About Some Clustering Algorithms in Evidence Theory

Alexander Lepskiy

National Research University – Higher School of Economics,
Moscow, Russia

International conference "Data Analysis, Optimization and their Applications",
on the occasion of Prof. **Boris Mirkin's** 80^{th} Birthday

January 30-31, 2023. Dolgoprudny, Moscow Institute of Physics and Technology

Research Motivation

The problem of clustering the so-called evidence bodies in the framework of the Dempster–Shafer theory is considered. The body of evidence is a pair $F = (\mathcal{A}, m)$, where

- $\mathcal{A}$ is the set of non-empty subsets (focal elements) of some base set X,
- m is a non-negative function of sets (mass function) defined on the set of all subsets of the base set.

The focal element $A \in \mathcal{A}$ describes the membership set of the true alternative $x \in A$ (for example, the air temperature forecast), and the mass $m(A)$ of this focal element A specifies the degree of belief that $x \in A$.

The body of evidence can be complex. For example, it may consist of many focal elements with a complex intersection structure.

We have the following problems:

- complex belief structures are difficult to interpret;
- high computational complexity of performing operations on complex belief structures.

Therefore, the following problems are relevant:

- analysis of the structure of the set of focal elements $\mathcal{A}$ of the body of evidence $F = (\mathcal{A}, m)$;
- finding an enlarged (simplified) structure of the set of focal elements $\widetilde{\mathcal{A}}$;
- redistribution of masses of focal elements of the set $\mathcal{A}$ to focal elements from $\widetilde{\mathcal{A}}$. As a result, we obtain a new mass function $\widetilde{m}$, etc.

Outline of Presentation

- Background of the Belief Function Theory;
- Basic Approaches for Clustering Body of Evidence:
- Hierarchical Inner and Outer Clustering;
- Clustering Based on Conflict Optimization;
  - Clustering Based on Conflict Density;
  - Redistribution of Focal Elements;
  - The k-means Algorithm for the Body of Evidence;
- Evaluation of the Internal Conflict;
- Summary and Conclusion.

Background of the Belief Function Theory

Dempster A.P. Upper and lower probabilities induced by multivalued mapping. Ann. Math. Statist. 38, 325–339 (1967)

Shafer G. A mathematical theory of evidence. Princeton Univ. Press (1976)

Let

- X be some set;
- $\mathcal{A} \subseteq 2^X$ be some finite subset of focal elements;
- $m : 2^X \to [0,1]$, $\sum_{A \in \mathcal{A}} m(A) = 1$ be some mass function, $m(A) > 0 \; \forall \; A \in \mathcal{A}$;
- the pair $F = (\mathcal{A}, m)$ is called a body of evidence;
- categorical evidence $F_A = (A, 1)$;
- if $F = (\mathcal{A}, m)$, then $F = \sum_{A \in \mathcal{A}} m(A) F_A$;
- in particular, simple evidence $F_A^\alpha = \alpha F_A + (1 - \alpha) F_X$, $\alpha \in [0, 1]$.

There is a one-to-one correspondence between the body of evidence $F = (\mathcal{A}, m)$ and the **belief function**

$$Bel(A) = \sum\nolimits_{B \subseteq A} m(B)$$

or the **plausibility function**

$$Pl(A) = \sum\nolimits_{B \cap A \neq \emptyset} m(B),$$

which can be considered as lower and upper bounds for the probability $P(A)$, respectively.

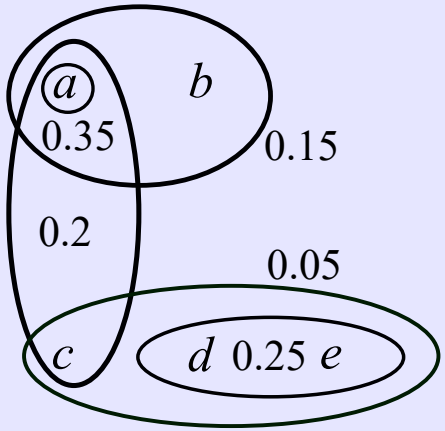
The body of evidence $F = (\mathcal{A}, m)$ on X can be represented as a weighted hypergraph with a set of vertices X, a set of hyperedges $\mathcal{A}$ and their weights $m(A)$, $A \in \mathcal{A}$.

