Clustering a Body of Evidence Based on Conflict Measures

Andrey G. Bronevich,
Alexander E. Lepskiy

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

The 11th Conference of the European Society for Fuzzy Logic and Technology,
9-13 September 2019, Prague, Czech Republic

The outline of the paper

1. In the paper, we propose the clustering based on conflict measures.
2. Clustering is understood as a decomposition of the body of evidence on pieces of evidence such that within pieces of evidence there is no conflict (or small conflict), and the most of conflict should be concentrated among pieces of evidence.
3. For this purpose, we use conflict measures whose computation is based on solving optimization problems. We show that the solution of such problems gives us the required clusterization.
4. Additionally, we give the axiomatic of such conflict measures.
5. We show how to use the evidence clusterization by analyzing political preferences of parties in Germany and their influence on popularity of such parties.

Some notations and definitions

- $X = \{x_1, ..., x_n\}$ is a finite set;
- 2^X is the the powerset of X;
- $Bel : 2^X \to [0, 1]$ is a belief function if there is a set function $m : 2^X \to [0, 1]$ with $m(\emptyset) = 0$ and $\sum_{A \in 2^X} m(A) = 1$ called the basic belief assignment (bba) such that

$$Bel(A) = \sum_{B \subseteq A | B \in 2^X} m(B);$$

- bba m defines also the plausibility function

$$Pl(A) = \sum_{B \cap A \neq \emptyset | B \in 2^X} m(B);$$

- Let Bel be a belief function with the bba m, then $A \in 2^X$ is a focal element for Bel if $m(A) > 0$. The set of all focal elements is called the body of evidence.;

Some notations and definitions

If a belief function has one focal element B, then it is called categorical and denoted by $\eta_{\langle B\rangle}$ and, obviously,

$$\eta_{\langle B\rangle}(A) = \begin{cases} 1, & B \subseteq A, \\ 0, & \text{otherwise.} \end{cases}$$

Let μ, μ_1, μ_2 are set functions on 2^X, then we write:

- $\mu = a\mu_1 + b\mu_2$ for $a, b \in \mathbb{R}$ if $\mu(A) = a\mu_1(A) + b\mu_2(A)$ for all $A \in 2^X$;
- $\mu_1 \le \mu_2$ if $\mu_1(A) \le \mu_2(A)$ for all $A \in 2^X$.

Every belief function can be represented as a convex sum of categorical belief functions

$$Bel = \sum_{B \in 2^X} m(B)\eta_{\langle B\rangle},$$

where m is the bba of Bel.

Some notations and definitions

- A belief function is a probability measure if its body of evidence consists of singletons.
- $M_{bel}(X)$ denotes the set of all belief functions on 2^X.
- $M_{pr}(X)$ denotes the set of all probability measures on 2^X.
  (If we do not identify the reference set X, or it can be identified by the context, then we write simply M_{bel} or M_{pr}.)
- Let $Bel \in M_{bel}(X)$. Then a set of probability measures

$$\mathbf{P}(Bel) = \{P \in M_{pr}(X) | P \geq Bel\},$$

  is called a credal set.

Three possible interpretations of conflict

Let $U_C : M_{bel} \to [0, +\infty)$ be a functional for measuring conflict.
We say that $\mu \in M_{bel}$ describes conflict-free information if $U_C(Bel) = 0$.

1. $U_C(Bel) = 0$ for $Bel \in M_{bel}(X)$ iff $Bel = \eta_{\langle B \rangle}$ for some $B \in 2^X$;
2. $U_C(Bel) = 0$ for $Bel \in M_{bel}(X)$ iff the body of evidence $\mathcal{A}$ of Bel is a chain of sets, i.e. the elements of $\mathcal{A}$ can be indexed such that $\mathcal{A} = \{B_1, ..., B_m\}$ and $B_i \subseteq B_j$, when $i \leq j$;
3. $U_C(Bel) = 0$ for $Bel \in M_{bel}(X)$ with a body of evidence $\mathcal{A} = \{B_1, ..., B_m\}$ iff $\bigcap_{i=1}^{m} B_i \neq \emptyset$.

In our investigation we will assume that the third interpretation is fulfilled.

Mappings of belief functions

Here we introduce the construction on belief functions that further allows us to formulate desirable properties of conflict measures.

