

**BÚSQUEDA DE UNA ESTRUCTURA COMÚN EN EL ESTUDIO DE  
LA DISTRIBUCIÓN DE GRASAS DE 2 GRUPOS DE MUJERES  
CUBANAS\***

*Lydia Lera Marqués<sup>1</sup>, Amparo Vallejo Arboleda<sup>2</sup> Maria Elena Diaz Sánchez<sup>3</sup>*

<sup>1</sup>Instituto de Cibernética, Matemática y Física Teórica, CITMA

Calle E # 309 entre 13 y 15, Vedado, La Habana, Cuba

[lydia@cidet.icmf.inf.cu](mailto:lydia@cidet.icmf.inf.cu)

<sup>2</sup>Universidad de Antioquia, Medellín

[avallejo@matematicas.udea.edu.co](mailto:avallejo@matematicas.udea.edu.co)

<sup>3</sup>Instituto de Nutrición e Higiene de los Alimentos, MINSAP

# BÚSQUEDA DE UNA ESTRUCTURA COMÚN EN EL ESTUDIO DE LA DISTRIBUCIÓN DE GRASAS DE 2 GRUPOS DE MUJERES CUBANAS\*

*Lydia Lera Marqués<sup>1</sup>, Amparo Vallejo Arboleda<sup>2</sup> Maria Elena Díaz Sánchez<sup>3</sup>*

<sup>1</sup>Instituto de Cibernética, Matemática y Física Teórica, CITMA  
Calle E # 309 entre 13 y 15, Vedado, La Habana, Cuba  
[lydia@cidet.icmf.inf.cu](mailto:lydia@cidet.icmf.inf.cu)

<sup>2</sup>Universidad de Antioquia, Medellín  
[avallejo@matematicas.udea.edu.co](mailto:avallejo@matematicas.udea.edu.co)

<sup>3</sup>Instituto de Nutrición e Higiene de los Alimentos, MINSAP

**Resumen:** En el trabajo se analizan las modificaciones de la composición corporal y distribución de grasas de 2 grupos de mujeres cubanas pre y post menopáusicas, a través de la búsqueda de una estructura de las variables antropométricas analizadas, común a los grupos. Para ello se utilizan los métodos factoriales: Análisis del Espacio Común de Krzanowski, Análisis de Componentes Principales Comunes, método Procrustes y método STATIS (Structuration des Tableaux A Trois Indices de la Statistique).

## INTRODUCCION

Se ha demostrado que el exceso de tejido adiposo y la distribución de la grasa en el cuerpo son factores de riesgo establecidos para las enfermedades crónicas y muertes, lo que hace necesario el estudio de métodos de análisis de la composición corporal en estudios epidemiológicos. La mayoría de estas técnicas no son accesibles a gran escala (por el costo de los equipos), por lo que la antropometría continúa siendo el método simple, preciso y económicamente posible para el trabajo de terreno de grupos numerosos de población. El estudio de la composición corporal por simples mediciones somáticas sigue siendo una temática de actualidad, que se debe acometer de forma multidimensional ya que el ser humano es la resultante de numerosos factores incidentes en la naturaleza (Díaz, 1999). Esto hace que los métodos multivariados, en particular, las técnicas factoriales sean adecuadas para estos fines.

En los últimos años han surgido una serie de técnicas de tipo factorial que permiten el análisis simultáneo de 2 o más grupos de datos, cuyo objetivo principal es la búsqueda de una estructura común entre diferentes conjuntos de datos. Entre estos métodos tenemos: el Análisis del Espacio Común de Krzanowski (Krzanowski, 1979), el Análisis de Componentes Principales Comunes (Flury, 1984), el STATIS (Lavit, 1988), y el método de Procrustes Generalizado (Gower, 1975; Lacourly et al., 1998).

---

\* Presentado en el IX Congreso Internacional de Biomatemática. Concepción. Chile

En el trabajo se aplican las técnicas factoriales anteriores a una serie de mediciones antropométricas de 2 grupos de mujeres pre y post menopáusicas para analizar las modificaciones de la composición corporal de estas mujeres.

## MATERIALES Y MÉTODOS

Se analizaron los datos formados por 2 grupos de 150 mujeres entre 35 y 60 años provenientes de un estudio sobre la etapa reproductiva y climaterio femenino.