Example

Let we have $X = \{a, b, c, d, e\}$ and the body of evidence

$$F = 0.35F_{\{a\}} + 0.15F_{\{a,b\}} + 0.2F_{\{a,c\}} + 0.25F_{\{d,e\}} + 0.05F_{\{c,d,e\}}$$

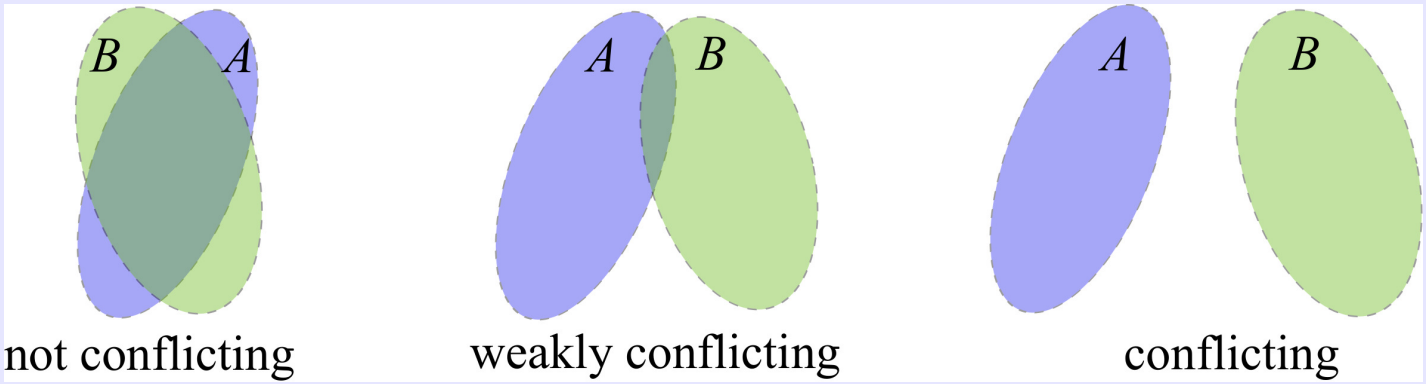
is given on X, i.e. $\mathcal{A} = \{\{a\}, \{a, b\}, \{a, c\}, \{d, e\}, \{c, d, e\}\}$. The hypergraph of the evidence body F is shown in Fig.



If two sources of information are represented by the bodies of evidence $F_1 = (\mathcal{A}_1, m_1)$ and $F_2 = (\mathcal{A}_2, m_2)$ on X, then the degree of conflict (contradiction) between these sources can be assessed using some functional (measure of external conflict) $Con : \mathcal{F}(X) \times \mathcal{F}(X) \to [0, 1]$, which takes on greater values the more pairs of non-overlapping (or 'weakly over-lapping') focal elements of two evidence bodies with large masses exist. The classical measure of external conflict is

$$Con(F_1, F_2) = \sum_{A \cap B = \emptyset} m_1(A) m_2(B),$$

which we will use below.



Basic Approaches for Clustering Body of Evidence

The clustering of the body of evidence $F = (\mathcal{A}, m)$ is primarily related to the clustering of the set of its focal elements $\mathcal{A}$. There are two formulations of the problem of clustering a set of focal elements.

1. It is required to find such a subset of $\mathcal{A}' \subseteq 2^X$ that would be 'close' to $\mathcal{A}$ in some sense, but $|\mathcal{A}'| \ll |\mathcal{A}|$. The new mass function $m'(A)$, is found either by a local redistribution of the masses $m(B)$ of the sets B involved in the formation of a new focal element $A \in \mathcal{A}'$, or by a global redistribution that minimizes the discrepancy functional between $F = (\mathcal{A}, m)$ and $F' = (\mathcal{A}', m')$.
2. It is required to find such a partition (or cover) of the set $\mathcal{A}$ of focal elements into subsets (clusters) $\{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$ that would correspond in some sense to the structure of the set $\mathcal{A}$.

The first type of clustering is used to reduce the computational complexity of algorithms for processing evidence bodies or solving other approximation problems. The second type of clustering is used to identify the structure of a set of focal elements, to estimate the degree of heterogeneity, inconsistency, etc.

Next, we consider some implementations of clustering of these two types, namely:

1. hierarchical clustering;
2. clustering based on conflict optimization.