- $\varphi : X \to Y$ be a mapping between finite sets X and Y,
- $Bel \in M_{bel}(X)$.
- Bel^{φ} is the belief function on 2^Y defined by

  $$Bel^{\varphi}(B) = Bel(\varphi^{-1}(B)),$$

  where $B \in 2^Y$ and $\varphi^{-1}(B) = \{x \in X | \varphi(x) \in B\}$.
- if m is the bba for Bel, then $m^{\varphi}(B) = m(\varphi^{-1}(B))$ for $B \in 2^Y$.

Axioms for conflict measures on M_{bel}

Axiom 1

$U_C(Bel) = 0$ for $Bel \in M_{bel}(X)$ with a body of evidence $\mathcal{A} = \{B_1, ..., B_m\}$ iff $\bigcap_{i=1}^{m} B_i \neq \emptyset$.

Axiom 2

$U_C(Bel_1) \leq U_C(Bel_2)$ for $Bel_1, Bel_2 \in M_{bel}(X)$ if $Bel_1 \leq Bel_2$.

Axiom 3

Let $\varphi : X \to Y$ be a mapping between finite sets X and Y, then $U_C(Bel^{\varphi}) \leq U_C(Bel)$ for every $Bel \in M_{bel}(X)$. In addition, $U_C(Bel^{\varphi}) = U_C(Bel)$ if φ is an injection.

Axiom 4

Let $Bel = aBel_1 + (1-a)Bel_2$, where $a \in [0,1]$ and $Bel_1, Bel_2 \in M_{bel}(X)$, then $U_C(Bel) \geq aU_C(Bel_1) + (1-a)U_C(Bel_2)$.

The construction of conflict measures

Lemma 1

Let a functional U_C satisfies Axioms 1-4, then for every $Bel \in M_{bel}(X)$, we have $U_C(Bel) \leq U_C(P_n)$, where the measure P_n defines the uniform probability distribution on $X = \{x_1, ..., x_n\}$, i.e. $P_n(\{x\}) = 1/n$ for every $x \in X$.

Theorem 1

Let the functional $\Phi : M_{pr} \to \mathbb{R}$ satisfies Axioms 1-4 on the set M_{pr} of all possible probability measures. Then Φ might be extended to the set M_{bel} of all belief functions by

$$U_C(Bel) = \inf \{\Phi(P) | P \in \mathbf{P}(Bel)\}.$$

Moreover, U_C satisfies Axioms 1-4.

The construction of conflict measures

Let us describe a measure of conflict on M_{pr} using functions defined on $\mathbb{R}^n$. For this purpose, consider an arbitrary $P \in M_{pr}(X)$, where $X = \{x_1, ..., x_n\}$. Obviously, such a P is uniquely defined by the vector $(P(\{x_1\}, ..., P(\{x_n\})$.

Thus, a conflict measure U_C on M_{pr} can be defined by a system of functions $f_n : \Omega_n \to [0, +\infty)$, $n = 1, 2, ...$, such that

1. $\Omega_n = \{(t_1, ..., t_n) \in \mathbb{R}^n | \sum_{i=1}^{n} t_i = 1,\ t_i \geq 0,\ i = 1, ..., n\}$;
2. $U_C(P) = f_n(P(\{x_1\}, ..., P(\{x_n\})$.

The construction of conflict measures

Theorem 2

A system of functions $f_n : \Omega_n \to [0, +\infty)$, $n = 1, 2, ...$, defines a measure U_C of conflict on M_{pr} iff

1. $f_n(t_1, ..., t_n) = 0$ if $t_1 = 1$ and $f_n(t_1, ..., t_n) > 0$ if $t_1 \in (0, 1)$;
2. $f_{n+1}(t_1, ..., t_n, 0) = f_n(t_1, ..., t_n)$ for every $(t_1, ..., t_n) \in \Omega_n$;
3. $f_n(t_{\varphi(1)}, ..., t_{\varphi(n)}) = f_n(t_1, ..., t_n)$ for every one-to-one mapping $\varphi : \{1, ..., n\} \to \{1, ..., n\}$;
4. $f_{n-1}(t_1 + t_2, t_3, ..., t_n) \leq f_n(t_1, t_2, t_3, ..., t_n)$ for every $(t_1, ..., t_n) \in \Omega_n$;
5. f_n is a concave function on Ω_n, i.e.

$$f_n(a\mathbf{t}_1 + (1-a)\mathbf{t}_2) \geq a f_n(\mathbf{t}_1) + (1-a) f_n(\mathbf{t}_2)$$

for every $a \in [0, 1]$ and every $\mathbf{t}_1, \mathbf{t}_2 \in \Omega_n$.

