Classification Algorithms Based on Core Seeking in Sequences of Nested Monotone Systems *

M. N. Dumbadze UDC 519.237.8

Some properties of minimal (by inclusion) cores of a monotone system [1] are investigated. A procedure is proposed for the construction of a classification algorithm based on successive extraction of minimal (by inclusion) cores of nested monotone systems. Special monotone systems are considered, for which the minimal (by inclusion) cores can be found relatively quickly.

1. General outline of the algorithms

The application of the apparatus of monotone systems for classification of multidimensional empirical data was proposed in [2]. Let $W=\{1,2,...,N\}$ be the set of objects being analyzed ($|W|=N$), and on the couples (i,H), where $i\in H\subseteq W$, define a scalar-valued function $\pi(i,H)$ that satisfies the condition

$$\pi(i,H)\geq\pi(i,H\setminus k)\;\;\forall i,k\in H\subseteq W. \tag{1}$$

Then an effective algorithm exists [3] for finding an extremal subset $G\subseteq W$, defined by the relationship

$$F(G)=\min_{i\in G}\pi(i,G)=\max_{H\subseteq W}\min_{i\in H}\pi(i,H). \tag{2}$$

The set G is called the core of the monotone system $\langle W,\pi,F\rangle$. [1] We stress that the extremal subset, in the sense of (2), generated by the algorithm of [3] is maximal by inclusion. [2]

The classification algorithm proposed has been investigated in [2,4] for various monotone systems and applied problems. It is reducible to the following successive exhaustion procedure on W.

---

* Moscow. Translated from Avtomatica i Telemekhanika, No. 3, pp. 127 – 133, March, 1990. Original article submitted February 13, 1989.
005 – 1179/90/5103 – 0382 $12.50

[1] In [1] the core is called "kernel", but sometimes "nuclei" is used as well. Moreover, [1] uses "monotonic system" vocabulary instead of "monotone system", noticed by JM.

[2] It is shown in [1] that the family of cores of a monotone system is closed under union and maximal core is therefore always unique.

In the first step of the procedure, we construct the core G_1 of the system $\langle W,\pi,F\rangle$, and then we construct the restriction of this system $\langle W\setminus G_1,\pi,F\rangle$. Next we find the core of this restriction. We denote it by G_2, after which the process of restriction and the core seeking on the restricted monotone system is continued until W is exhausted. The sequence of cores $G_1,G_2,\ldots$ constructed in this way determines the sought classification.

Practical experience shows [2] that, even for clearly distinguishable classes, the sequence constructed in this way suffers from a deficiency, namely the set G_1 accounts for most of the cardinality of W, and all other $G_{i\neq 1}$ are sets of low cardinality or even single-element sets.

In order to avoid this shortcoming, some parametric family of such functions replaces the single measure of association of an element with a subset $\pi(i,H)$, and multiple application of the proposed procedure of construction of the sequence $G_1,G_2,\ldots$for various parameter values eventually produces a classification compatible with the researcher's goal. The application of this approach, of course, is severely limited by computational complexity, lack of a goal-directed organization of exact enumeration, and ambiguity of the applied evaluation of the result typical for classification problems. This is the primary felt in cases when constructed classification is used as a working tool for compact representation of large empirical arrays, and the same technique has to be repeatedly applied to the same data [2].

In this context, it is desirable to modify the procedure of successive extraction of cores from contracting monotone systems so as to "protect" it from the deficiency of excessive cardinality of the first core G_1.

The fundamental feasibility of such a modification was obvious: by the properties of the known algorithm [3], G_1 is the largest of the subsets satisfying (2); it would be necessary to construct an algorithm capable of finding cores smaller cardinality, and best of all cores that are smallest by inclusion. One of the goals of our paper is to solve the last problem. Given an exact algorithm, the procedure of successive exhaustion of W by such minimal cores is constructed along the same lines as the successive procedure described above. [3]

[3] Unlike the procedure [2], where by construction $F(G_i)>F(G_{i+1})$ for all i, the proposed procedure may produce whole collections of cores $\{G_i,G_{i+1},\ldots,G_{i+r}\}$ such that $F(G_i)=F(G_{i+1})=\ldots=F(G_{i+r})$.
We point the readers attention to that the procedure exactly in form [2] was actually first described in http://www.datalaundering.com/download/modular.pdf . Noticed by JM

2. Structural properties of the set of cores of monotone systems used in classification algorithms with successive identification of classes

The initial successive procedure proposed in Sec. 1 is based on a property, which was described in detail in [5] in the form of a general polynomial-time algorithm that finds the largest by inclusion core of a monotone system. In this section we construct new algorithm that find one of the minimal by inclusion cores, [4] also for the general case.