Hierarchical Inner and Outer Clustering

There are several approaches to hierarchical clustering of evidence bodies. For example, the following algorithm is one of the popular [Denœux T. Inner and outer approximation of belief structures using a hierarchical clustering approach. Int. J. of Uncert., Fuzz. and Know.-Based Syst. 9(4), 437–460 (2001)].

Two clusterings are the result of this algorithm. One of them is internal in the form of evidence body $F^- = (\mathcal{A}^-, m^-)$, the other is external in the form $F^+ = (\mathcal{A}^+, m^+)$, where

$$B = \bigcap A \in \mathcal{A}^-, \;\; C = \bigcup A \in \mathcal{A}^+, \;\; m^{\pm}(B) = \sum\nolimits_A m(A).$$

A pair (A, B) of focal elements is chosen for union/intersection, which delivers the minimum increment of the measure of imprecision

$$f(F) = \sum\nolimits_{A \in \mathcal{A}} m(A)\,|A|.$$

The increments of this measure at the union/intersection of two sets and will be equal

$$\delta_{\cup}(C,D) = (m(C)+m(D))\,|C\cup D| - m(C)\,|C| - m(D)\,|D|$$

and

$$\delta_{\cap}(C,D) = m(C)\,|C| + m(D)\,|D| - (m(C)+m(D))\,|C\cap D|\,,$$

respectively. Therefore, the algorithm finds at each step a pair

$$(A^-,B^-) = \arg\min_{C\neq D}\delta_{\cap}(C,D)$$

for intersection and a pair

$$(A^+,B^+) = \arg\min_{C\neq D}\delta_{\cup}(C,D)$$

for union.

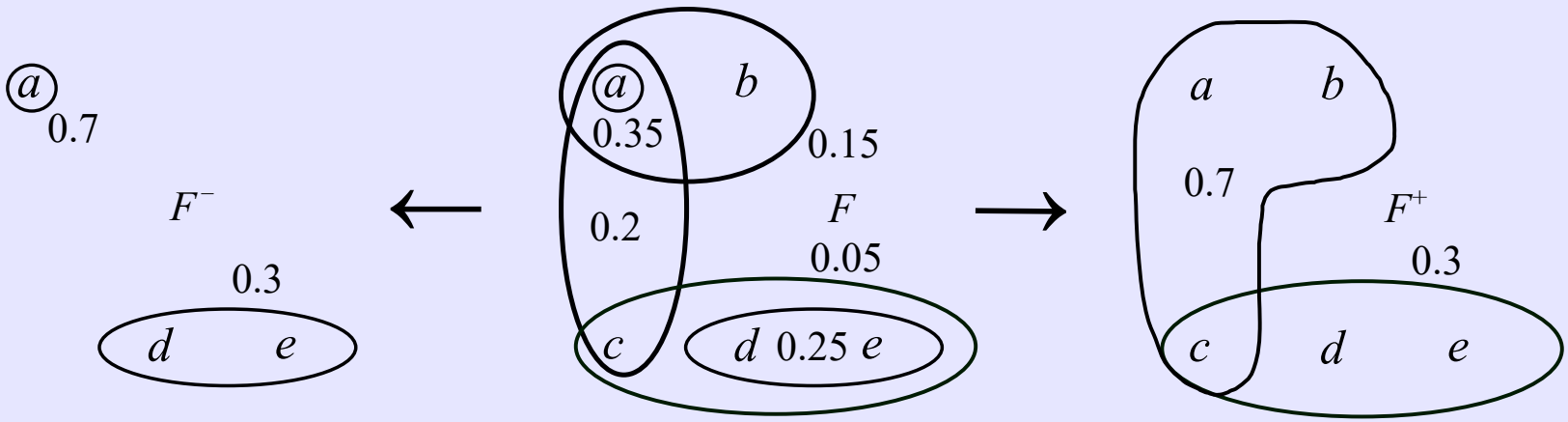
Example

We have the following transformations of sets of focal elements:

$$\mathcal{A}=\{\{a\},\{a,b\},\{a,c\},\{\boldsymbol{d},\boldsymbol{e}\},\{\boldsymbol{c},\boldsymbol{d},\boldsymbol{e}\}\}\to\{\{\boldsymbol{a}\},\{\boldsymbol{a},\boldsymbol{b}\},\{a,c\},\{c,d,e\}\}\to$$
$$\to\{\{\boldsymbol{a},\boldsymbol{b}\},\{\boldsymbol{a},\boldsymbol{c}\},\{c,d,e\}\}\to\{\{a,b,c\},\{c,d,e\}\}=\mathcal{A}^{+},$$
$$\mathcal{A}=\{\{a\},\{a,b\},\{a,c\},\{\boldsymbol{d},\boldsymbol{e}\},\{\boldsymbol{c},\boldsymbol{d},\boldsymbol{e}\}\}\to\{\{\boldsymbol{a}\},\{a,b\},\{\boldsymbol{a},\boldsymbol{c}\},\{d,e\}\}\to$$
$$\to\{\{\boldsymbol{a}\},\{\boldsymbol{a},\boldsymbol{b}\},\{d,e\}\}\to\{\{a\},\{d,e\}\}=\mathcal{A}^{-}.$$

We obtain the outer and inner approximations

$$F^{+}=0.7F_{\{a,b,c\}}+0.3F_{\{c,d,e\}}\ \text{ and }\ F^{-}=0.7F_{\{a\}}+0.3F_{\{d,e\}}.$$



Clustering Based on Conflict Optimization

This approach is based on the assumption that the heterogeneity of the set of focal elements is due to the fact that this information could be obtained from conflicting sources.

Therefore, the conflict between focal elements of one cluster should be small, and the conflict between focal elements of different clusters should be large. Non-overlapping focal elements are called conflicting.

Let's consider two clustering methods related to conflict optimization:

1. a method of selecting a small set $\mathcal{A}' \subseteq \mathcal{A}$ of the most conflicting focal elements (cluster centers) and (possibly) redistributing the remaining focal elements among these centers, maximizing conflict between clusters;
2. a method for finding a partition (coverage) of a set of focal elements that minimizes the average intracluster conflict (evidential k-means).

Conflict Density

The first method uses the concept of conflict density and was proposed in [Bronevich A., Lepskiy A. Measures of conflict, basic axioms and their application to the clusterization of a body of evidence. Fuzzy Sets and Syst. 446, 812–832 (2022)].

A function $\psi_F : 2^X \to [0,1]$ is called the conflict density distribution of the evidence body $F = (\mathcal{A}, m)$ if it satisfies the conditions:

1. $\psi_F(A) = 0$ if $B \cap A \neq \emptyset \ \forall B \in \mathcal{A}$;
2. $\psi_F(A) = 1$ if $B \cap A = \emptyset \ \forall B \in \mathcal{A}$;
3. $\psi_{\alpha F_1 + \beta F_2} = \alpha \psi_{F_1} + \beta \psi_{F_2} \ \forall F_1, F_2 \in \mathcal{F}(X)$, $\alpha + \beta = 1$, $\alpha \geq 0$, $\beta \geq 0$.

It can be shown that $\psi_F(A) = \sum_{B:A\cap B=\emptyset} m(B) = 1 - Pl(A)$. In addition, we are primarily interested in elements with a large mass. Therefore, we will use the function

$$\varphi_F(A) = m(A)\psi_F(A), \quad A \in \mathcal{A}.$$

The distance $d(A, B)$ between focal elements is another characteristic that we will take into account. This distance should not be too small.

Algorithm 2.
Input data: $F = (\mathcal{A}, m)$, the minimum value $h_1 > 0$ of $\varphi_F(A)$ $\forall A \in \mathcal{A}'$; the minimum distance $h_2 > 0$ between focal elements from $\mathcal{A}'$.
Output data: the body of evidence $\mathcal{A}' \subseteq \mathcal{A}$.
1. Let the set $\mathcal{A}$ be ordered in descending order of the function φ_F: $\varphi_F(A_1) \geq \varphi_F(A_2) \geq \ldots \geq \varphi_F(A_k)$. Put $\mathcal{A}' = \{A_1\}$, $s := 2$.
2. If $\varphi_F(A_s) \leq h_1$, then the end. Otherwise, go to step 3.
3. If $\min\limits_{A\in\mathcal{A}'} d(A, A_s) > h_2$, then $\mathcal{A}' := \mathcal{A}' \cup \{A_s\}$, $s := s + 1$, go to step 2.

