Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2015) 1–8

Procedia Computer Science

An Algebra Description for Hard Clustering

Bo Wang[a,b], Yong Shi[a,b,*], Zhuofan Yang[c], Xuchan Ju[d]

[a] *Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences*
[b] *Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences*
[c] *School of Management, University of Chinese Academy of Sciences*
[d] *Collage of Mathematica Science, University of Chinese Academy of Sciences*

Abstract

Hard clustering algorithm partitions data set into several distinct regions. Clustering result offers a kind of characterization for the distribution of data relied on concentration. At the same time, the cluster structure can be regarded as a representation of knowledge in the form of data. However, as a sort of unsupervised learning task, due to a lack of overall criterion for evaluating the effect of clustering algorithms, different clustering algorithms lead to different results based on different considerations. Because of this uncertainty of single clustering result, by virtue of algebra tools, this paper tries to obtain a more reasonable cluster structure by combining various hard clustering results. Furthermore, based on the algebra representation and topologic description of clustering, lattice theory and latticized topology can be employed, which allows us to define algebra operations and discuss topology property on clustering results.

Keywords:
Hard Clustering, Algebra Description, Lattice Theory, Latticized Topology

1. Introduction

Different from goodness of fit based classification problem, as an unsupervised learning task, clustering algorithms offers the results based on several vague principles, which leads to the difficulty of evaluating for the cluster validity [1]. Traditionally, there are three different criteria for evaluating cluster validity, external criterion, internal criterion and relative criterion [2]. External criterion measures the difference between clustering result and predefined cluster structure. Based on similarity matrix, internal criterion compares clustering results under some test of statistical significance. Relative criterion selects the best parameter which is in accordance with the data set. At the same time, it can also tell whether the data set has a distinct cluster structure. However, because of the uncertainty of these criteria and the clustering analysis itself, we cannot ensure which result is the best one. Besides, according to the discussion in [3], clustering can be classified into hard clustering and soft clustering. Typical method in the former type is partition based clustering [4]. Soft clustering employs degree of membership to measure and define different clusters [5].

Lattice theory is an important topic of universal algebra, which mainly studies the algebra structure of arbitrary nonempty set [6]. Based on algebra operations on the set, equivalence relation can be developed to congruence

*Corresponding author
Email address: yshi@ucas.ac.cn (Yong Shi)

relation, which keeps the equivalence under the corresponding algebra operations. Furthermore, all the congruence relations on the set construct a sublattice of the lattice formed by all the equivalence relations on the same set with respect to inclusion. It is clear that every single hard clustering is equivalent to some partitions on the data space. As a result, by means of defining proper operations on the same data space, we found the correspondence between the lattice of congruence relations and all the clustering results on the data set. In this way, the properties of congruence lattice can be extended to the clustering results.

In addition, latticized topology offers the topological structure on lattice. Based on the establishment of 8 different topological bases, the convergence properties can be discussed [7]. On one hand, hard clustering result can be viewed as a non-intersection subset of the power set of data set, which is a typical Boolean algebra. On the other hand, every single clustering result can be seen as an element in the power set of product space formed by data set. Based on algebraic closure operator, clustering result equals to closed set under the operator, which can build another cluster related lattice.

In this paper, we focus on algebra description for hard clustering and also discuss fundamental properties under specific latticized topology. Especially, congruence lattice of a certain algebra system will be used to describe results obtained by one single clustering algorithm or a class of clustering algorithms. By doing this, we will show the relation between clustering results and the congruence lattice on a specific algebra. Then, topological basis built by clustering results can be proposed.

2. Preliminaries

In this section, some related knowledge on lattice theory and latticized topology will be given.

2.1. Algebraic Lattice

Definition 2.1. A binary relation $\leqslant$ defined on a nonempty set A is a partial order on A if the following conditions hold identically in A:

(1) $a \leqslant a$ (reflexivity);
(2) $a \leqslant b$ and $b \leqslant a$ imply $a = b$ (antisymmetry);
(3) $a \leqslant b$ and $b \leqslant c$ imply $a \leqslant c$ (transitivity).

