ESTADÍSTICA ESPAÑOLA
Vol. 39, Núm. 142, 1997, págs. 129 a 140

Nuevos algoritmos no jerárquicos en clasificación de datos

por
LUIS JAVIER LÓPEZ MARTÍN(1)
Departamento de Economía de las Instituciones y Estadística-Econometría
Universidad de La Laguna

MONTSERRAT HERNÁNDEZ LÓPEZ
Departamento de Economía de las Instituciones y Estadística-Econometría
Universidad de La Laguna

RESUMEN

En la literatura sobre clasificación, los métodos no jerárquicos han recibido menor atención que los métodos jerárquicos. Esta circunstancia es aún más acentuada en lo que se refiere a las técnicas de recubrimiento. En este artículo, presentamos dos nuevos algoritmos no jerárquicos que, además, generan intersecciones no vacías entre las clases resultantes.

Palabras clave: análisis cluster, métodos no jerárquicos, técnicas de optimización, técnicas de recubrimiento, algoritmos.

Clasificación AMS: 62H30

(1) Los autores de este artículo desean expresar su sincero agradecimiento al Doctor Miguel Sánchez García y a los evaluadores anónimos por sus útiles comentarios y sugerencias. Lógicamente, todos los errores que pudieran existir son de nuestra responsabilidad.

1. INTRODUCCIÓN

Entre las técnicas más usadas en el análisis de datos, fundamentalmente si se trabaja con datos cualitativos, se encuentran las técnicas de clasificación (Sokal y Sneath, 1963). Estas técnicas tienen como objetivo agrupar objetos en clases o clusters, internamente lo más homogéneos posible, de tal forma que los objetos pertenecientes a un cluster estén más próximos entre sí que los pertenecientes a grupos diferentes.

Para llevar a la práctica esta metodología se hace imprescindible, en muchos casos, disponer de una medida D de desemejanza entre los objetos.

Definición 1. Una medida de desemejanza D es una aplicación

$$D: OxO \to R^+ \cup \{0\} \text{ tal que: } \begin{cases} 1)\ D(o_i, o_i) = 0 & \forall o_i \in O \\ 2)\ D(o_i, o_j) = D(o_j, o_i) & \forall o_i, o_j \in O \end{cases}$$

siendo O el conjunto de n objetos a agrupar.

Si además D verifica la denominada *desigualdad métrica o triangular*

$$D(o_i, o_j) \leq D(o_i, o_k) + D(o_k o_j), \qquad \forall o_i, o_j, o_k \in O$$

se dice que D es una distancia.

Muchos son los métodos de clasificación propuestos a lo largo de la historia de esta metodología. En muchos casos, se trata de métodos jerárquicos, y en otros, menos numerosos, los métodos son no jerárquicos o de enlace único. En los primeros, el método genera, a distintos niveles de desemejanza ε, distintas soluciones de agrupamiento o particiones entre los objetos; por el contrario, para los segundos la solución se obtiene una vez fijado un nivel de desemejanza determinado (Bailey, 1994).

Los métodos jerárquicos actúan sobre la matriz de desemejanzas (o de semejanzas) para construir un árbol que describirá, mediante uniones o divisiones sucesivas, las relaciones entre los objetos. Los algoritmos que han seguido la vía de la jerarquización se distinguen entre sí en función del criterio adoptado para la unión o división de los grupos. Johnson (1967) realiza una profunda revisión de este tipo de métodos de clasificación(2).

