Slávka Bodjanová

Hard and fuzzy classification within the framework of hierarchical and optimization clustering

Kybernetika, Vol. 28 (1992), No. Suppl, 81--84

Persistent URL: http://dml.cz/dmlcz/124208

SUPPLEMENT TO KYBERNETIKA VOLUME 28 (1992), PAGES 81–84

HARD AND FUZZY CLASSIFICATION WITHIN THE FRAMEWORK OF HIERARCHICAL AND OPTIMIZATION CLUSTERING

SLAVKA BODJANOVA

The ultimate goal of any clustering technique is to divide a given set of objects into a few classes (clusters) in such a way that the data points assigned to one cluster be as similar as possible to each other and as dissimilar as possible to points assigned to other clusters. Most clustering algorithms are based either on hierarchical or nonhierarchical techniques. In this paper attention will be given to fuzzy and hard optimization partitioning and to divisive hierarchical partitioning. A general form of algorithm including fuzzy, hard, hierarchical and optimization approaches to classification of finite data sets is proposed.

1. BASIC NOTATIONS

Let $X = \{x_1, x_2, \dots, x_n\}$ be a set of n objects (observations). The partition space including the hard partitions of X into k clusters $(2 \le k \le n)$ was introduced by Bezdek [2] as follows:

$$P_k = \left\{ U \in V_{kn} :\ u_{ij} \in \{0,1\} \text{ for all } i, j;\ \sum_i u_{ij} = 1 \text{ for all } j;\ \sum_j u_{ij} > 0 \text{ for all } i \right\},$$

where V_{kn} is the usual vector space of real $k \times n$ matrices.
Analogously, the partition space of fuzzy partitions was defined by

$$P_{fk} = \left\{ U \in V_{kn} :\ u_{ij} \in [0,1] \text{ for all } i, j;\ \sum_i u_{ij} = 1 \text{ for all } j;\ \sum_j u_{ij} > 0 \text{ for all } i \right\}.$$

Dumitrescu defined a hierarchical sequence of fuzzy partitions [3] as follows:

Definition 1. A sequence $U^{(0)}, U^{(1)}, \dots, U^{(q)}$ of fuzzy partitions is called hierarchical if $U^{(t)}$ is a refinement of $U^{(t-1)}$ for $t = 1, \dots, q-1$.

Definition 2. Let $U \in P_{fk}$ and $W \in P_{fm}$ be two fuzzy partitions of $X = \{x_j\}$, $j = 1, \dots, n$. The partition W is said to be a refinement of U if

i) $m \ge k$,

ii) there exists a partition $I^{(1)}, \ldots, I^{(k)}$ of $\{1, \ldots, n\}$ such that

$$u_i = \bigcup_{s \in I^{(i)}} w_s \qquad \text{and} \qquad u_i \cap \bigcup_{s \notin I^{(i)}} w_s = \emptyset,$$

where for $i = 1, \ldots, k$ and $j = 1, \ldots, n$:

$$(u_i \cap w_s)(x_j) = \max\{u_{ij} + w_{sj} - 1,\, 0\} \quad \text{and} \quad (u_i \cup w_s)(x_j) = \min\{u_{ij} + w_{sj},\, 1\}.$$

2. SEQUENCES OF HARD AND FUZZY PARTITIONS

Most real-world classes of objects are fuzzy in nature, however, from a practical point of view we often want to replace them by their hard approximation. Two main ways of mutual combination of hard and fuzzy partitioning result from practice: fuzzification and defuzzification.

Fuzzification can be described by a mapping $\xi_1 : P_k \longrightarrow P_{fk}$ which assigns to each hard partition $W \in P_k$ its fuzzy version $U \in P_{fk}$. On the other hand, defuzzification is realized by a mapping $\xi_2 : P_{fk} \longrightarrow P_k$ which assigns to each $U \in P_{fk}$ its hard approximation $W \in P_k$.

2.1. Iterative sequences of hard and fuzzy partitions

Using fuzzification as well as defuzzification in any optimization clustering technique we can obtain an iterative sequence of hard and fuzzy partitions $W^{(t)} \in P_k$, $U^{(t)} \in P_{fk}$ where $t = 0, 1, \ldots, q$ is the number of iterations. Obviously, $U^{(t)} = \xi_1\left(W^{(t)}\right)$, $W^{(t+1)} = \xi_2\left(U^{(t)}\right)$ and $W^{(0)}$ $\left(U^{(0)}\right)$ is the initial hard (fuzzy) partition. In this case a criterion (optimization) function is defined on $P_k \times P_{fk}$ so that $\psi : P_k \times P_{fk} \longrightarrow \mathsf{R}$, $\psi\left(W^{(t)}, U^{(t)}\right) = \psi\left(\xi_2(U^{(t-1)}, \xi_1(W^{(t)})\right)$. We can use the following general algorithm to obtain an iterative sequence of hard and fuzzy partitions.

Algorithm 1.

Input: Give the data set X, the number of clusters k, the mappings ξ_1, ξ_2 and the criterion function ψ.

Step 1: Initialization
Guess a partition $W^{(0)} \in P_k$ as the initial estimate of the partition of X. Put $t = 0$.

Step 2: Fuzzification
Assign a fuzzy partition $U^{(t)} \in P_{fk}$ to the partition $W^{(t)} \in P_k$ according to ξ_1.

Step 3: Compute value of criterion function $\psi\left(W^{(t)}, U^{(t)}\right)$.

Step 4: Defuzzification
Assign a hard partition $W^{(t+1)} \in P_k$ to the partition $U^{(t)} \in P_{fk}$ with respect to ξ_2 and ψ.

Step 5: Compare partitions
If $W^{(t)} = W^{(t+1)}$ then stop. Else increase t by 1 and go to Step 2.