La muestra fue seleccionada de los consultorios del Plan del Médico de la Familia de un área de salud del municipio Habana Vieja en la Ciudad de la Habana, Cuba, entre 1991 y 1992. La muestra fue clasificada en mujeres pre menopáusicas y mujeres post menopáusicas según si habían alcanzado el período climatérico (Díaz, 1999).

Las variables antropométricas medidas fueron: diámetros biacromial (BIAC) y bicrestal (BICR); circunferencias (cm.) cintura (CCIN), abdomen (CABD), muslo (superior) (CMUS) y cadera (CCAD); pliegues (mm.) tríceps (PTRI), pierna (media) (PPIE), medioaxilar (PMAX), subescapular (PSE) y suprailíaco (PSID).

## DESCRIPCIÓN DE LOS MÉTODOS

Los métodos de tipo factorial pueden utilizarse en cualquiera de las situaciones siguientes:

- Los mismos individuos medidos en diferentes ocasiones para las mismas variables.
- Los mismos individuos medidos en diferentes ocasiones para diferentes variables.
- Diferentes conjuntos de individuos a los que se les miden las mismas variables.

Nuestro problema se refiere a la situación 3, dos conjuntos de mujeres a las que se les miden las mismas variables.

## ANÁLISIS DEL ESPACIO COMÚN

Este método tiene como objetivo fundamental la búsqueda de estructuras comunes entre  $k$  grupos de observaciones a los que se les miden las mismas variables y se basa en la aplicación del Análisis de Componentes Principales (ACP) a cada uno de los grupos y su descripción en términos de un número menor de componentes principales (CP). A continuación se comparan los componentes principales de los  $k$  grupos diferentes calculando los ángulos ( $\theta_i$ ) entre los subespacios generados por las  $k$  primeras componentes principales de cada grupo. Como medida de similitud se utiliza  $\sum_{i=1}^k \cos^2 \theta_i$ . Se ve que  $0 \leq \sum_{i=1}^k \cos^2 \theta_i \leq k$  donde  $k$  corresponde a espacios coincidentes y 0 a espacios ortogonales (Sánchez, 1995).

Este método tiene la desventaja que no brinda la posibilidad de gráficos y no da una estimación de la matriz de varianzas y covarianzas.

## ANÁLISIS DE COMPONENTES PRINCIPALES COMUNES

El Análisis de Componentes Principales Comunes (ACPC) también busca estructuras comunes entre diferentes grupos de datos.

Flury (1984) encuentra una transformación ortogonal simple del sistema coordinado para 2 ó más grupos simultáneamente, es decir busca una rotación que diagonalice a las matrices de varianzas y covarianzas de los grupos (Flury y Riedwyl, 1988).

Cómo los componentes principales comunes (CPC) no son exactamente incorrelacionados, la cuestión es si el modelo de CPC es apropiado para los datos. Para ello se construye una prueba estadística y se obtienen estimadores máximo verosímil.

Se definen las cantidades:

$A_i$  = producto de las varianzas de las CPC del grupo  $i$

$B_i$  = producto de las varianzas de las CP originales del grupo  $i$

El estadístico para un número de  $k$  grupos y  $p$  variables será

$$X_{\text{CPC}}^2 = \sum_{i=1}^k (n_i - 1) \log \frac{A_i}{B_i}$$

Si el modelo de CPC y los tamaños muestrales  $n_i$  son suficientemente grandes, el estadístico se distribuye aproximadamente como chi-cuadrado con  $(k-1)p$  g.l.

Este modelo tiene las siguientes ventajas:

1. En lugar de  $k$  conjuntos de transformaciones de componentes principales sólo tiene que ser interpretada una matriz ortogonal.
2. Los coeficientes de la transformación CPC son más estables (en el sentido de menor error estándar) que los CP ordinarios.

Tiene la desventaja de no brindar la posibilidad de gráficos.