La principal ventaja de los métodos jerárquicos es la facilidad de interpretación del árbol resultante. Entre sus desventajas se suelen citar las siguientes: el algorit-

(2) Algunos de estos métodos han sido desarrollados por Sokal y Sneath (1963) —Enlace Simple—, Sokal y Sneath (1963) —Enlace Completo—, Sokal y Michener (1958) —Enlace Medio en el Grupo—, McQuitty (1965, 1967) —Enlace Medio Ponderado—, Sokal y Michener (1958) y Gower (1967) —Centroide— y Ward (1963) —Método de Ward—.

mo de clasificación *sólo pasa una vez por los datos*, por lo que una partición inicial desacertada no puede ser subsanada; en el caso del método del Enlace Simple, aparece el efecto del *encadenamiento*; pueden generar soluciones diferentes simplemente reordenando los datos en la matriz de desemejanzas o semejanzas, lo que causa problemas de *no estabilidad* de las soluciones (Aldenderfer y Blashfield, 1984).

Dentro de los métodos no jerárquicos se encuentran, entre otros, los métodos de partición iterativa o de optimización y las técnicas de recubrimiento.

La característica esencial que distingue a los métodos de optimización es que producen una única partición de los objetos en un número particular k, especificado a priori, de clusters no solapados, como resultado de la minimización o maximización de alguna función objetivo (Jardine y Sibson, 1971; Anderberg, 1973).

Generalmente, estos métodos empiezan con una partición inicial del conjunto de objetos en k clusters, para cada uno de los cuales se define un centroide; se localiza, entonces, cada objeto en el cluster que tenga su centroide más cerca, para, posteriormente, calcular los nuevos centroides y volver a relocalizar cada objeto. Así sucesivamente hasta que no se produzcan cambios en los clusters (MacQueen, 1966).

Por su parte, los métodos de recubrimiento conducen a clusters no disjuntos, que permiten que un objeto cualquiera pueda aparecer en más de una clase final a la vez. Su objetivo básico es recubrir el conjunto de objetos O con una familia de clusters maximales (Everitt, 1993). Estos métodos están muy condicionados por los objetivos del estudio, las características intrínsecas de las variables observadas para los objetos, así como por la medida de desemejanza elegida para relacionar los objetos.

La idea de formar clusters no disjuntos puede no parecer atractiva, ya que la interpretación de los mismos suele ser complicada. Quizás sea ésta la razón por la que estos métodos constituyen una de las áreas menos desarrolladas en el ámbito del análisis de clasificación.

A pesar de ello, creemos que, en determinados contextos, la aplicación de algoritmos que produzcan clases no disjuntas puede ser beneficiosa. Así puede ocurrir, por ejemplo, si se estudia la proximidad comercial de espacios económicos distintos. Supóngase que el conjunto de objetos lo formen las islas de un archipiélago y que se desea clasificarlas en función de sus semejanzas en cuanto a los problemas de transporte que sufren sus empresas en el comercio exterior, de modo que pueda valorarse la existencia o no de la llamada *doble insularidad*, derivada del hecho de que las islas mayores económicamente están mejor comunicadas con el exterior

que las demás. Supóngase, además, que existen dos islas mayores, A y B, (mejor establecidas para el comercio exterior) y que una de las restantes islas menores, C, realice su comercio con el exterior exclusivamente a través de la isla mayor A. En este contexto, parece acertada la aplicación de un algoritmo de recubrimiento, ya que tendría explicación la aparición de una clase formada por las dos islas mayores (objetos A y B) y otra clase formada por los objetos A y C, apareciendo así el objeto A en dos clases.

En este artículo, se presentan dos nuevos algoritmos que permiten hallar los clusters maximales del grafo de proximidad $G_\varepsilon = (V,E)$, definido en el conjunto de objetos O por una desemejanza D a un nivel prefijado ε, y en los que se acepta el recubrimiento o las intersecciones no vacías entre los clusters resultantes.

El primero es el Algoritmo Inductivo en los Objetos, y el segundo el Algoritmo de Búsqueda de Objetos Representativos de las clases resultantes. Este último presenta además el aspecto novedoso de aportar el elemento representativo de las clases, definido éste como el elemento cuya suma de distancias al resto de elementos de la clase es menor. Por este hecho, se puede decir que este segundo algoritmo es un híbrido entre la filosofía que envuelve a los métodos de recubrimiento y la propia de los métodos de partición iterativa.