In addition, if $\forall a, b \in A$, we have $a \leqslant b$ or $b \leqslant a$, then, $\leqslant$ is a total order relation. We call a nonempty set A with a partial order $\leqslant$ poset for short, denoted by $(A, \leqslant)$.

Definition 2.2. A poset $(L, \leqslant)$ is a lattice iff for every $a, b \in L$, both $sup\{a, b\}$ and $inf\{a, b\}$ exist (in L). Here, $sup\{a, b\}$ and $inf\{a, b\}$ are denoted by $a \vee b$ and $a \wedge b$ respectively.

Moreover, a sublattice of L is a nonempty subset L' of L, which is closed under $\vee$ and $\wedge$.

Dedekind and Birkhoff characterized nonmodular lattice and nondistributive lattice in the sense of isomorphism respectively [8]. Properties of general lattice are also studied in [9].

Definition 2.3. A poset $(P, \leqslant)$ is complete if for every subset A of P both $sup\ A$ and $inf\ A$ exist (in P). All complete posets are lattices, and a lattice L which is complete as a poset is a complete lattice.

Remark 2.1. Let $Eq(A)$ denote all the equivalence relations on A. It can be proved that poset $(Eq(A), \subseteq)$ is a complete lattice. In addition, if $\{\theta_i\}_{i\in I}$ is a subset of $Eq(A)$, we have $\bigwedge_{i\in I} \theta_i = \bigcap_{i\in I} \theta_i$ and

$$\vee_{i\in I}\theta_i = \cup\{\theta_{i_0} \circ \theta_{i_1} \circ \theta_{i_2} \circ \cdots \circ \theta_{i_k} : i_0, i_1, i_2, \cdots, i_k \in I, k < \infty\}.$$

Particularly, if θ_1 and θ_2 is permutable under the operation $\circ$, we have $\theta_1 \vee \theta_2 = \theta_1 \circ \theta_2$.

Remark 2.2. All the partitions of set A is denoted by $\Pi(A)$. There is a bijection between $\Pi(A)$ and $Eq(A)$.

Definition 2.4 (Compact Element)**.** An element a in lattice L is compact if $\forall a \leqslant \bigvee A$, there is a finite subset B of A, such that $a \leqslant \bigvee B$. L is compactly generated iff every element in L is a supremum of compact elements.

Definition 2.5 (Algebraic Lattice)**.** A lattice is an algebraic lattice iff it is complete and compactly generated.

2.2. Congruence Lattice

For a nonempty set A, we can define language on it.

Definition 2.6. A language (type) $\mathcal{F}$ is a collection of function symbols such that a nonnegative integer n is assigned to each member $f \in \mathcal{F}$. This integer is called the arity (or rank) of f, and f is said to be an n-ary function symbol. The subset of n-ary function symbols in $\mathcal{F}$ is denoted by $\mathcal{F}_n$.

A nonempty set A with a certain language is called algebra, which is denoted by $(\boldsymbol{A}, \mathcal{F})$.

Definition 2.7. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two algebras of the same type. Then $\boldsymbol{B}$ is a subalgebra of $\boldsymbol{A}$ if $B \subseteq A$ and every fundamental operation on $\boldsymbol{B}$ is the restriction of the corresponding operation of $\boldsymbol{A}$, i.e., for each function symbol f, $f^{\boldsymbol{B}}$ is $f^{\boldsymbol{A}}$ restricted to B, simply denoted by $B \leqslant A$. A subuniverse of A is a subset B of A which is closed under the fundamental operations of $\boldsymbol{A}$.

Now, we define congruence lattice and point out that it is an algebraic lattice.

Definition 2.8 (Congruence Lattice)**.** Let $\boldsymbol{A}$ be an algebra of type $\mathcal{F}$ and $\theta \in Eq(A)$. Then θ is a congruence on $\boldsymbol{A}$ if θ satisfies the following compatibility property.