## STATIS

El método STATIS es un método exploratorio del Análisis de Datos, que se aplica a datos cuantitativos, siendo su principal objetivo la búsqueda de una estructura común entre  $k$  grupos que pueden estar formados por los mismos individuos a los que se les han hecho mediciones a través del tiempo o por individuos diferentes a los que se les miden las mismas variables. Esta estructura (intraestructura) está descrita por las distancias mutuas entre las observaciones ( $W_k = X_k \cdot X_k^t$ ,  $X_k$ : matriz de datos en la ocasión  $k$ ). Contrariamente a este término se define la interestructura como las relaciones entre grupos de observaciones, descritas por las distancias entre  $W_k$  y a partir de ellas se construye una imagen euclídiana plana de los grupos, si  $G_1, \dots, G_k$  es la nube de puntos formada por los grupos, el coseno del ángulo entre los vectores  $OG_k$  y  $OG_l$  es la aproximación del producto escalar normado entre  $W_k$  y  $W_l$  y se le llama coeficiente RV que representa un índice de similitud. Si el coeficiente RV se aproxima a 1 se dice que las observaciones tienen la misma estructura en el interior de las matrices  $X_k$  y  $X_l$ . El método es como una generalización del Análisis de Componentes Principales y brinda la posibilidad de gráficos.

## MÉTODO DE PROCRUSTES GENERALIZADO (MPG)

El MPG busca un compromiso C, las rotaciones y los factores de escala minimizando:

$$\sum_{k=1}^T \|X_k - \tilde{X}_k\|^2 \text{ con } \tilde{X}_k = \rho_k C P_k$$

$$\min_{C, \rho_k, P_k} \sum_{k=1}^T \|X_k - \rho_k C P_k\|^2 = \min_{C, \rho_k, P_k} \sum_{k=1}^T \text{tr}(X_k - \rho_k C P_k)(X_k - \rho_k C P_k)'$$

$$= \min_{C, \rho_k, P_k} \sum_{k=1}^T \text{tr}(\rho_k C P_k - X_k)(\rho_k C P_k - X_k)'$$

donde  $X_k$  : k-ésima tabla de dimensión nxm y  $\| \cdot \|$  es la norma de Frobenius, (n=no. de observaciones y m=no. de variables);

$\rho_k$  : k-ésimo coeficiente de escala;

$P_k$  : k-ésima matriz de rotación de dimensión mxm tal que  $P_k P_k' = I$ ;

C : matriz compromiso de dimensión mxn.

Este método brinda la posibilidad de gráficos.

En general, estos métodos factoriales siguen un enfoque descriptivo centrándose en la estructura algebraica de los datos y no tienen en cuenta supuestos distribucionales.

## RESULTADOS Y DISCUSIÓN

La tabla 1 muestra las medias y desviaciones estándar de las variables medidas.

**Tabla 1: Medias y desviaciones estándar de las variables antropométricas analizadas**

### GRUPO 1

	BIAC	BICR	CCIN	CABD	CCAD	CMUS	PTRI	PSE	PMAX	PSID	PPIE
<b>MED</b>	35.95	26.51	81.47	96.52	99.91	61.54	24.12	28.07	21.31	27.16	23.294
<b>D.E.</b>	2.24	2.92	11.58	12.47	11.53	7.7375	10.63	12.21	7.97	11.37	8.73

### GRUPO 2

	BIAC	BICR	CCIN	CABD	CCAD	CMUS	PTRI	PSE	PMAX	PSID	PPIE
<b>MED</b>	35.65	26.44	85.68	100.18	101.48	61.72	25.76	30.15	22.79	28.79	23.74
<b>D.E.</b>	2.13	3.11	12.29	12.46	12.58	7.73	8.85	12.21	7.97	11.37	8.73

El grupo de mujeres post menopáusicas posee valores ligeramente más altos en sus dimensiones físicas pero sólo son significativos cintura, abdomen, cadera y el pliegue cutáneo de pierna.

El grupo de mujeres pre menopáusicas tiene mayor desarrollo magro dado por los diámetros biacronial y bicrestal.

Para cada uno de los grupos se hallaron los componentes principales (tablas 2 y 3) observándose que en los dos grupos el primer componente explica la composición corporal fundamentalmente la grasa cutánea, interna y periférica. Los componentes 2 y 3 indican la relación entre las variables, como se diferencian. En el primer componente principal las variables de los diámetros biacronial y bicrestal, que representan la masa magra, son las que menos peso tienen. En general, este es un componente de la masa corporal total.