2. ALGORITMO DE CLASIFICACIÓN INDUCTIVO EN LOS OBJETOS

Definición 2. Sea $O = \{o_1,\ldots,o_n\}$ el conjunto de objetos, y D una función de desemejanza definida sobre O. Entonces, dado un nivel de desemejanza $\varepsilon > 0$ arbitrario, una clase o cluster sobre O es un subconjunto C de O tal que:

$$\forall o_i, o_j \in C, \quad D(o_i, o_j) \leq \varepsilon$$

Definición 3. Dado $\varepsilon > 0$, sea el grafo $G_\varepsilon = G_\varepsilon^n \{V_n, A\}$ donde

$$V_n = \{1,2,\ldots,n\} \text{ y } A = \{(i,j) \in V_n^2 \ / \ D(o_i,o_j) \leq \varepsilon\}$$

Definición 4. Se dice que un subconjunto de objetos es maximal si no está contenido en ningún otro.

Según estas definiciones, la clasificación de los objetos O, con nivel de desemejanza $\varepsilon > 0$, está formada por los subconjuntos o cliques maximales en el grafo G_ε.

Se denota con L_ε^k, los cliques maximales del grafo parcial G_ε^k, formado por el conjunto de vértices $V_k = \{1,2,...,k\}$ correspondientes a los objetos $\{o_1,...,o_k\}$. El algoritmo construye L_ε^{k+1}, una vez conocido L_ε^k.

Sean:

$U_k = \{i \leq k$ tal que $D(o_i,o_{k+1}) \leq \varepsilon\}$

$L_\varepsilon^k = \{C_1,C_2,...,C_m\}$ los m clusters maximales de V_k

$F_k = \{C_1',C_2',...,C_m'\}$ donde $C_j' = (C_j \cap U_k) \cup \{k+1\}$

$G_k = \{C_1,C_2,...,C_m,C_1',C_2',...,C_m'\}$

Teorema 1. Sea H_{k+1} la familia de elementos maximales en G_k, respecto de la relación de inclusión. Se verifica que $H_{k+1} = L_\varepsilon^{k+1}$.

Demostración:

a) Sea $B \in L_\varepsilon^{k+1}$ y k+1 un vértice del grafo. Se pueden dar dos situaciones:

i) $(k+1) \notin B \Rightarrow B \in L_\varepsilon^k \Rightarrow \exists\, C_j \in G_k$ maximal tal que $B = C_j \Rightarrow B \in H_{k+1}$.

ii) $(k+1) \in B$ y sea $B' = B - \{k+1\} \subseteq U_k$. Por ser $B \in L_\varepsilon^{k+1} \Rightarrow \exists\, C_j \in L_\varepsilon^k$ tal que $B' \subseteq C_j$ y, por tanto, $B' \subseteq C_j \cap U_k$. Pero $B' \not\subset C_j \cap U_k$, $\forall j$, ya que, entonces, B no sería un clique maximal. Por tanto, $B' = C_j' = C_j \cap U_k$ para al menos un j, y, como consecuencia, $B \in H_{k+1}$.

b) Es claro que cualquier elemento H_{k+1} es un clique maximal del grafo G_ε^{k+1}, y, por tanto, pertenece a L_ε^{k+1}.

El Algoritmo Inductivo en los Objetos es convergente y consta de 6 pasos:

Paso 0: Se toma k = 1, $L_\varepsilon^k = \{1\}$.

Paso 1: Se calcula U_k.

Paso 2: Se calcula F_k.

Paso 3: Se halla la familia $F_k^* = \{D_1,D_2,...,D_{m'}\}$, formada por los elementos maximales de F_k. Obviamente, $m' \leq m$. Sea $I_{m'} = \{D_1,D_2,...,D_{m'}\}$ y hacemos $L_\varepsilon^{k+1} = \{\varnothing\}$.