Examples of conflict measures

Example 1.

$f_n(t_1, ..., t_n) = \min\{1 - t_1, ..., 1 - t_n\}.$

If we extend this conflict measure on M_{bel}. Then we get the result

$$U_C(Bel) = 1 - \max_{x \in X} Pl(\{x\}),$$

where Pl is the plausibility function associated with $Bel \in M_{bel}(X)$.

Examples of conflict measures

Proposition 1

Let $g : [0,1] \to [0,+\infty)$ be a concave function with the following properties:

1. $g(0) = g(1) = 0$;
2. $g(t)$ is strictly decreasing at $t = 1$.

Then the system of functions

$$f_n(t_1, ..., t_n) = \sum_{i=1}^{n} g(t_i), \quad (t_1, ..., t_n) \in \Omega_n,$$

obeys the conditions from Theorem 2.

Examples of conflict measures

Example 2.

If $g(t) = -t \ln t$, then $f_n(t_1, ..., t_n) = -\sum_{i=1}^{n} t_i \ln t_i$ defines the Shannon entropy. The corresponding conflict measure on M_{bel} is called the minimal Shannon entropy. g is concave, since $g''(t) = -(1/t) < 0$ for $t \in (0, 1]$.

Example 3.

We can evaluate conflict within a probability measure $P \in M_{pr}$ using Dempster's rule of aggregation.

$$k(P, P) = 1 - \sum_{i=1}^{n} P(\{x_i\})P(\{x_i\}) = \sum_{i=1}^{n} P(\{x_i\})(1 - P(\{x_i\}))$$

In this case, $f_n(t_1, ..., t_n) = \sum_{i=1}^{n} t_i(1 - t_i)$, i.e. $g(t) = t(1 - t)$.

Clustering a body of evidence: the problem statement

Let $Bel \in M_{bel}(X)$. Then the problem is how to disaggregate Bel on parts $Bel_1, ..., Bel_k \in M_{bel}$ such that

$$Bel = \sum_{i=1}^{k} a_i Bel_i,$$

where $\sum_{i=1}^{k} a_i = 1$, $a_i \geq 0$, such that

- the conflict within belief functions Bel_i should be minimal,
- the external conflict among them should be maximal.

The internal conflict can be evaluated by

$$\sum_{i=1}^{k} a_i U_C(Bel_i).$$

The external conflict an be evaluated by

$$U_C(Bel) - \sum_{i=1}^{k} a_i U_C(Bel_i).$$

The solution of the optimization problem

We will solve the optimization problem, when the internal conflict is equal to zero, i.e. $U_C(Bel_i) = 0$, $i = 1, \ldots, k$. The following algorithm allows us to minimize the number of clusters k.

Algorithm 1.

1. Output data: $Bel \in M_{bel}(X)$, A is a body of evidence of Bel.
2. To find a subset $B \subseteq X$ with the smallest cardinality such that $Pl(B) = 1$.
3. Assume that $B = \{y_1, ..., y_k\}$, then

$$\begin{aligned}
\mathcal{A}_1 &= \{A \in \mathcal{A} | y_1 \in A\}, \\
\mathcal{A}_2 &= \{A \in \mathcal{A} \backslash \mathcal{A}_1 | y_2 \in A\}, \\
&\vdots \\
\mathcal{A}_k &= \{A \in \mathcal{A} \backslash (\mathcal{A}_1 \cup ... \cup \mathcal{A}_{k-1} | y_k \in A\}.
\end{aligned}$$

The solution of the optimization problem

It easy to see that Algorithm 1 finds the required representation

$$Bel_i = \frac{1}{c_i} \sum_{B \in \mathcal{A}_i} m(B) \eta_{\langle B \rangle},$$

where $c_i = \sum_{B \in \mathcal{A}_i} m(B)$ but there are some suspicions that this clusterization does not exactly reflect the structure of given data. Let us have a partition $\{\mathcal{A}_1, ..., \mathcal{A}_k\}$ of $\mathcal{A}$ such that

$$y_1 \in \bigcap\nolimits_{A_i \in \mathcal{A}_1} A_i, ..., y_k \in \bigcap\nolimits_{A_i \in \mathcal{A}_k} A_i.$$

Then we say that this clusterization keeps the structure of given data if the belief function $Bel_B = \sum_{A_i \in \mathcal{A}} m(A_i) \eta_{\langle B \cap A_i \rangle}$ has the same inner conflict as Bel.