**Tabla 2: Vectores y valores propios del grupo de mujeres pre menopáusicas**

<b>COMPONENTES PRINCIPALES (CP)</b>			
	<b>CP 1</b>	<b>CP 2</b>	<b>CP 3</b>
<b>BIAC</b>	.470279	-.796533	.037707
<b>BICR</b>	.536081	.230054	-.719835
<b>CCIN</b>	.929667	-.074113	.026561
<b>CABD</b>	.935020	-.109082	-.079582
<b>CCAD</b>	.918228	-.130718	-.117850
<b>CMUS</b>	.875797	-.182865	-.139875
<b>PTRI</b>	.883525	.108852	.097796
<b>PSE</b>	.899990	.126606	.103099
<b>PMAX</b>	.743432	.228023	.444020
<b>PSID</b>	.874035	.175145	.194991
<b>PPIE</b>	.687965	.246555	-.045250
<b>Expl.Var</b>	7.237774	.926646	.817493
<b>Prop.Total</b>	.657979	.084241	.074318

Los dos conjuntos de transformaciones obtenidos por el ACP, parecen ser muy similares por lo que es razonable asumir que puedan ser idénticas. En estos casos es recomendable buscar una transformación común de los CP para los 2 grupos y estimar sólo una matriz ortogonal para los datos. Es en estos casos donde radica la ventaja de aplicar las técnicas factoriales descritas anteriormente.

En otros casos los coeficientes de la transformación presentan alguna variabilidad muestral y puede suceder que las transformaciones obtenidas en los grupos sean diferentes sólo por el error muestral y sin embargo, los vectores propios presenten relaciones idénticas en cada uno de los grupos (Flury y Riedwyl, 1988).

**Tabla 3: Vectores y valores propios del grupo de mujeres post menopáusicas**

<b>COMPONENTES PRINCIPALES (CP)</b>			
	<b>CP 1</b>	<b>CP 2</b>	<b>CP 3</b>
<b>BIAC</b>	.461195	-.651543	-.480785
<b>BICR</b>	.506757	-.515397	.418587
<b>CCIN</b>	.909703	-.069374	-.130617
<b>CABD</b>	.930008	-.118522	-.006101
<b>CCAD</b>	.871933	-.152409	.051179
<b>CMUS</b>	.818816	-.194681	.044785
<b>PTRI</b>	.850249	.136497	.122644
<b>PSE</b>	.823908	.210514	-.061460
<b>PMAX</b>	.640794	.440428	-.377529
<b>PSID</b>	.795621	.393297	-.069785
<b>PPIE</b>	.653332	.218365	.454003
<b>Expl.Var</b>	6.464924	1.229422	.800428
<b>Prop.Total</b>	.587720	.111766	.072766

En general, el ACP arroja que el tipo físico de las mujeres pre y post menopáusicas es de forma androide, con alta incidencia de obesidad. Además, no se diferencian mucho entre ellas debido a que en la muestra de pre menopáusicas hay algunas mujeres que ya presentan síndrome climatérico y otras son pre menopáusicas tardías lo que hace que ya su figura comience a transformarse en los años peri menopáusicos.

En la tabla 4 se tienen los componentes principales comunes para dichos grupos. Se ve que el primer componente representa también la composición corporal, de forma similar a los resultados obtenidos por los componentes principales ordinarios. Luego interpretando sólo una matriz ortogonal se logra caracterizar la composición corporal y la distribución de grasas de las mujeres.

Este modelo asume que las direcciones de los p ejes principales (en este caso 11), son idénticas para ambos grupos por lo que se tiene que buscar un compromiso con las transformaciones individuales de cada grupo.

Los valores propios simples y los valores propios comunes son relativamente cercanos y lo que los une esencialmente es la grasa subcutánea, interna y periférica, dada por las mediciones de las circunferencias y pliegues.

El valor de la prueba estadística es

$X^2_{CPC} = 35.27$ , luego se cumple la hipótesis de existencia de ejes principales comunes para los 2 grupos (la chi-cuadrado para 55 g.l. y 95 % es aproximadamente 39).

Podemos concluir que el modelo de componentes principales comunes es apropiado.

**Tabla 4: Análisis de Componentes Principales Comunes de los grupos de mujeres pre y post menopáusicas**

**COMPONENTES PRINCIPALES COMUNES**

	CP 1	CP 2
<b>BIAC</b>	0.4657	-0.7240
<b>BICR</b>	0.5215	-0.3729
<b>CCIN</b>	0.9197	-0.0717
<b>CABD</b>	0.9325	-0.1138
<b>CCAD</b>	0.8951	-0.1416
<b>CMUS</b>	0.8470	-0.1888
<b>PTRI</b>	0.8669	0.1227
<b>PSE</b>	0.8619	0.1686
<b>PMAX</b>	0.6921	0.3342
<b>PSID</b>	0.8348	0.2842
<b>PPIE</b>	0.6707	0.3120
<b>% VAR</b>	58.52	75.45

En la tabla 5 se tienen los valores y vectores propios obtenidos por el método de Krzanowski para las variables analizadas. Se ve que con 2 componentes principales se logra una buena aproximación del espacio común, dado por las mediciones de circunferencia y pliegues. Al analizar el primer componente principal se obtienen los mismos resultados que se obtuvieron con los métodos anteriores.