These algorithms are based on the following theorems.

Theorem 1. $\tilde{G}$ is a minimal by inclusion core of a monotone system $\langle W,\pi,F\rangle$ if and only if for any $i\in\tilde{G}$ we have

$$F(\tilde{G}_i) < F(G), \tag{3}$$

where $\tilde{G}_i$ is the core of the system $\langle\tilde{G}\setminus i,\pi,F\rangle$, and G is the largest by inclusion core of the system $\langle W,\pi,F\rangle$.

Note that the core G_H of the restricted system $\langle G\setminus H,\pi,F\rangle$, where H is an arbitrary subset of G ($H\subset G$), satisfies one of two mutually exclusive conditions: $F(G_H)=F(G)$ or $F(G_H)<F(G)$, where G is the largest core of the system $\langle W,\pi,F\rangle$.

Theorem 2. Let G and $K=\{G_1,G_2,...,G_\ell\}$ be respectively the largest by inclusion core and the set of all cores of the system [5] $\langle W,\pi,F\rangle$, and let G_H be the core of the restricted system $\langle G\setminus H,\pi,F\rangle$ for any $H\subset G$. If $F(G_H)<F(G)$, then for any i we have $H\cap G_i\neq\varnothing$.

Theorems 1 and 2 are proved in the Appendix.

Recall [5] that the algorithm to find the largest by inclusion core G of the system $\langle W,\pi,F\rangle$ reduces to the construction of the so-called defining sequence $J=\langle i_1,i_2,...,i_N\rangle$ of elements from W and extraction of the largest right-hand segment $J_\mu=\langle i_\mu,i_{\mu+1},...,i_N\rangle$ from this sequence such that

$$\pi(i_\mu,J_\mu)=\max_{\nu=\overline{1,N}}\pi(i_\nu,J_\nu) \tag{4}$$

and μ is the least index for which (4) holds.

---

[4] Any such core, because the number of all minimal by inclusion cores of the system $\langle W,\pi,F\rangle$ may depend exponentially on the cardinality of the set W.

[5] $\forall i\;\; G\supseteq G_i$, where $G_i\in K$.

It follows from [5] that for all $\nu = 1,...,N$,

$$\pi(i_\nu, J_\nu) = \min_{i \in J_\nu} \pi(i, J_\nu) = F(J_\nu), \qquad (5)$$

and therefore any segment J_ν satisfying (4) defines a certain core $\tilde{G}$ of the system $\langle W,\pi,F \rangle$. Identify the smallest among all such cores extracted simultaneously with the identification of the largest by inclusion core G. Denote this smallest core by $\overline{G}$ and the corresponding element – the first element $\overline{G}$ in the sequence J – by $i_{\overline{\mu}}$. Subsequently, use this core as the initial restriction $\langle \overline{G},\pi,F \rangle$ to search for one of the minimal by inclusion cores. This may reduce the amount of computation in our algorithm.

Corollary. For any core $\tilde{G}$ of the system $\langle \overline{G},\pi,F \rangle$, we have $i_{\overline{\mu}} \in \tilde{G}$.
The proof follows directly from Theorem 2.

The corollary and Theorem 2 easily lead to the sought algorithm. We will describe it step by step (we refer to this algorithm as A1).

Step1. Construct a sequence J that satisfies condition (5) for all $\nu = 1,...,N$; identify the largest index $\overline{\mu}$ among those that satisfy condition (4). It defines the core $\overline{G}$ (we treat it as a core $\overline{G}(\alpha)$ of level $\alpha = 1$) and the element $i_{\overline{\mu}} \in \overline{G}$. By the corollary, the sought minimal core G_{min} included in $\overline{G}$ necessarily contains $i_{\overline{\mu}}$ ($i_{\overline{\mu}} \in G_{min} \subseteq \overline{G}$).