Compatibility Property (CP): for each n-ary function symbol $f^{\boldsymbol{A}} \in \mathcal{F}$ and $a_i, b_i \in A$, if $a_i \theta b_i$, $1 \leqslant i \leqslant n$, then,

$$f^{\boldsymbol{A}}(a_1, a_2, \cdots, a_n) \theta f^{\boldsymbol{A}}(b_1, b_2, \cdots, b_n)$$

holds.

Remark 2.3. The compatibility property is a necessary condition for introducing an algebra structure on the set of equivalence classes A/θ, a quotient structure which is inherited from nonempty set A. In Fig.1, we show a sketch for congruence relation. The dotted lines subdivide A into the equivalence classes of θ. Then selecting a_1, b_1 and a_2, b_2 in the same equivalence class respectively. For binary operation $f^{\boldsymbol{A}}$, compatibility property guarantees $f^{\boldsymbol{A}}(a_1, a_2)$ and $f^{\boldsymbol{A}}(b_1, b_2)$ to be in the same equivalence class.

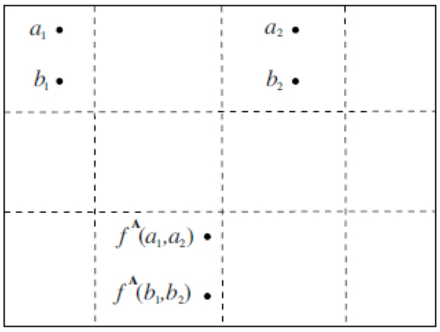

Fig. 1. Congruence Relation

Denote the congruence lattice on $\boldsymbol{A}$ as **Con A**. Based on algebraic closure operator, we have the following theorem.

Theorem 2.1. Con A *is an algebraic lattice.*

2.3. General latticized topology

Let L be a complete lattice. We construct latticized topology on L, which can induce common topology on L.

Definition 2.9. Any element x in a complete lattice L can define a principal filter:

$$[x) = \{y \in L | x \leq y\}.$$

Then, more generalized than principal filter, order filter $N \subsetneqq L$ can be defined as follows: for all $x \in N, x \leq y$, we have $y \in N$.

Remark 2.4. Notice that any principal filter is an order filter. Particularly, for any $N \subset L$, if N is a order filter, then, $\forall x \in N, [x) \subset N$.

Definition 2.10 (Lower Oriented Set)**.** Let L be a complete lattice. N is a lower oriented set iff $N \subset L$ and $\forall x, y \in N, \exists z \in N$, such that $z \leq x, z \leq y$.

Definition 2.11 (Filter)**.** Let L be a complete lattice and $N \subset L$. N is a filter on L iff
(1) N is a order filter on L (full);
(2) N is a lower oriented set (tail).
Denote the set of all the filters on L as $\tau^*(L)$.

Now, we denote the maximum element and minimum element of a complete lattice L as 1 and 0 respectively. Next, define upper neighbor element structure on L.

Definition 2.12. Let $n^* : L \rightarrow \tau^*(L)$. n^* is an upper neighbor element structure on L iff
(1) $n^*(0) = L$;
(2) $\beta \in n^*(\alpha) \Rightarrow \alpha \leq \beta$;
(3) $\beta \in n^*(\alpha) \Rightarrow \exists \gamma \in L, s.t. \alpha \leq \gamma \leq \beta$ and $\forall \delta \leq \gamma, \gamma \in n^*(\delta)$;
(4) $n^*(\bigvee_{t \in T} \alpha_t) = \bigcap_{t \in T} n^*(\alpha_t)$.
Here, $n^*(\alpha)$ is the upper neighbor element system of α.

Definition 2.13. $\gamma \in L$ is an upper uniformly neighbor element with respect to n^* iff $\forall \delta \leq \gamma, \gamma \in n^*(\delta)$.