Paso 4: Si existe un $j \in \{1,...,m'\}$ tal que $C_i \subset D_j$, se realizan las siguientes igualdades: $L_\varepsilon^{k+1} = L_\varepsilon^{k+1} \cup \{D_j\}$, $m' = m'-1$, $I_{m'} = \{D_1,D_2,...,D_{j-1},D_{j+1},...,D_{m'}\}$, con reordenación de índices. Se vuelve al paso 5. Si no existe un $j \in \{1,...,m'\}$ tal que $C_i \subset D_j$, se deja como estaba definido $I_{m'}$, se toma $L_\varepsilon^{k+1} = L_\varepsilon^{k+1} \cup \{C_i\}$ y se va al paso 5.

Paso 5: Se pone i = i+1. Si $i \leq m$, se va al paso 4; en otro caso, $L_\varepsilon^{k+1} = L_\varepsilon^{k+1} \cup \{I_{m'}\}$; se reordenan los nuevos clusters, y se va al paso 6.

Paso 6: Se hace k = k+1. Si k < n, se va al paso 1. En caso contrario, el algoritmo finaliza.

Ejemplo 1. Aplicación del algoritmo inductivo en los objetos sobre una matriz de desemejanzas dada. Supóngase la siguiente matriz de desemejanzas arbitraria referida a 10 objetos:

	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}
O_1	0	12	23	34	45	11	23	22	16	9
O_2	12	0	13	17	12	21	47	32	19	24
O_3	23	13	0	7	13	35	27	29	15	14
O_4	34	17	7	0	20	18	16	39	32	27
O_5	45	12	13	20	0	25	36	39	26	13
O_6	11	21	35	18	25	0	11	15	17	41
O_7	23	47	27	16	36	11	0	32	23	18
O_8	22	32	29	39	39	15	32	0	14	29
O_9	16	19	15	32	26	17	23	14	0	8
O_{10}	9	24	14	27	13	41	18	29	8	0

En esta matriz de desemejanzas, la desemejanza media es igual a 22.644, la desemejanza mediana es 21. Sin embargo, existen tres desemejanzas modales: 13, 23 y 32. Son estas cuatro últimas desemejanzas las elegidas para la presentación de las clases resultantes tras la aplicación del Algoritmo Inductivo en los Objetos sobre la matriz de desemejanzas. Si se toma $\varepsilon = 13$, la secuencia de pasos sería la siguiente:

Paso 0: k=1. $L^1_{13} = \{1\}$.

Paso 1: $U_1 = \{\varnothing\}$.

Paso 2: $F_1 = \{2\}$, ya que $\{1\} \cap \{\varnothing\} \cup \{k+1\} = \{\varnothing\} \cup \{2\} = \{2\}$.

Paso 3: $F_1^* = \{2\}$. $I_1 = \{2\}$. $L^2_{13} = \{\varnothing\}$, i=1.

Paso 4: Como no hay algún $C_1 \subset D_j$, entonces: $I_{m'} = \{2\}, L^2_{13} = \{1\}$.

Paso 5: i=2. Como 2>1, entonces $L^2_{13} = \{1\}$.

Paso 6: k=2. Como 2<10, se vuelve al paso 1.

Y así sucesivamente, hasta comparar todos los objetos. Los resultados para todos los ε considerados fueron:

$\varepsilon = 13$	$\varepsilon = 21$	$\varepsilon = 23$	$\varepsilon = 32$
$\{O_1,O_2\}$	$\{O_2,O_3,O_4,O_5\}$	$\{O_2,O_3,O_4,O_5\}$	$\{O_2,O_4,O_5,O_6\}$
$\{O_3,O_4\}$	$\{O_2,O_4,O_6\}$	$\{O_2,O_4,O_6\}$	$\{O_4,O_6,O_7,O_9\}$
$\{O_2,O_3,O_5\}$	$\{O_4,O_6,O_7\}$	$\{O_4,O_6,O_7\}$	$\{O_1,O_6,O_7,O_8,O_9\}$
$\{O_1,O_6\}$	$\{O_2,O_3,O_9\}$	$\{O_1,O_2,O_3,O_9\}$	$\{O_1,O_2,O_6,O_8,O_9\}$
$\{O_6,O_7\}$	$\{O_1,O_2,O_6,O_9\}$	$\{O_1,O_2,O_6,O_9\}$	$\{O_2,O_3,O_4,O_5,O_9,O_{10}\}$
$\{O_8\}$	$\{O_6,O_8,O_9\}$	$\{O_1,O_6,O_7,O_9\}$	$\{O_3,O_4,O_7,O_9,O_{10}\}$
$\{O_1,O_{10}\}$	$\{O_3,O_5,O_{10}\}$	$\{O_1,O_6,O_8,O_9\}$	$\{O_1,O_2,O_3,O_8,O_9,O_{10}\}$
$\{O_5,O_{10}\}$	$\{O_7,O_{10}\}$	$\{O_3,O_5,O_{10}\}$	$\{O_1,O_3,O_7,O_8,O_9,O_{10}\}$
$\{O_9,O_{10}\}$	$\{O_3,O_9,O_{10}\}$	$\{O_1,O_3,O_9,O_{10}\}$	
	$\{O_1,O_9,O_{10}\}$	$\{O_1,O_7,O_9,O_{10}\}$	

3. ALGORITMO DE BÚSQUEDA DE OBJETOS REPRESENTATIVOS

Las técnicas algorítmicas más usuales entre las de partición iterativa son de dos tipos, según se conozca o no el número de clusters que se desean hallar. Cuando este número es conocido, por ejemplo k, las técnicas más habituales son las del tipo k-medias. No obstante, estas técnicas no parecen haber sido desarrolladas admitiendo la existencia de clases finales no disjuntas.

Las técnicas del tipo k-medias (Anderberg, 1973; Hartigan, 1975; Sánchez García, 1978) parten inicialmente de k clusters (o de k-objetos representativos de los k-clusters), para iterativamente calcular los k-objetos representativos de los clusters (o los k-clusters que mejor configuran los k-objetos representativos de los clusters) hasta obtener la convergencia de este proceso.

Definición 5. Si se denota por C_i un cluster genérico, el objeto representativo de C_i es O_i^* ($O_i^* \in C_i$) si minimiza la suma de las desemejanzas a los objetos de C_i, es decir,

$$\sum_{o_j \in C_i} D(o_i^*, o_j) \leq \sum_{o_j \in C_i} D(o, o_j), \qquad \forall o \in C_i$$

Definición 6. Conocidos los k objetos representativos de los clusters, es decir, $o_1, o_2, \ldots, o_k$, la clase o cluster C_j se forma como

$$C_j = C_j(o_j^*) = \left\{ o_i \in O \,/\, D(o_i, o_j^*) \leq D(o_i, o_s^*), \qquad \forall s, \quad 1 \leq s \leq k \right\}$$

Por convenio, cuando la desemejanza de un objeto o_j, con respecto a varios objetos o_i^* es la misma, se asigna el objeto o_j al cluster cuyo objeto representativo presente menor subíndice.

El problema que plantean este tipo de técnicas es el carácter de mínimo local de la solución, hecho que obliga a ampliar la técnica hasta obtener una convergencia más general, de tipo global. Este principio es el que subyace en otras técnicas en las que el número de clusters no se fija a priori, por lo cual las restricciones que operan en el procedimiento de búsqueda del óptimo son más débiles.

Ahora bien, cuando no se conoce de antemano el número de clases o clusters, el proceso para buscar representativos es más complejo. Expondremos seguidamente una técnica de clasificación cuyos fundamentos son similares a la tradicional técnica ISODATA (Anderberg,1973; Sánchez García, 1978), aunque admite, y en ello radica su originalidad, la posibilidad de que un elemento pueda pertenecer a más de una clase.