Step 2. Set $\alpha = \alpha + 1$. Take the system $\langle \overline{G},\pi,F \rangle$ and consider the family of its restrictions $\langle \overline{G} \setminus i,\pi,F \rangle$ for all $i \in \overline{G}$ except $i_{\overline{\mu}}$, and on each restriction find the core $\overline{G}_i$ (it is extracted by the effective algorithm proposed in [5]).

Two cases are possible:

1) There exists i such that $F(\overline{G}_i) = F(\overline{G})$. This situation may be detected in Step 2 without examining all the systems $\langle \overline{G} \setminus i,\pi,F \rangle$. The first system, which gives the equality $F(\overline{G}_i) = F(\overline{G})$ establishes this fact.

2) For all $i \neq i_{\overline{\mu}}$, $F(\overline{G}_i) < F(\overline{G})$.

1. In case 1), select any element $j \neq i_{\overline{\mu}}$ from the set $\overline{G}_i$ (for definiteness, take the element with the least index in the initial list of elements W) Denote this element by i_1. Replace the system $\langle \overline{G},\pi,F \rangle$ constructed in Step 1 with the system $\langle \overline{G}(\alpha-1),\pi,F \rangle$, where $\overline{G}(\alpha-1) = \overline{G}_i$. Then apply Step 2 on this system.

2. If case 2) applies, this means by Theorem 1 that the minimal by inclusion core has been found. It is the core $\overline{G}$ (on the level $\alpha = 1$) or $\overline{G}(\alpha-1)$ on level α.

In order to improve the computational efficiency of the algorithm, we should strive to detect the first case as soon as possible. At the end of this, we describe the algorithm A2, which is a slight modification of the algorithm A1. Algorithm A2, like A1, consists of two steps, where Step 1 is the same for both algorithms. Step 2 of A2 is described as follows.

After Step 1 of A1, partition $\overline{G}$ into two parts: $H_1 \cup i_{\overline{\mu}}$ and $H_2 = \overline{G} \setminus \left(H_1 \cup i_{\overline{\mu}}\right)$ ($H_1 \cup H_2 \cup i_{\overline{\mu}} = \overline{G}$) with maximally close cardinalities. Here we consider the ordered set $\overline{G}$ produced by Step 1 of A1, and therefore it is easily partitioned by selecting the boundary between the sets H_1 and H_2 in the sequence of the elements of $\overline{G}$. Determine the core $\overline{G}(H_1)$ of the system $\left\langle H_1 \cup i_{\overline{\mu}}, \pi, F \right\rangle$. If $F(\overline{G}(H_1)) = F(\overline{G})$, then $\overline{G}(H_1)$ again should be partitioned into two parts with closest possible cardinalities, and this partitioning is successively continued until case 2 is obtained, $F(\overline{G}_i) < F(\overline{G})$. Case 2 obviously implies that H_2 contains at least one element that belongs to the sought core. Therefore, the search for $F(\overline{G}_i) = F(\overline{G})$ in step 2 of A2 is reduced by selecting an element $i \in H_1$ for the examination of the system $\left\langle \overline{G}, \pi, F \right\rangle$, instead of an arbitrary next element i as in A1. We denote this element by $i(H_1)$. Then all the elements from H_1 are successively examined, and only after that the elements from H_2.

It is easy to see that A2 differs from A1 in that the system $\left\langle \overline{G} \setminus i, \pi, F \right\rangle$ considered for any element $i \in \overline{G}$ ($i \neq i_{\overline{\mu}}$), in Step 2 of A1 is replaced in A2 with the system $\left\langle \overline{G} \setminus H, \pi, F \right\rangle$, where H is a certain subset of the set $\overline{G}$.

Both Algorithms A1 and A2 were coded in FORTRAN IV. The results of machine experiments established definitive superiority of A2. Other more radical opportunities for reducing the computational effort can be realized by considering special monotone systems.

Thus, the proposed algorithm to find a minimal by inclusion core produces a classification algorithm in the form of the following successive procedure. First the minimal by inclusion core G_1 is found in the system $\left\langle W, \pi, F \right\rangle$. Then the minimal by inclusion core G_2 is found in the restriction $\left\langle W \setminus G_1, \pi, F \right\rangle$, and the process of restrictions is continued until W is exhausted.