Examples of metrics between focal elements:

1. cardinality (measure) of the symmetric difference of sets $d_S(A,B) = |A \Delta B|$;
2. Hausdorff metric $d_H(A,B) = \max\left\{\sup\limits_{a\in A}\inf\limits_{b\in B}\rho(a,b), \sup\limits_{b\in B}\inf\limits_{a\in A}\rho(a,b)\right\}$, if X is a metric space;
3. $d_J(A,B) = \rho_J(F_A^{m(A)}, F_B^{m(B)})$, where $\rho_J(F_1,F_2) = \sqrt{\frac{1}{2}\sum_{A,B} s_{A,B}(m_1(A)-m_2(A))(m_1(B)-m_2(B))}$, $s_{A,B} = \frac{|A\cap B|}{|A\cup B|}$ is the Jaccard index.

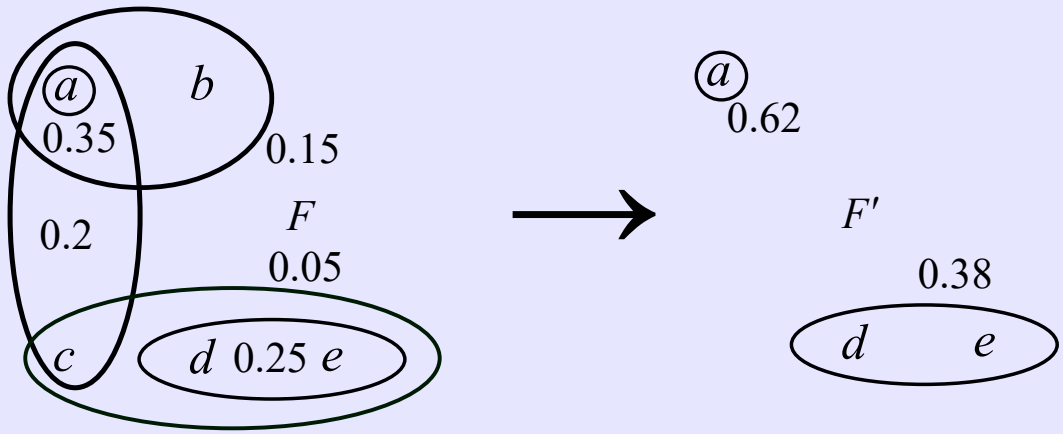
This algorithm can be considered as an evidential analogue of the 'point' the DBSCAN algorithm (Density Based Spatial Clustering of Applications with Noise [Ester M., Kriegel H-P., Sander J., Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proc. of the 2^{nd} Int. Conf. on Knowl. Discov. and Data Mining, 226–231. AAAI Press (1996)]).

Example

Algorithm will give the new set of focal elements $\mathcal{A}' = \{\{d, e\}, \{a\}\}$ for the evidence body from Example using the metric d_J, $h_1 = 0.1$, $h_2 = 0.2$. Then the general form of the body of evidence with the set of focal elements $\mathcal{A}'$ will be as follows

$$F'(x) = xF_{\{a\}} + (1 - x)F_{\{d,e\}}, \quad x \in [0, 1].$$

The masses of the body of evidence F' can be found from the condition of minimizing the distance $\rho_J(F, F'(x)) \to \min$. Finally we will get $F' = 0.62F_{\{a\}} + 0.38F_{\{d,e\}}$.



Redistribution of Focal Elements Among New Clusters

If we have a body of evidence $F = (\mathcal{A}, m)$, then we need to find such a partition (or cover) of the set of focal elements $\mathcal{A}$ into subsets (clusters) $\mathcal{C} = \{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$ in order to maximize the external conflict between evidence clusters:

$$Con(F(\mathcal{A}_1), \ldots, F(\mathcal{A}_l)) \to \max\,.$$

Here $F(\mathcal{A}_i)$ means the redistribution of the masses of focal elements from the set $\mathcal{A}$ to $\mathcal{A}_i \subseteq \mathcal{A}$, which will be carried out according to the rule (extension procedure)

$$F(\mathcal{A}_i) = (\mathcal{A}_i, m_i): \quad m_i(A) = m(A) \;\; \forall A \in \mathcal{A}_i,$$

