Plano principal y círculo de correlaciones  $\alpha = 0$ , de  $I$  a  $D_{1/\sigma^2}$ .

vector de  $\mathbb{R}^p$  es vector propio de  $VM$  ya que  $VM = I_5$ , y que la componente principal asociada tendrá varianza 1 y una inercia explicada de  $1/5$ .

- En el plano principal, se observa una “mezcla” de los individuos que anteriormente formaban grupos bien diferenciados.

De los resultados de la generación de  $I_5$  a  $D_{1/\sigma^2}$ , se observa que:

- Para la primera componente principal, entre Educación Física y las materias académicas (ya no son sólo las científicas), y para la segunda componente principal la oposición es entre las materias científicas y las humanísticas.
- En el plano principal se mantienen los mismos tres grupos de individuos anteriormente interpretados.

Podemos concluir que los resultados mejor interpretables, a la luz de la tabla de datos, se obtienen para cuando  $M = D_{1/\sigma^2}$ , esto es, cuando se estandarizan las variables y se usa la métrica usual. Además se ve que la interpretación es difícil cuando  $M = V^{-1}$ .

## Referencias

- [1] Bolaños, F (1996) *Contribución al estudio de las métricas en Análisis de Datos*. Tesis de Licenciatura.
- [2] Cailliez, F.; Pagès, J.P. (1976) *Introduction à l'Analyse de Données*. SMASH, París.

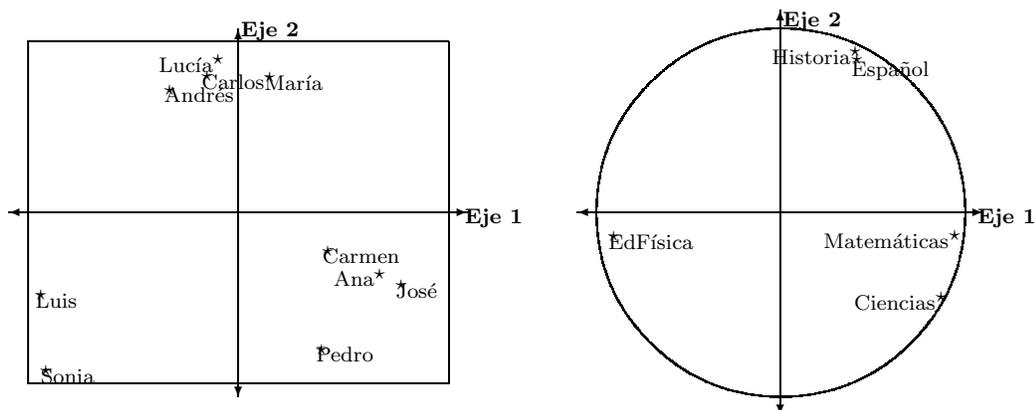


Figura 4: Plano principal y círculo de correlaciones  $\alpha = 6$ , de  $I$  a  $D_{1/\sigma^2}$ .

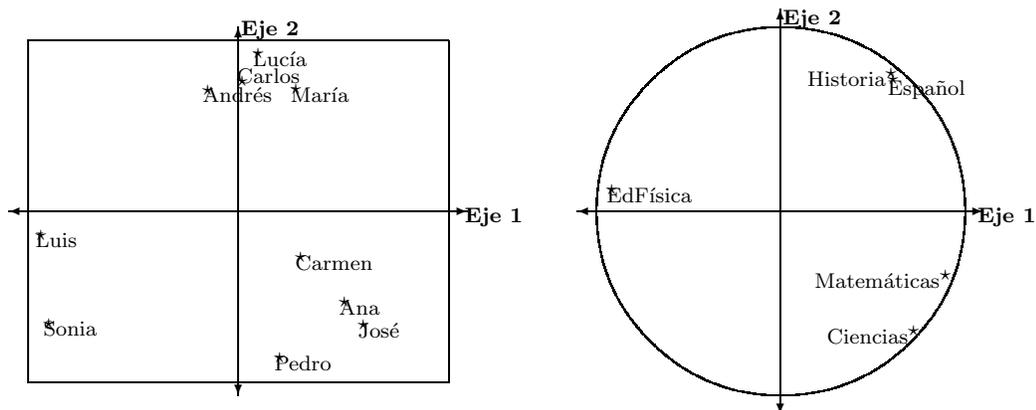


Figura 5: Plano principal y círculo de correlaciones  $\alpha = 12$ , de  $I$  a  $D_{1/\sigma^2}$ .

- [3] Fabre, C. (1986) *Contribution à la Protection des Méthodes Relationnelles*. Tesis de doctorado de tercer ciclo. Universidad Paul Sabatier, Toulouse.
- [4] Kazmierczak, J.B. (1987) "Sur l'usage d'un principe d'invariance pour aider au choix d'une métrique." *Statistique et Analyse des Données*, vol.12(3): 37-57.
- [5] Labrèche, S. (1995) "Selección de métricas en A.C.P. y en clasificación." *IX Simposio Métodos Matemáticos Aplicados a las Ciencias*, J. Trejos (ed.), U.C.R.-I.T.C.R., Turrialba, pp 87-96.
- [6] Maurin, M. (1987) "Sur les qualités des métriques en A.C.P." En: *Statistique et Analyse des Données*, vol.12(3): 58-74.
- [7] Maurin, M. (1991) "Un paramétrage des métriques en ACP, applications et conséquences." *XXIII Journées de Statistique*, ASU, Estrasburgo, pp. 358-360.
- [8] Piza, E. (1989) "La métrica inducida por un Análisis en Componentes Principales." *Ciencia y Tecnología*, Vol XIII(2): 167-179.

- [9] Schektman, Y. (1983) *Quelques réflexions sur les métriques. Une approche originale pour la protection des méthodes relationnelles et l'étude de la régression multidimensionnelle*. Nota interna, Laboratorio MLAD, Universidad Paul Sabatier, Toulouse.
- [10] Schektman, Y. (1984) “Métriques et analyse des données, contributions et synthèses,” *II Colloque Franco-Soviétique de Statistique*, CNET, Paris.
- [11] Trejos, J. (1993) *Contribution à l'Acquisition Automatique de Connaissances à Partir de Données Qualitatives*. Tesis de doctorado, Universidad Paul Sabatier, Toulouse.
- [12] Trejos, J. (1994) *Génération de produits scalaires pour l'Analyse des Données*. Nota interna, PIMAD, Escuela de Matemática, Universidad de Costa Rica. 31pp.
- [13] Trejos, J.; Bolaños, F. (1995) “Generación de métricas útiles en análisis de datos,” *IX Simposio Métodos Matemáticos Aplicados a las Ciencias*, J. Trejos (Ed.), U.C.R.-I.T.C.R., Turrialba, pp 173–182.