In conclusion note, that the cores obtained in this way are nonintersecting and we have

Proposition. Let $G_1, G_2, \ldots, G_{|S|}$ be minimal by inclusion cores of the monotone systems $\left\langle W, \pi, F \right\rangle$, $\left\langle W \setminus G_1, \pi, F \right\rangle, \ldots, \left\langle W \setminus G_1 \cup G_2 \cup \ldots \cup G_{|S|-1}, \pi, F \right\rangle$, respectively. Then $F(G_1) \geq F(G_2) \geq \ldots \geq F(G_{|S|})$, where S is some index set.

The proposition is easily proved from the definition of monotone system.

3. Special monotone systems

As we have noted in the previous section, the use of special systems may produce additional computational savings in the search of minimal by inclusion cores. Moreover, the use of special systems sometimes makes it possible to describe the entire core structure, thus largely broadening the capabilities of the algorithms trough introduction of a priori information for the construction of the sought classification.

In this section, we give three examples of such special systems. They have been used in various applications [5,6].

1. Consider the monotone system $\langle W, \pi_1, F \rangle$, where the function $\pi_1(i,H) = \max_{j \in H} a_{ij}$ is defined on the association matrix $A = \| a_{ij} \|$, $i, j \in W$. For all $i, j \in W$ we assume that $a_{ij} \geq 0$, $i \neq j$, $a_{ii} = 0$. Let $S = \{1,2,...,k\}$ be some index set and let G_S for any $s = 1,...,k$ be the diameter of the matrix A, i.e., $G_S = \{i_s, j_s\}$, and $a_{i_s j_s} = \max_{i,j \in W} a_{ij}$.

Then the amount of computations performed by the proposed algorithm in order to find the minimal by inclusion core of such monotone systems is reduced by the following theorem.

Theorem 3. For any $s = 1,...,k$, G_S is minimal by inclusion cores of the monotone system $\langle W, \pi_1, F \rangle$.

Theorem 4. Let $\{G_S, s = 1,...,k\}$ be the set of all minimal by inclusion cores of the monotone system $\langle W, \pi_1, F \rangle$, then $G = \bigcup_{s \in S} G_s$, where $S = \{1,2,...,k\}$.

Theorems 3 and 4 are proved in the Appendix.

2. Consider the monotone system $\langle W, \pi_2, F \rangle$ [1,5], where the function $\pi_2(i,H) = \sum_{j \in H} a_{ij}$ is defined on the association matrix $A = \| a_{ij} \|$, $i, j \in W$. For all $i, j \in W$ we assume that $a_{ij} \geq 0$, $i \neq j$, $a_{ii} = 0$ as before.

Then we have the following theorem.

Theorem 5. A unique minimal by inclusion core exists in the monotone system $\langle W, \pi_2, F \rangle$.

Theorem 5 is proved in the Appendix.

The amount of computations in the proposed algorithm for system $\langle W, \pi_2, F\rangle$ is thus reduced.

Note that minimal by inclusion cores of the monotone system $\langle W, \pi_1, F\rangle$ and $\langle W, \pi_2, F\rangle$ are also cores of smallest cardinality.

3. We know [6] that for P-monotone systems the defining sequence $J^p = \langle i_1^p, i_2^p, ..., i_N^p\rangle$, which coincides with the sequence $J = \langle i_1, i_2, ..., i_N\rangle$, is constructed not by relationship (5) but according to some given order P, which is determined by the values of the functions $\pi(i, W)$ on the set W.

Thus, Step 1 of the algorithm executes much faster for P-monotone systems than for general monotone systems.

APPENDIX

Proof of the Theorem 1. *Necessity*. Let $\tilde{G}$ be a minimal by inclusion core, i.e., for $\tilde{G}$ we have $F(\tilde{G}) = F(G)$ and for any $H \subset \tilde{G}$ we have $F(H) < F(\tilde{G})$. Hence $F(\tilde{G}_H) < F(G)$. If H for any $i \in \tilde{G}$ is chosen as the set $\tilde{G} \setminus i$, then for any core $\tilde{G}_i$ of the system $\langle \tilde{G} \setminus i, \pi, F\rangle$ we obtain that $F(\tilde{G}_i) < F(G)$. This completes the proof of necessity.