$$m_i(X) = 1 - \sum\nolimits_{A \in \mathcal{A}_i} m(A).$$

Algorithm

Input data: $F = (\mathcal{A}, m)$, a selected set $\mathcal{A}' = \{A_1, \ldots, A_l\} \subseteq \mathcal{A}$.
Output data: partition (cover) $\mathcal{C} = \{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$ of the set $\mathcal{A}$.
1. Let $\mathcal{A}_i^{(0)} = \{A_i\}$, $i = 1, \ldots, l$.
2. The focal element $B \in \mathcal{A} \setminus \left\{\mathcal{A}_1^{(0)}, \ldots, \mathcal{A}_l^{(0)}\right\}$ will be assigned to that cluster $\mathcal{A}_i^{(0)}$ for which the maximum conflict measure is reached:

$$\mathcal{A}_i^{(0)} = \underset{j: B \in \mathcal{A}_j^{(0)}}{\arg\max} Con\left(F\left(\mathcal{A}_1^{(0)}\right), \ldots, F\left(\mathcal{A}_j^{(0)} \cup \{B\}\right), \ldots, F\left(\mathcal{A}_l^{(0)}\right)\right).$$

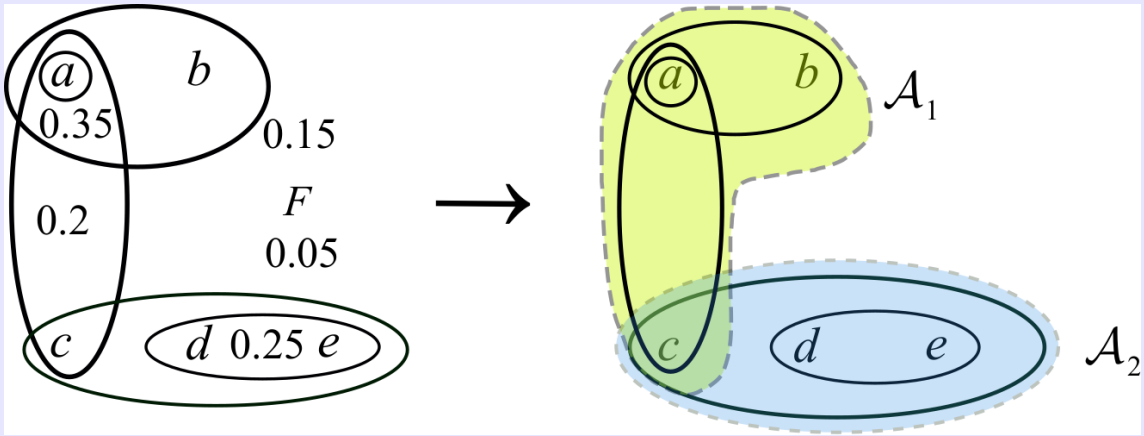
If equal maximum conflict values are obtained when assigning the element B to several clusters $\mathcal{A}_j^{(0)}$, $j \in J$, then this element B is included in all these clusters, and the mass value $m(B)$ is evenly distributed over the updated clusters, i.e. element B will be included in each cluster $\mathcal{A}_j^{(0)}$, $j \in J$ with weight $m(B)/|J|$.

Example

Let's redistribute the remaining focal elements $\mathcal{A}\backslash\mathcal{A}' = \{\{a,b\},\{a,c\},\{c,d,e\}\}$ for the body of evidence from Example and the the selected set of focal elements $\mathcal{A}' = \{\{d,e\},\{a\}\}$.

We get a partition $\mathcal{C} = \{\mathcal{A}_1, \mathcal{A}_2\}$, where

$$\mathcal{A}_1 = \{\{a\},\{a,b\},\{a,c\}\}, \quad \mathcal{A}_2 = \{\{d,e\},\{c,d,e\}\}.$$



The Evidential k-means Algorithm

Suppose we have a body of evidence $F = (\mathcal{A}, m)$. It is required to find a partition (or cover) of the set $\mathcal{A}$ into subsets $\{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$ such that minimize the overall conflict between the centers of clusters C_i and the bodies of evidence formed by the focal elements of these clusters

$$\Phi = \sum_{i=1}^{l} \sum_{B \in \mathcal{A}_i} Con\left(F(\{B\}), C_i\right) \to \min,$$

where $F\left(\{B\}\right) = m(B)F_B + (1 - m(B))F_X$.