Proposition 2.1. *$\gamma \in L$ is an upper uniformly neighbor element with respect to n^* iff $n^*(\gamma) = [\gamma)$, i.e. $\gamma \in n^*(\gamma)$.*

Proposition 2.2. *Let Γ_u be the set of all the upper uniformly neighbor elements with respect to n^*, which is also called upper latticized topology on L. It satisfies the following three fundamental properties:*
(1) $0, 1 \in \Gamma_u$;
(2) If $\gamma_1, \gamma_2 \in \Gamma_u$, then $\gamma_1 \wedge \gamma_2 \in \Gamma_u$;
(3) If $\gamma_t \in \Gamma_u, t \in T$, then $\bigvee_{t \in T} \gamma_t \in \Gamma_u$.

Theorem 2.2. *Let $T_{10} = \{(\gamma] | \gamma \in \Gamma_u\} \cup \emptyset \subseteq P(L)$ and $P(L)$ be the power set of L. Then, T_{10} is a topological basis of L.*

Proof. Fist, $\bigcup T_{10} = L$.
Next, let $(\gamma_1], (\gamma_2]$ be two arbitrary elements in T_{10}. Since $(\gamma_1] \cap (\gamma_2] = (\gamma_1 \wedge \gamma_2]$ and $\gamma_1 \wedge \gamma_2 \in \Gamma_u$, we have $\forall \alpha \in (\gamma_1] \cap (\gamma_2], \alpha \in (\gamma_1 \wedge \gamma_2] = (\gamma_1] \cap (\gamma_2]$.
Now, let $\mathfrak{I} = \{U \subset L | \exists \{\gamma_t, t \in T\} \subseteq T_{10}, U = \bigcup_{t \in T} (\gamma_t]\}$. It can be proved that $(L, \mathfrak{I})$ is the unique topology space with respect to T_{10} as its topological basis. □

Remark 2.5. Denote Γ_d as the set of all the lower uniformly neighbor elements with respect to n_*. Here, n_* denotes the lower neighbor element structure. Similarly, we can also dually define $T_{01} = \{[\gamma) | \gamma \in \Gamma_d\} \cup \emptyset \subseteq P(L)$ as a topological basis of L.

3. Algebra characterization for clustering result

In this section, data set is denoted by X and the data space which X belongs to is denoted by X^∞. Now we consider the algebraic representation for clustering result. Especially, for a hard clustering algorithm, it is easy to see that a group of partitions on X^∞ can be corresponded to the clustering result, in the sense of restriction. More specifically, according to *Remark* 2.2, each partition is equal to a certain equivalence relation on X^∞, denoted by R. And the clustering result can be denoted by P, which is also a partition on data set X. In the sense of restriction, we have $P = (X^\infty / R) \bigcap P(X)$. Here, $P(X)$ denotes the collection of all the power sets of X.

3.1. Description for single clustering result

Let C be an operation which can assign every single data point to its corresponding center (or medoid), based on the nearest Euclidean distance. Define a mapping from the data space X^{∞} to the collection of central points X_i, $\alpha_i : X^{\infty} \to X_i, \alpha_i(x) = C(x)$. Now, we need to construct languages on X^{∞} and X_i, which can ensure the mapping α_i above to be a homomorphism. Here, we suppose X^{∞} to be an Euclidean space.

Definition 3.1. Define an n-ary on X^{∞}:

$$f^{\boldsymbol{X^{\infty}}}(x_1, x_2, \cdots, x_n) = C(x_1) + C(x_2) + \cdots + C(x_n), \forall x_i \in X^{\infty}.$$

Remark 3.1. Because X^{∞} is closed under vector addition, the definition above is well-defined.

Definition 3.2. Define an n-ary on X_i:

$$f^{\boldsymbol{X_i}}(c_1, c_2, \cdots, c_n) = C(c_1 + c_2 + \cdots + c_n).$$

Here, $c_1, c_2, \cdots c_n \in X_i$.