Sufficiency is proved by contradiction. Assume that for all $i \in \tilde{G}$, the cores $\tilde{G}_i$ satisfy condition (3). Suppose that the core $\tilde{G}$ is not minimal by inclusion. Then $\exists\, i \in \tilde{G}$, such that for the core $\tilde{G}_i$ of the system $\langle \tilde{G} \setminus i, \pi, F\rangle$ we have $F(\tilde{G}_i) = F(\tilde{G})$, but on the other hand $F(\tilde{G}) = F(G)$, i.e. $F(\tilde{G}_i) = F(G)$, a contradiction with condition (3). The assumption is false. ■

Proof of the Theorem 2. By contradiction. Assume that $\exists\, G_i \in K$ – a core of the system $\langle W, \pi, F\rangle$ such that $H \cap G_i = \varnothing$, where H is some set satisfying the condition of the theorem. This, in turn, implies that G_i is a core of the system $\langle G \setminus H, \pi, F\rangle$, i.e., $F(G_i) = F(G_H) = F(G)$, which obviously excludes the condition $F(G_H) < F(G)$. The assumption is false. ■

Proof of theorem 3. First we will show that the set $G_S = \{i_s, j_s\}$ is a core.

Let G be an arbitrary core of the monotone system $\langle W, \pi_1, F \rangle$. Then by (2), $F(G) = \min_{i \in G} \pi(i, G) = \min_{i \in G} \max_{j \in G} a_{ij} = a_{i_l j_q}$ for some pair of elements $i_l, j_q \in W$.

Clearly,

$$F(G) = a_{i_l j_q} \le \max_{i,j \in W} a_{ij} = a_{i_s j_s}. \qquad (6)$$

On the other hand, we have $F(G) \ge F(H)$, for any $H \subseteq W$ (the function $F(H)$ by definition attains a maximum on the set G). This relationship clearly also holds for $H = G_S = \{i_s, j_s\}$ such that $a_{i_s j_s} = \max_{i,j \in W} a_{ij}$; since $F(H) = \min_{i \in \{i_s, j_s\}} \pi(i, \{i_s, j_s\}) = \min_{i \in \{i_s, j_s\}} \max_{j \in \{i_s, j_s\}} a_{ij} = a_{i_s j_s}$, we have

$$F(G) \ge a_{i_s j_s}. \qquad (7)$$

Using (6) and (7), we obtain $F(G) = a_{i_s j_s}$, i.e., the set G_S is a core.

Moreover, $F(G_S \setminus i_s) = F(\{j_s\}) = 0 < F(G)$ (or $F(G_S \setminus j_s) < F(G)$, i.e. G_S is a minimal by inclusion core of the monotone system $\langle W, \pi_1, F \rangle$. ■

Proof of the Theorem 4. We use the following definition of the largest by inclusion core G of the system $\langle W, \pi, F \rangle$:

$$F(H) < F(G) \ \ \forall H \supset G, \qquad (8)$$

$$F(H) \le F(G) \ \ \forall H \subset G \qquad (9)$$

Let us prove inequality (8) for the set L, where $L = \bigcup_{s \in S} G_s$.

By definition $F(L) = \min_{i \in L} \pi(i, L) = \min_{i \in L} \max_{j \in L} a_{ij} = a_{i_s j_s}$, $s \in \overline{1,k}$. By the central theorem of the theory of monotone systems [1], and also by Theorem 3, the set L is a core, i.e., we have $F(G) = F(\bigcup_{s \in S} G_s)$ or $F(G) = F(L) = a_{i_s j_s}$, $s \in \overline{1,k}$.

Consider some element $t \in W \setminus L$ (if $W \setminus L = \varnothing$, then it remains to prove only (9), i.e., that W is the largest by inclusion core). As H take the set $L \cup t \supset L$. Then for any $s = 1, \ldots, k$ we have $F(H) = \min_{i \in L \cup t} \pi(i, L \cup t) = \min_{i \in L \cup t} \max_{j \in L \cup t} a_{ij} = \max_{j \in L \cup t} a_{tj}$; on the other hand, by definition of the elements i_s, j_s, we obtain $\max_{j \in L \cup t} a_{tj} < a_{i_s j_s}$, i.e., $F(H) < F(L)$. We similarly show that $F(H) < F(L)$ for any $H \supset L \cup t$. We have thus proved (8) for the set L. Let us now prove (9).