By the center of the i-th cluster $\mathcal{A}_i$, we mean some body of evidence C_i constructed from the pair $(\mathcal{A}_i, m_i)$, where m_i is the restriction of the mass function to $\mathcal{A}_i \subseteq \mathcal{A}$, $i = 1, \ldots, l$. Let the center C_i has the form

$$C_i = \sum_{A \in \mathcal{A}_i} \alpha_i(A) F_A, \tag{1}$$

where $\alpha_i = (\alpha_i(A))_{A \in \mathcal{A}_i} \in S_{|\mathcal{A}_i|}$, $S_k = \left\{(t_1, \ldots, t_k) : t_i \geq 0, \sum_{i=1}^{k} t_i = 1\right\}$ is an k-dimensional simplex.

Theorem

Let $Pl_{\mathcal{A}_i}(A) = \sum_{\substack{B \in \mathcal{A}_i: \\ A \cap B \neq \emptyset}} m(B)$. *Then the minimum of the functional* Φ *for a fixed cover* $\mathcal{C} = \{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$ *will be achieved at*

$$\alpha_i = (\alpha_i(A))_{A \in \overline{\mathcal{A}}_i} \in S_{|\overline{\mathcal{A}}_i|}, \quad i = 1, \ldots, l, \tag{2}$$

where $\overline{\mathcal{A}}_i = \left\{ A \in \mathcal{A}_i : A = \arg\max_{A \in \mathcal{A}_i} Pl_{\mathcal{A}_i}(A) \right\}$.

The Algorithm

1. Let's choose the number of clusters l. Let's assign some evidence bodies as initial cluster centers $C_i^{(0)}$, $i = 1, \ldots, l$. We fix the threshold of maximum conflict within clusters $Con_{\max} \in [0, 1]$. Put $s = 0$.

2. We redistribute focal elements among clusters according to the principle of minimizing the conflict between evidence clusters and cluster centers. The focal element $B \in \mathcal{A}$ is assigned to the cluster

$$\mathcal{A}_i^{(s)} = \arg\min_j Con\left(F(\{B\}), C_j^{(s)}\right)$$

and $\min_i Con\left(F(\{B\}), C_i^{(s)}\right) \leq Con_{\max}$. If $\min_i Con\left(F(\{B\}), C_i^{(s)}\right) > Con_{\max}$, then the focal element B is assigned as the center of the new cluster. We get clusters $\mathcal{A}_i^{(s)}$, $i = 1, \ldots, l$.

3. Let us calculate new cluster centers using the formulas (1), (2). We increase the counter $s \leftarrow s + 1$.

4. Steps 2 and 3 are repeated until the clusters (or their centers) stabilize.

Proposition.

Algorithm converges in a finite number of steps.

Cluster centers may depend on parameters $\alpha = (\alpha(A))_{A \in \overline{\mathcal{A}}_i} \in S_{|\overline{\mathcal{A}_i}|}$.
In this case, it is necessary to use additional procedures for choosing parameters. The selection criteria can be considered, for example:

1. coverage minimization, i. e., we choose the parameters so that the coverage $\mathcal{C} = \{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$ is 'closer' to the partition. For example, $\sum_{i=1}^{l} |\mathcal{A}_i| \to \min$.
2. minimizing the uncertainty of evidence-centers of clusters C_i, $i = 1, \ldots, l$. For example, this can be done using the generalized Hartley measure $H(C_i) = \sum_{A \in \overline{\mathcal{A}_i}} \alpha_i(A) \ln |A| \to \min$.
3. minimizing the distance between cluster centers and the original evidence body: $d(C_i, F) \to \min$, $i = 1, \ldots, l$;
4. maximizing distance between cluster centers $d(C_i, C_j) \to \max$ or maximizing conflict $Con(C_i, C_j) \to \max$, $i, j = 1, \ldots, l$ $(i \neq j)$ etc.