Proposition 3.1. *Based on the above definition of n-ary languages on X^{∞} and X_i, α_i is a homomorphism.*

Proof. The proposition can be verified by the following:

$$\begin{aligned} \alpha_i(f^{\boldsymbol{X^{\infty}}}(a_1, a_2, \cdots, a_n)) &= \alpha_i(C(a_1) + C(a_2) + \cdots + C(a_n)) \\ &= C(C(a_1) + C(a_2) + \cdots + C(a_n)) \\ &= f^{\boldsymbol{X_i}}(\alpha_i(a_1), \alpha_i(a_2), \cdots, \alpha_i(a_n)), \forall a_i \in X^{\infty}. \end{aligned} \tag{1}$$

Until now, we obtain a surjective homomorphism mapping α_i between $\boldsymbol{X^{\infty}}$ and $\boldsymbol{X_i}$. □

Remark 3.2. Denote **Con** $\boldsymbol{X^{\infty}}$ as all the congruence relations on the algebra system $\boldsymbol{X^{\infty}}$, where X^{∞} is the universe. As a sublattice of the lattice constituted by all the equivalence relations, **Con** $\boldsymbol{X^{\infty}}$ is a complete lattice.

According to fundamental homomorphism theorem, the kernel of the surjective homomorphism mapping $ker(\alpha_i)$ is an element in **Con** $\boldsymbol{X^{\infty}}$. Let X be the data set, which is a nonempty subset in the vector space. When restricting the universe in X, we obtain a relation between **Con** $\boldsymbol{X^{\infty}}$ and partitions on X. Particularly, every partitions on X can be arranged with a class of elements in **Con** $\boldsymbol{X^{\infty}}$, which can induce an equivalence relation on **Con** $\boldsymbol{X^{\infty}}$. Consequently, $ker(\alpha_i)$ is a representative element in the equivalence class. In addition, the restriction of every coset under $ker(\alpha_i)$ in X is a cluster in the clustering result.

3.2. Description for a class of clustering results

Here, we extend our algebra method to describe a class of clustering results. Without loss of generality, we consider two different Euclidean distance based clustering algorithms on X, whose collections of central points are denoted by X_i and X_j respectively. We define algebra operation on X directly, which makes X be an algebra system.

Definition 3.3. Let α_i be the homomorphism mapping from $\boldsymbol{X^{\infty}}$ to $\boldsymbol{X_i}$. For arbitrary $\{x_1, x_2, \cdots, x_n\} \subseteq X^{\infty}$, define set $S_i(x_1, x_2, \cdots, x_n)$, which satisfies $\forall s \in S_i(x_1, x_2, \cdots, x_n)$:

$$\alpha_i s = f^{\boldsymbol{X_i}}(\alpha_i x_1, \alpha_i x_2, \cdots, \alpha_i x_n).$$

When $S_i(x_1, x_2, \cdots, x_n) \bigcap S_j(x_1, x_2, \cdots, x_n) \neq \emptyset$, for all $\{x_1, x_2, \cdots, x_n\} \subseteq X^{\infty}$, n-ary operation on X^{∞} can be defined as follows: $f^{\boldsymbol{X^{\infty}}}(x_1, x_2, \cdots, x_n) \in S_i(x_1, x_2, \cdots, x_n) \bigcap S_j(x_1, x_2, \cdots, x_n)$.

Remark 3.3. Suppose that there are two different Euclidean distance based clustering algorithms, whose collections of central points are denoted by X_i and X_j respectively. Consider $f^{\boldsymbol{X_i}}$ as the following: $f^{\boldsymbol{X_i}}(c_1, c_2) = C^i(c_1 + c_2), i = 1, 2$, $S_1(x_1, x_2) \bigcap S_2(x_1, x_2) \neq \emptyset$ is equal to the intersection of the equivalence classes of the center corresponding to $C^1(x_1) + C^1(x_2)$ in X_1 and $C^2(x_1) + C^2(x_2)$ in X_2 is nonempty. Here, $C^i, i = 1, 2$ are the mappings arranging point to its nearest center according to two different clustering results respectively. In this case, we can define the value of $f^{\boldsymbol{X^{\infty}}}(a_1, a_2)$ in the intersection. Here, α_1 and α_2 are the surjective homomorphisms from $\boldsymbol{X^{\infty}}$ to $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$, respectively.

Then, this method can be generalized to the case for a class of clustering results as follows.