Puede observarse que la  $\sum_{i=1}^k \cos^2 \theta_i$  se aproxima al número de tablas (2 en este caso) lo que sugiere que efectivamente hay características comunes entre los 2 grupos.

La figura 1 muestra una representación euclidiana de los grupos. Los productos escalares normados de los grupos son cercanos a 1, lo que dice que las matrices de correlación son homotéticas. El % de varianza explicado por el primer eje es elevado (50.51%).

La interestructura, es decir, la relación entre grupos, puede ser considerada como una matriz de correlación promedio entre las variables (tabla 6). Representa la matriz de correlación compromiso y es un buen reflejo de la correlación de los 2 grupos.

**Tabla 5: AEC de las medidas antropométricas de los grupos de mujeres pre y post menopáusicas**

**VECTORES Y VALORES PROPIOS**

<b>BIAC</b>	0.1739	-0.7707
<b>BICR</b>	0.1945	-0.1244
<b>CCIN</b>	0.3512	-0.0789
<b>CABD</b>	0.3558	-0.1228
<b>CCAD</b>	0.3411	-0.1514
<b>CMUS</b>	0.3224	-0.2014
<b>PTRI</b>	0.3324	0.1255
<b>PSE</b>	0.3309	0.1719
<b>PMAX</b>	0.2676	0.3425
<b>PSID</b>	0.3220	0.2890
<b>PPIE</b>	0.2578	0.2441
<b>VALORES PROPIOS</b>	1.9995	1.6922

	<b>CP1</b>		<b>CP2</b>	
	$\cos^2 \theta$	$\theta$	$\cos^2 \theta$	$\theta$
<b>GRUPO 1</b>	0.9998	0.0220	0.8461	0.5621
<b>GRUPO 2</b>	0.9998	0.0220	0.8461	0.5621
$\sum_{i=1}^k \cos^2 \theta_i$	1.9996		1.6922	

**Tabla 6: Matriz de correlación**

	1.0000	0.8546
	0.8546	1.0000

De forma general se tiene que las variables están significativamente correlacionadas.

Al aplicar el método de Procrustes también se obtuvo una fuerte relación entre los 2 grupos analizados ( $R^2 = 0.9993$ ).

Los resultados obtenidos confirman que la menopausia está asociada con un aumento de la masa grasa y una disminución de la masa magra (Aloia et al., 1995).