2.2. Hierarchical sequences of hard and fuzzy partitions

Using fuzzification in a hierarchical clustering we can obtain a hierarchical sequence of hard and fuzzy partitions. Many divisive hierarchical clustering methods are based on binary splitting. Using this technique, a chosen cluster $w_m \in W \in P_k$ is split into two clusters w_{m_1} and w_{m_2} according to a binary splitting rule $\mathcal{R}_B$. Let $U^{(t)} = \xi_1\left(W^{(t)}\right) \in P_{ft}$ be a tth term of a hierarchical sequence $U^{(0)}, U^{(1)}, \ldots, U^{(q)}$ of fuzzy partitions of $X = \{x_j\}$, $j = 1, \ldots, n$. If the cluster w_m is split into two clusters w_{m_1}, w_{m_2} then the cluster $u_m \in U^{(t)}$ is split into two clusters u_{m_1}, u_{m_2} such that

$$u_{m_1}(x_j) = \xi_1\left(w_{m_1}(x_j)\right) \cdot u_m(x_j) \qquad \text{and} \qquad u_{m_2}(x_j) = \xi_1\left(w_{m_2}(x_j)\right) \cdot u_m(x_j).$$

A hierarchical sequence of hard and fuzzy partitions can be obtained by the following general algorithm.

Algorithm 2.

Input: Give the data set X, the mapping ξ_1 (ξ_2), the spliting rule $\mathcal{R}_B$, the rule $\mathcal{R}_C$ for choice of cluster to be split and the rule $\mathcal{R}_H$ for halting the process of splitting.

Step 1: Initialization
Let $W^{(0)} = U^{(0)} \in P_1$ such that $w_1^{(0)}(x_j) = 1$ for all $x_j \in X$. Put $t = 1$.

Step 2: Splitting
Split the cluster $w_t^{(t-1)}$ into two clusters $w_t^{(t)}$ and $w_{t+1}^{(t)}$ according to the rule $\mathcal{R}_B$. Create a partition

$$W^{(t)} = \left\{W^{(t-1)} - \left\{w_t^{(t-1)}\right\}\right\} \cup \left\{w_t^{(t)}, w_{t+1}^{(t)}\right\}.$$

Step 3: Fuzzification
Create a partition $U^{(t)} = \xi_1\left(W^{(t)}\right)$ such that $U^{(t)} = \left\{U^{(t-1)} - \left\{u_t^{(t-1)}\right\}\right\} \cup \left\{u_t^{(t)}, u_{t+1}^{(t)}\right\}$ where $u_t^{(t)} = \xi_1\left(w_t^{(t)}\right) \cdot u_t^{(t-1)}$ and $u_{t+1}^{(t)} = \xi_1\left(w_t^{(t)}\right) \cdot u_t^{(t-1)}$.

Step 4: (optional) Defuzzification
It is possible to replace the partition $W^{(t)} \in P_{t+1}$ by the partition $W_*^{(t)} = \xi_2\left(U^{(t)}\right)$ created as follows:

$$W_*^{(t)} = \left\{W^{(t-1)} - \left\{w_t^{(t-1)}\right\}\right\} \cup \left\{\xi_2(w_t^{(t)}), \xi_2(w_{t+1}^{(t)})\right\}.$$

Step 5: Testing
If $W^{(t)} \in P_n$ or $W^{(t)}$ satisfies the rule $\mathcal{R}_H$ then stop. Else use rule $\mathcal{R}_C$ to choose the cluster $w_m^{(t)} \in W^{(t)}$ to be split. Then increase t by 1, interchange clusters $w_m^{(t)}$ and $w_t^{(t)}$, and go to Step 2.

2.3. Iterative sequences and hierarchical sequences of hard and fuzzy partitions

Algorithm 1 and Algorithm 2 lead to the method (Algorithm 3) which produces an hierarchical sequence of q hard and fuzzy partitions (Algorithm 2) such that each partition $W^{(t)} \in P_t \quad (U^{(t)} \in P_{ft})$ is a result of an iterative sequence of hard and fuzzy partitions (Algorithm 1) at each hierarchical level $t = 0, 1, \ldots, q$.

Algorithm 3.

Input: Input from Algorithm 2 including mapping ξ_2.

Step 1: Initialization
(Step 1 from Algorithm 2)

Step 2: Splitting
(Step 2 from Algorithm 2)

Step 3: Fuzzification
(Step 3 from Algorithm 2)

Step 4: Compute criterion function
(Step 3 from Algorithm 1)

Step 5: Defuzzification
(Step 4 from Algorithm 2)

Step 6: Compare partitions
If $W_*^{(t)} \neq W^{(t)}$ then replace $W^{(t)}$ by $W_*^{(t)}$ and go to Step 3.
Else go to Step 7.

Step 7: Testing
(Step 5 from Algorithm 2)

We used Algorithm 3 (Algorithm 1, Algorithm 2) for creating a computer program SPLITAF and for the classification of 42 Slovak districts according to their values observed in 6 demographic features. Backer's method of optimal decomposition of induced fuzzy sets [1] was applied to specify fuzzification (ξ_1) and defuzzification (ξ_2) and the MacNaughton-Smith algorithm with several different rules $\mathcal{R}_C$ and $\mathcal{R}_H$ was used to perform binary splitting.

REFERENCES

[1] E. Backer: Cluster Analysis by Optimal Decomposition of Induced Fuzzy Sets. Delftse Universitaire Press, Delft 1978.

[2] J. C. Bezdek: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York 1981.

[3] D. Dumitrescu: On fuzzy partitions in cluster analysis. Seminar Angeluta, pp. 105–108. Cluj-Napoca 1983.

Slavka Bodjanova, Department of Statistics, Institute of Economics, Bratislava. Czechoslovakia.