Definition 3.4. Let f^{X_i} be an n-ary on X_i. If

$$\bigcap_{i \in I} S_i(x_1, x_2, \cdots, x_n) \neq \emptyset, \forall \{x_1, x_2, \cdots, x_n\} \subseteq X^\infty, \tag{2}$$

then the algebra operation $f^{\boldsymbol{X^\infty}}$ on X can be induced according to α_i and $f^{\mathbf{X_i}}, i \in I$. Furthermore, the homomorphism from $\boldsymbol{X^\infty}$ to $\prod_{i \in I} \boldsymbol{X_i}$ can be built, i.e., α satisfying $\alpha(a)(i) = \alpha_i(a), \forall a \in \boldsymbol{X^\infty}$.

Remark 3.4. The condition (2) is also called compatibility property. Only when a class of clustering results satisfy compatibility property, algebra operation can be defined on X^∞.

Remark 3.5. The kernel of homomorphism $ker(\alpha) = \bigcap_{i \in I} ker(\alpha_i)$ can build a new congruence on $\boldsymbol{X^\infty}$, in which elements in the same equivalence class restricted in data set X are also partitioned in the same cluster by every single algorithm. Particularly, when α is injective and homomorphic, we denote $ker(\alpha) = \Delta$, which leads to a trivial case.

Example 3.1. *Here, we illustrate how the compatibility property can be satisfied. In Fig.3, two clustering results are displayed. In the first one, data set is partitioned into two regions, i.e., inner layer and outer layer, which can be obtained by some density based algorithm. In the second one, data set is partitioned into four regions, which are indicated by the sectors. This result can be gained by Euclidean distance based algorithm, for example K-means. Consequently, two-tuples are used to mark the cluster labels of all the separated areas in two results respectively. For example,* $(1, 2)$ *represents that the corresponding area is in cluster 1 and cluster 2 according to two clustering results respectively. In this way, 8 different regions are marked.*

Fig. 2. Compatibility Property

In Fig.3, because every cluster obtained by one algorithm has a nonempty intersection with all the cluster obtained by the other algorithm, two clustering results satisfy the compatibility property naturally. In fact, if the congruences corresponding to two results are denoted by θ_1 *and* θ_2*, we have* $\theta_1 \bigcup \theta_2 = \nabla$*, which means all the points are clustered in the same cluster. Additionally, as we have pointed out,* $\theta_1 \bigcap \theta_2$ *will induce a new congruence, containing 8 distinct clusters. Especially, when there is only one point in every single region, we have* $\theta_1 \bigcap \theta_2 = \Delta$*, i.e., different points in different regions. In this special case, we call* θ_1 *and* θ_2 *a pair of factor congruences on* $\boldsymbol{X^\infty}$.

Remark 3.6. Now, we would like to discuss some properties on this algebra structure. Denote $\alpha|_X$ as α restricting in X. First of all, the homomorphism of every single α_i guarantee the homomorphism of α, which also indicates the homomorphism of $\alpha|_X$. Secondly, for an infinite index of I and a finite data set X, $\alpha|_X$ cannot be a mapping from X onto $\prod_{i \in I} X_i$. In this case, we can only discuss when $\alpha|_X$ is an embedding. As is shown above, $\alpha|_X$ is an embedding iff α is an embedding iff $ker(\alpha) = \bigcap_{i \in I} ker(\alpha_i) = \Delta$.

As is discussed above, here, we define the ability of a class of clustering results in separating different data points.

Definition 3.5. Mapping $\alpha_i : \boldsymbol{X^\infty} \to \boldsymbol{X_i}, i \in I$ separate points iff for any arbitrary pair $x_1, x_2 \in X^\infty$, exists α_i, $\alpha_i(a_1) \neq \alpha_i(a_2)$ holds.

Remark 3.7. It can be proved that $\alpha_i : \boldsymbol{X^\infty} \to \boldsymbol{X_i}, i \in I$ separate points iff $\bigcap\limits_{i \in I} ker\alpha_i = \Delta$ iff α is injective.