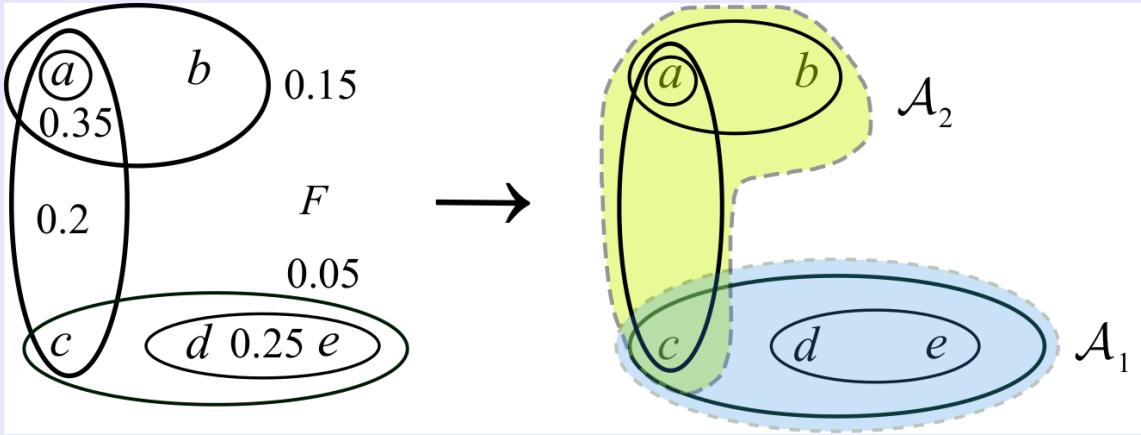
Example

Let's apply this algorithm for clustering into two clusters the body of evidence

$$F = 0.35F_{\{a\}} + 0.15F_{\{a,b\}} + 0.2F_{\{a,c\}} + 0.25F_{\{d,e\}} + 0.05F_{\{c,d,e\}}$$

on $X = \{a, b, c, d, e\}$. As a result, we get clusters

$$\mathcal{A}_1^{(1)} = \{\{d, e\}, \{c, d, e\}\}, \quad \mathcal{A}_2^{(1)} = \{\{a\}, \{a, b\}, \{a, c\}\}.$$



Evaluation of the Internal Conflict Based on Clustering

Clustering a body of evidence $F = (\mathcal{A}, m)$ can be used to evaluate its internal conflict. If $\mathcal{C} = \{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$ is a cover (or partition) of the set of focal elements $\mathcal{A}$, then the internal conflict can be estimated by the formula

$$Con_{in}(F) = Con(F(\mathcal{A}_1), \ldots, F(\mathcal{A}_l)).$$

For example, the measure of internal conflict of the body of evidence from Example will be equal to

$$Con_{in}(F) = Con(F(\{d, e\}, \{c, d, e\}), F(\{a\}, \{a, b\}, \{a, c\})) = 0.2.$$

Summary and Conclusion

- the following classes of algorithms of evidence body clustering are considered:
  a) hierarchical clustering algorithms;
  b) clustering algorithms based on the density function;
  c) clustering algorithms based on conflict optimization.;
- many of the considered algorithms are analogues of the corresponding algorithms for "point" data;
- the dual frequency-multiple nature of the bodies of evidence imposes peculiar restrictions, the need to use "one's own" measures of proximity (for example, based on measures of conflict), etc;
- It shows how clustering can be used to evaluate the internal conflict of a body of evidence;
- all these features leave a lot of room for creativity in the development of algorithms for clustering bodies of evidence.

References


Bronevich, A., Lepskiy, A.: Measures of conflict, basic axioms and their application to the clusterization of a body of evidence. Fuzzy Sets and Systems 446, 812–832 (2022)


Denœux, T.: Inner and outer approximation of belief structures using a hierarchical clustering approach. Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems 9(4), 437–460 (2001)


Harmanec, D.: Faithful approximations of belief functions. In: Laskey, K. B. and Prade, H., (eds) Uncertainty in Artificial Intelligence 15 (UAI99), Stockholm, Sweden, (1999)


Lepskiy, A.: Analysis of Information Inconsistency in Belief Function Theory. Part I: External Conflict. Control Sciences 5, 2–16 (2021)


Lepskiy, A.: Analysis of Information Inconsistency in Belief Function Theory. Part II: Internal Conflict. Control Sciences 6, 2–12 (2021)


Petit-Renaud, S., Denœux, T.: Handling different forms of uncertainty in regression analysis: a fuzzy belief structure approach. In: Hunter, A. and Pearsons, S., (eds) Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'99), pp. 340–351. Springer Verlag (1999)

Thanks for you attention

Happy Birthday Boris Grigorievich

alex.lepskiy@gmail.com
alepskiy@hse.ru
https://www.hse.ru/en/org/persons/10586209