As H take the set $L\setminus i_q$ (or $L\setminus j_q$) for any $q\in S$. Clearly,

$$F(L\setminus i_q)=\min_{i\in L\setminus i_q}\pi(i,L\setminus i_q)=\min_{i\in L\setminus i_q}\max_{j\in L\setminus i_q}a_{ij}\le a_{i_s j_s}=F(L).$$

Also clearly $F(L\setminus i_q)\le F(L)$. We similarly show that $F(H)\le F(L)$ for any $H\subset L\setminus i_q$ (or $H\subset L\setminus j_q$).

Thus, the set L satisfies conditions (8) and (9), i.e., $G=L=\bigcup_{s\in S}G_s$. ■

Proof of Theorem 5. We first prove the following lemma.

LEMMA. Let G^λ be the largest by inclusion core of the system $\langle W,\pi_2,F\rangle$ and i_λ the corresponding element, i.e., $\min_{i\in G^\lambda}\pi(i,G^\lambda)=\pi(i_\lambda,G^\lambda)$. Then any subset $H\subset G^\lambda$ that contains the element i_λ is not a core of the monotone system $\langle W,\pi_2,F\rangle$.

Proof of the LEMMA. Consider the set H that satisfies the condition of the lemma, i.e., $i_\lambda\in H\subset G^\lambda$. Since $a_{ij}>0$, $i,j\in W$, $i\ne j$ and $H\subset G^\lambda$, we have

$$\pi(i_\lambda,H)<\pi(i_\lambda,G^\lambda). \qquad (10)$$

On the other hand,

$$\min_{j\in H}\pi(j,H)\le\pi(i_\lambda,H). \qquad (11)$$

Using (10) and (11), we have $F(G^\lambda)=\pi(i_\lambda,G^\lambda)>\min_{j\in H}\pi(j,H)=F(H)$. ■

Let us now prove Theorem 5. Take the set $G^\lambda\setminus i_\lambda$ and find its largest core. Denote it by G_1^λ. Clearly, $G_1^\lambda\subseteq G^\lambda\setminus i_\lambda$. If $F(G_1^\lambda)<F(G^\lambda)$, then the lemma and Theorem 1 have found the sought minimal core; it is the set G^λ. If $F(G_1^\lambda)=F(G^\lambda)$, then apply the lemma to G_1^λ, and so on, until all the cores have been exhausted. We obtain a finite chain of nested sets $G^\lambda\supset G_1^\lambda\supset G_2^\lambda\ldots$. The last core is minimal by inclusion.

Theorem 5 is thus a corollary of the lemma. ■

LITERATURE CITED

1. J. E. Mullat, "Extremal subsystems of monotone systems, I" Avtomat. Telemekh., No. 5, 130-139 (1976), http://www.datalaundering.com/download/extrem01.pdf .
2. L. K. Vykhandu, "Some Methods to Order Objects and Variables in Data Systems", Tr. Tallinsk. Politekh. Inst., No. 482, 43-50 (1980), http://www.datalaundering.com/download/variable.pdf .
3. J. E. Mullat, "Extremal subsystems of monotone systems, II," Avtomat. Telemekh., No. 8, 169-179 (1976), http://www.datalaundering.com/download/extrem02.pdf .
4. R. V. Éeremaa, "Clustering using monotone systems," in Theoretical and Applied Topics of Mathematics, II, Abstracts of Papers at Republican Conf. [in Russian], Tartu (1985).
5. E. N. Kuznetsov, "Analyzing the Structure of the Coupling Matrix with the Aid of the Associated Monotonic System", Avtomat. Telemekh. No. /, pp.128-136 (1980), http://www.datalaundering.com/download/couplem.pdf .
6. E. N. Kuznetsov, I. B. Muchnik and L. V. Shvartser, "Local transformations of monotone systems, I," Avtomat. Telemekh., No. 12, pp. 85-95 (1985), http://www.datalaundering.com/download/ltransfo01.pdf .