3.3. The relation between clustering results and congruences on algebra $\boldsymbol{X}$

For data space constructed algebra $\boldsymbol{X^\infty}$, we point out when a class of clustering results, which are denoted by **Clu** $\boldsymbol{X}$, satisfying compatibility property, they can be corresponded to a subset of all the congruences on $\boldsymbol{X^\infty}$, denoted by **Con** $\boldsymbol{X^\infty}$, by an injective mapping.

Here, any element in **Clu** $\boldsymbol{X}$ can be represented by a homomorphism mapping α_i or can be viewed as a quotient space $(\boldsymbol{X^\infty}/ker(\alpha_i)) \bigcap P(X)$. For arbitrary $\theta_i \in$ **Con** $\boldsymbol{X^\infty}$ and $\alpha_i \in$ **Clu** $\boldsymbol{X}$, there is a partition (equivalence relation) of data set X can be induced. Next, $\forall \{x_1, x_2, \cdots, x_n\} \subseteq \boldsymbol{X^\infty}$, define:

$$f^{\boldsymbol{X^\infty}}(a_1, a_2, \cdots, a_n) \in \bigcap_{i \in I} \{s | \alpha_i s = f^{\boldsymbol{X_i}}(\alpha_i a_1, \alpha_i a_2, \cdots, \alpha_i a_n)\}.$$

Obviously, $\forall i \in I, ker(\alpha_i) \in$ **Con** $\boldsymbol{X^\infty}$. Then, we have the following theorem.

Theorem 3.1. *For a class of clustering results on data set X, if they satisfy compatibility property, we can define algebra operation on X^∞. There is a bijective between* **Clu** $\boldsymbol{X}$ *and a subset of* **Con** $\boldsymbol{X^\infty}$*. The intersection of all the kernels under homomorphism with respect to every single clustering result products new congruence on* $\boldsymbol{X^\infty}$.

Remark 3.8. According to this theorem, we can construct the one-to-one correspondence between a class of clustering results and a subset of **Con** $\boldsymbol{X^\infty}$. Then, according to disjunction and conjunction operations on **Con** $\boldsymbol{X^\infty}$, we can study clustering result as congruence. Particularly, every clustering result can be viewed as compactly generated element.

4. Latticized topology for clustering generated lattice

In this section, latticized topology will be employed to study the lattice corresponding to clustering result.

4.1. Latticized topology in congruence lattice

Here, we show the latticized topology in congruence lattice. Firstly, according to *Theorem* 3.1, a correspondence can be built between clustering result in the data set and congruence relation on data space. In this case, we can study the topological property of clustering result by considering topological basis on congruence lattice of the algebra constructed by data space. Now, $\boldsymbol{A}$ is denoted an arbitrary algebra. According to algebraic closure operator Θ, all the congruence relations on $\boldsymbol{A}$ can be viewed as a closed set with respect to Θ. Every $\theta \in$ **Con** $\boldsymbol{A}$ is a closure of some subset of $A \times A$ with respect to Θ. As a result, we consider the topological structure of lattice $(\Theta(P(A \times A)), \vee, \wedge)$. Denote $\Theta(P(A \times A))$ as L. $\tau_*(L)$ is the ideal of L, which is dual to filter $\tau^*(L)$ of L. As is discussed in *Remark* 2.5, T_{01} is a topological basis of L.

Notice that $T_{01} = \{[\gamma) | \gamma \in \Gamma_d\} \cup \emptyset \subseteq P(L)$. For every lower uniformly neighbor elements $\theta \in$ **Con** $\boldsymbol{A}$ with respect to n_*, we can define a new congruence on $\boldsymbol{A}/\theta$, which can be induced by an element ϕ in $[\theta, \nabla]$ as follows:

$$\phi/\theta = \{\langle a/\theta, b/\theta \rangle \in (\boldsymbol{A}/\theta)^2 | \langle a, b \rangle \in \phi\}.$$

Lemma 4.1. *For any pair $\phi, \theta \in$* **Con** $\boldsymbol{A}$*, if $\theta \leq \phi$, then ϕ/θ is a congruence relation on $\boldsymbol{A}/\theta$.*