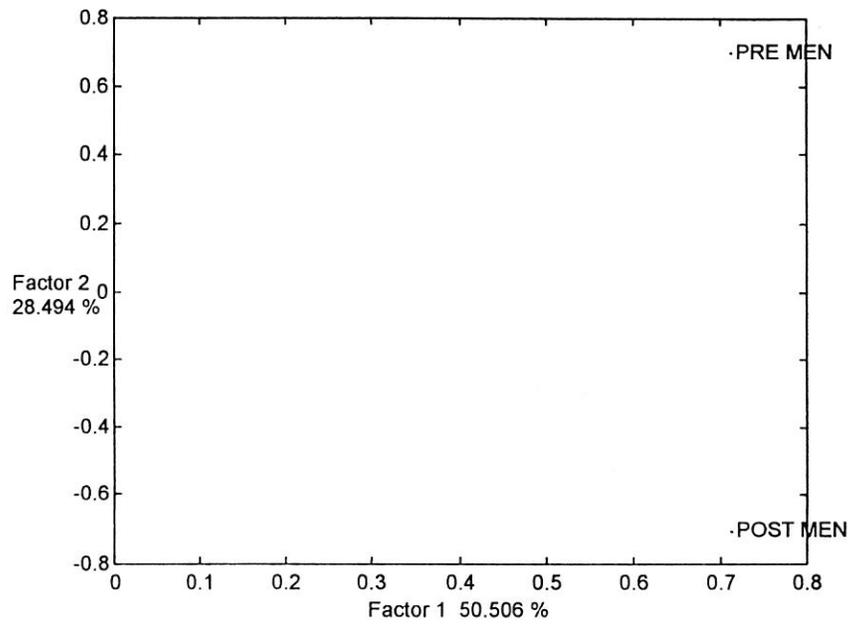


Figura 1: Imagen euclidiana de los grupos de mujeres pre y post menopáusicas

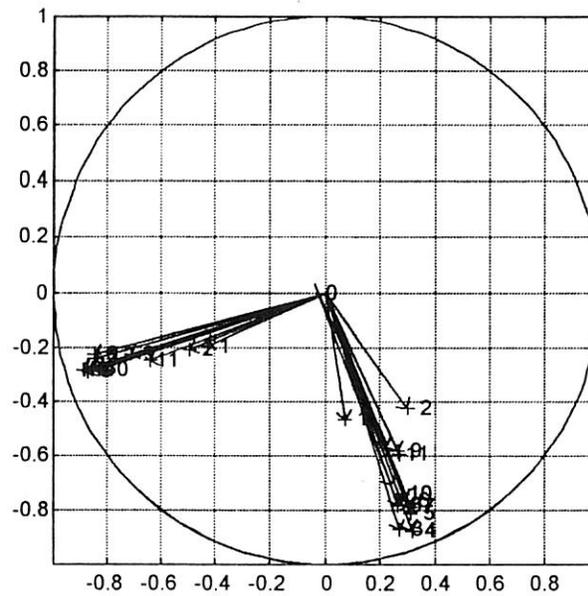


Figura 2: Círculo de correlaciones

## CONCLUSIONES

1. La utilización combinada de los métodos factoriales resulta muy útil para interpretar la variabilidad presente en los 2 grupos de mujeres desde puntos de vista diferentes, complementándose entre ellos y corroborando los resultados.
2. Se encontró una estructura común para las variables antropométricas dada fundamentalmente por la grasa cutánea, interna y periférica.
3. Se caracterizaron los grupos a través de una estructura común a diferencia de ACP que caracteriza a cada grupo por separado.

**Agradecimientos:** A la Universidad de Antioquia, Medellín, Colombia por financiar una parte de esta investigación.

## REFERENCIAS

- Aloia, J. F., Vaswani, A., Russo, L., Sheehan, M., Flaster, E. (1995): The influence of menopause and hormonal replacement therapy on body cell mass and body fat mass. *Amer. J. Obst. Ginecol.*, 172.
- Carlier, A., Lavit, C., Pages, M., Pernin, M.O., Turlot, J.C. (1989): A comparative review of methods which handle a set of methods which handle a set of indexed data tables. In: *Multivariate Data Analysis*, eds: Coppy, R. y Bolasco, S. Elsevier Science Publishers B.V.
- Díaz Sánchez, M. E. (1999): *Composición Corporal y Distribución de Grasa en grupos de la población cubana*. Tesis de doctorado. Instituto de Nutrición e higiene de los alimentos.
- Ferrales, M. (1992): *Biología y Aspectos Sociales de la Mujer en la Etapa Reproductiva y el Climaterio*. Trabajo de Diploma. Universidad de la Habana. Cuba
- Flury, B. (1984): *Common Principal Components in k groups*. *JASA*, 79.
- Flury, B. y Riedwyl, H. (1988): *Multivariate Statistics. A practical approach*. Chapman and Hall, New York.
- Gower, J.C. (1975): *Generalized Procrustes Analysis*. *Psychometrika* 40, 33-51.
- Krzanowski, W.J. (1979): *Between group comparison of principal components*. *JASA* 74, pp. 703-707.
- \_\_\_\_\_ (1981): *Correction of 1979, JASA* 78, p.1022.
- \_\_\_\_\_ (1988): *Principles of Multivariate Analysis. A user's Perspective*, Clarendon Press, Oxford.
- Lacourly, N., Lera, L. y Farías, F. (1999): *An Algorithm for Generalized Procrustes Analysis*. Enviado a publicar.
- Lavit, Christine (1988): *Analyse Conjointe de Tableaux Quantitatifs*. Masson, Paris.
- Mardia, K. V., Kent, J. T. y Bibby, M. 1979: *Multivariate Analysis*, Academic Press, Londres.
- Sánchez, Silvia (1995): *El Análisis de Componentes Principales Comunes*. Tesis de Maestría. Universidad de Guadalajara, México.