El Algoritmo de Búsqueda de Objetos Representativos en las clases resultantes consta de 7 pasos, siendo r un parámetro que permite garantizar un número mínimo (n/r) de clusters, y del que dependen las posibilidades de solapamiento.

Paso 1: Se parte de los n objetos, como representativos de los n clusters iniciales y se elige $\varepsilon > 0$, $2 \leq r < n/2$.

Paso 2: Para cada objeto o_j, se construye el cluster $C_{1j} = C_1(o_j)$ que es el conjunto de los o_i tal que $D(o_j,o_i) \leq \varepsilon$, prescindiendo de los grupos que no sean maximales.

Paso 3: Si el número de clusters hallados en el paso 2 es n, se toma $\varepsilon = \varepsilon + \Delta\varepsilon$ y se regresa al paso 2(3). Si el número de clusters es menor que n/r, se toma $\varepsilon = \varepsilon - \Delta\varepsilon$ y se vuelve al paso 2. En otro caso, se va al paso 4.

(3) El incremento en el parámetro ε debe especificarse en función de la magnitud de las desemejanzas.

Paso 4: Se asigna a cada cluster C_{1j}, el objeto $o_j^{'}$ que sea representativo, es decir, el objeto o_j que minimice la suma de las desemejanzas de los restantes elementos del cluster a dicho objeto.

Paso 5: Se asigna cada objeto o_j a todos los clusters que tengan por objetos representativos $o_i^{'}$, siempre que se verifique que $D(o_j, o_i^{'}) \leq \varepsilon$.

Los pasos 4 y 5 se repiten hasta que los clusters que se obtienen, añadiendo objetos a los representativos, sean los mismos clusters para los que se obtuvieron dichos representativos. Una vez alcanzada la convergencia, se va al paso 6.

Paso 6: Unir los clusters cuyos objetos representativos presenten una desemejanza menor o igual que ε con respecto al objeto representativo del primer cluster resultante del paso 5. A continuación, y entre los clusters que no hayan participado en la unión anterior, se unen aquéllos cuyo objeto representativo presente una desemejanza menor o igual que ε con respecto al primero de ellos. Esta operación se repite hasta que no es posible realizar más uniones. Si no se produce alguna unión entre los clusters resultantes del paso 5, el algoritmo termina; en caso contrario, se va al paso 7.

Paso 7: Se hallan los objetos representativos de los nuevos clusters, y se asigna cada objeto o_i a los clusters cuyo objeto representativo presente una desemejanza menor o igual que ε con respecto a o_i. Los objetos que no cumplen esta condición quedan sin asignar, formando cada uno de ellos un nuevo cluster. En el caso de que todos los objetos estén asignados, el algoritmo para; en otro caso, cada uno de los objetos no asignados forma un nuevo cluster y se vuelve al paso 4.

La convergencia del algoritmo es consecuencia de que los clusters que se han unido alguna vez en el paso 6, no pueden volver a unirse si el algoritmo conduce a ello cuando se repita el paso 6, y, por tanto, el número de elementos de estos clusters no puede disminuir, apareciendo a partir de una determinada iteración los mismos clusters, momento en el cual el algoritmo finaliza.

Ejemplo 2. Aplicación del algoritmo que busca los objetos representativos en las clases resultantes sobre la matriz de distancias dada en el ejemplo 1.

Los resultados, en este caso, para r=2, son los siguientes, teniendo en cuenta que en el caso de que una clase posea dos objetos, el representativo es cualquiera de los dos.