4.2. Hierarchial structure of congruence relations

Lemma 4.2. *For any pair $\phi, \theta \in$ **Con** $\boldsymbol{A}$, if $\theta \leq \phi$, then $(\boldsymbol{A}/\theta)/(\phi/\theta) \cong \boldsymbol{A}/\phi$.*

Remark 4.1. The lemmas above is easy to prove. According to *Lemma* 4.2, a hierarchial structure of clustering can be abstracted. Suppose θ_1 and θ_2 are two clustering results which satisfy compatibility property. If $\theta_1 \vee \theta_2 \neq \nabla$, then $\theta_1 \vee \theta_2 \in \mathbf{Con}\,\boldsymbol{A}$. Because $\theta_1, \theta_2 \leq \theta_1 \vee \theta_2$ and $(\boldsymbol{A}/\theta_i)/(\theta_1 \vee \theta_2/\theta_i) \cong \boldsymbol{A}/(\theta_1 \vee \theta_2), i = 1, 2$, the clusters in $\theta_1 \vee \theta_2$ can be viewed as grouping on clusters in $\theta_i, i = 1, 2$.

More generally, we have the following theorem:

Theorem 4.1. *For arbitrary element θ in congruence lattice* **Con** $\boldsymbol{A}$*, sublattice $[\theta, \nabla]$ is isomorphic to* **Con** $\boldsymbol{A}/\theta$.

In Fig.3, we indicate the total order hierarchial structure of congruence relations.

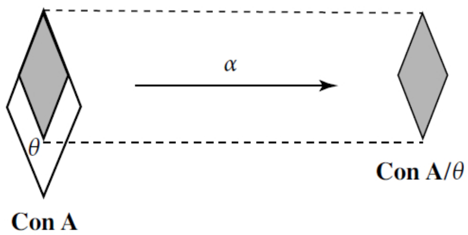

Fig. 3. Hierarchial Structure of Congruence

Remark 4.2. This theorem shows the rationality of clustering on equivalence classes, which is also the thought of granular computing.

5. Conclusion

In this paper, we show the algebra structure of general hard clustering result. From single result to a class of results, homomorphic relation is built between data space and center points space, which can generate congruence of the algebra of data space. Furthermore, by virtue of fundamental homomorphism theorem, we describe single and a class of clustering results under algebra structure. Also, we prove the bijective correspondence between clustering results and the subset of congruences on data space. As a result, the property of algebraic lattice can be transfer onto clustering results, which will be one part of future work. Besides, we build topological basis on lattice to indicate the topological structure of clustering result. In the continued work, specific topology properties will be equipped in latticized topology based clustering analysis.

Acknowledgement

This work has been partially supported by the following Grants: No. 61472390, No. 11271361, Key Project (No. 71331005) and Major International Joint Research Project (No. 71110107026) from the National Natural Science Foundation of China.

References

[1] J.W. Han and M. Kamber. Data Mining Concepts and Techniques. Morgan Kaufmann publishers, USA, 2001.
[2] S. Theodoridis and K. Koutroumbas. Pattern Recognition (4th ed.). Boston: Academic Press, 2008.
[3] A.K. Jain, M.N. Murty, and P.J. Flynn. Data Clustering: A Review. ACM Computing Surveys, 31(3): 264-323, 1999.
[4] J.B. MacQueen. Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1: 281-297, 1967.
[5] E.R. Ruspini. A new Approach to Clustering. Information and Control, 15(1): 22-32, 1969.
[6] S. Burris and H.P. Sankappanavar. A Course in Universal Algebra. New York: Springer, 1981.
[7] P.Z. Wang. Neighbor Element Structure of Latticized Structure and Convergence Ralation (in Chinese). Journal of Beijing Normal University, 2, 1984.
[8] G. Birkhoff. Lattice Theory (3rd ed.). Colloq. Publ., vol. 25, Amer. Math. Soc., Providence, 1967.
[9] R. Balbes and P. Dwinger. Distributive Lattices. Univ. of Missouri Press, Columbia, 1974.