$\varepsilon = 13$	$\varepsilon = 21$
$\{O_1,O_2,O_6,O_{10}\}$ Representativo: O_1	$\{O_5\}$ Representativo: O_5
$\{O_2,O_5\}$ Representativo: O_2 u O_5	$\{O_7,O_4,O_6,O_{10}\}$ Representativo: O_7
$\{O_3,O_4\}$ Representativo: O_3 u O_4	$\{O_9,O_1,O_2,O_3,O_6,O_8,O_{10}\}$ Representativo: O_9
$\{O_6,O_7\}$ Representativo: O_6 u O_7	
$\{O_8\}$ Representativo: O_8	
$\{O_9,O_{10}\}$ Representativo: O_9 u O_{10}	

$\varepsilon = 23$	$\varepsilon = 32$
$\{O_4,O_2,O_3,O_5,O_6,O_7\}$ Representativo: O_4	$\{O_9,O_1,O_2,O_3,O_4,O_5,O_6,O_7,O_8,O_{10}\}$ Representativo: O_9
$\{O_9,O_1,O_2,O_3,O_6,O_8,O_{10}\}$ Representativo: O_9	

4. CONCLUSIONES

Siendo conscientes de que los métodos de clasificación que permiten el solapamiento entre las clases obtenidas no tienen demasiada presencia en la literatura sobre clasificación, como consecuencia, al menos en parte, de la dificultad que plantea su interpretación, se proponen dos nuevos algoritmos convergentes cuya aplicación, sin embargo, es beneficiosa en ciertos contextos.

REFERENCIAS

ALDENDERFER, M.S. y BLASHFIELD, R.K. (1984): *Cluster Analysis.* Series: Quantitative Applications in the Social Sciences. California: Sage Publications, Inc.

ANDERBERG, M.R. (1973): *Cluster Analysis for Applications.* New York: Academic Press.

BAILEY, K.D. (1994): *Typologies and Taxonomies. An introduction to Classification Techniques.* London: Sage Publications, Inc.

EVERITTT, B.S. (1993): *Cluster Analysis.* Great Britain: Edward Arnold.

GOWER, J.C. (1967): «A comparison of some methods of cluster analysis», *Biometrics*, 23, pp. 623-628.

HARTIGAN, J.A. (1975): *Clustering Algorithms.* New York: John Wiley and Sons.

JARDINE, N. y SIBSON, R. (1971): *Mathematical taxonomy.* New York: John Wiley and Sons.

JOHNSON, S.C. (1967): «Hierarchical Clustering Schemes», *Psychometrika*, 32, 3, 241-254.

MACQUEEN, J. (1966): «Some methods for classfication and analysis of multivariate observations», *Proceedings of 5th Symposium of Mathematics, Statistics and Probability*, vol. 1, pp. 281-297. Universtity of California Press. Berkeley.

MCQUITTY, L.L. (1965): «A conjunction of rank order typal analysis and item selection», *Educational and Psychological Measurement*, 25, pp. 949-961.

MCQUITTY, L.L. (1967): «Expansion of similarity analysis by reciprocal pairs of discrete and continuous data», *Educational and Psychological Measurement*, 27, pp. 253-255.

SÁNCHEZ GARCÍA, M. (1978): *Métodos estadísticos aplicados al tratamiento de datos.* Madrid: CCVC.

SOKAL, R.R. y MICHENER, C.D. (1958): «A statistical method for evaluating systematic relationships», *University of Kansas Science Bulletin*, 38, pp. 1409-1438.

SOKAL, R.R. y SNEATH, P.H.A. (1963): *Principles of Numerical Taxonomy.* W.H. Freeman. San Francisco.

WARD, J.H. (1963): «Hierarchical grouping to optimize an objective function», *Journal of American Statistical Association*, 58, pp. 236-244.

NEW NON-HIERARCHICAL ALGORITHMS IN CLASSIFICATION OF DATA

SUMMARY

Unlike hierarchical methods, non-hierarchical methods of cluster analysis have not been extensively used or examined. One kind of non-hierarchical methods, clumping techniques, often suffers this problem. We, therefore, present in this article, two new algorithms which, on the one hand, are non hierarchical and, on the other hand, generate non-empty intersections between final classes.

Key words: Cluster Analysis, Non-hierarchical Methods, Clumping Techniques, Optimization Techniques, Algorithms.

AMS Classification: 62H30