The solution of the optimization problem

Proposition 2

Let $P \in M_{pr}$ be a solution of the optimization problem for finding $U_C(Bel)$, i.e. $P \in \mathbf{P}(Bel)$ and $U_C(P) = U_C(Bel)$. Then for $B = \{x \in X | P(\{x\}) > 0\}$ we have $U_C(Bel_B) = U_C(Bel)$.

Algorithm 2

1. To find $P \in \mathbf{P}(Bel)$ such that $U_C(P) = U_C(Bel)$.
2. Let $\{y_1, ..., y_k\} = \{x \in X | P(\{x\}) > 0\}$ and $P(\{y_1\}) \geq P(\{y_2\}) \geq ... \geq P(\{y_k\})$. Then

$$\begin{aligned} \mathcal{A}_1 &= \{A \in \mathcal{A} | y_1 \in A\}, \\ \mathcal{A}_2 &= \{A \in \mathcal{A} \backslash \mathcal{A}_1 | y_2 \in A\}, \\ &\vdots \\ \mathcal{A}_k &= \{A \in \mathcal{A} \backslash (\mathcal{A}_1 \cup ... \cup \mathcal{A}_{k-1} | y_k \in A\}. \end{aligned}$$

The solution of the optimization problem

Proposition 3

Let $Bel \in M_{bel}$ and $P \in \mathbf{P}(Bel)$ with $U_C(P) = U_C(Bel)$. Let us denote $\{y_1, ..., y_k\} = \{x \in X | P(\{x\}) > 0\}$ and let $\mathcal{A}$ be the body of evidence of Bel. Then there is a partition $\{\mathcal{A}_1, ..., \mathcal{A}_k\}$ of $\mathcal{A}$ such that $y_1 \in \bigcap_{A_i \in \mathcal{A}_1} A_i, \ldots, y_k \in \bigcap_{A_i \in \mathcal{A}_k} A_i$ and

$$Bel = \sum_{i=1}^{k} P(\{y_i\}) Bel_i,$$

where $P(\{y_i\}) Bel_i = \sum_{B \in \mathcal{A}_i} m(B) \eta_{\langle B \rangle}$, $i = 1, ..., k$.

The solution of the optimization problem

A function $g : [0, 1] \to [0, +\infty)$ is called strictly concave if $g(x + \Delta x + \Delta y) - g(x + \Delta x) - g(x + \Delta y) + g(x) < 0$ for every $x, x + \Delta x, x + \Delta y, x + \Delta x + \Delta y \in [0, 1]$ and $\Delta x, \Delta y > 0$. If g is twice differentiable on [0,1], then g is strictly concave if $g''(x) < 0$ for any $x \in [0, 1]$.

Proposition 4

Let U_C on M_{bel} be constructed using the system of functions f_n from Proposition 1, and let g be a strictly concave function. Let $Bel \in M_{bel}(X)$ with a body of evidence $\mathcal{A}$ and assume that $U_C(Bel) = U_C(P)$, where $P \in \mathbf{P}(Bel)$. Then we can obtain the partition from Proposition 4, using Algorithm 2.

Positions of parties and their influence on voting results in Germany

Let X be the set of questions. Parties of Germany should answer on them before elections. Lepskiy and developed a methodic that allows to evaluate the importance of any group of questions based on voting results.

This information can be represented by the bba $m : 2^X \longrightarrow [0, 1]$ or the corresponding belief function.

We tried to analyze this information using conflict measures and other characteristics known in the theory of belief functions.

At first, we choose the 8 most valuable questions, i.e. $X = \{x_1, ..., x_8\}$. The Shapley value known also as the pignistic transformation gives us the result

$$V = (0.11, 0.175, 0.152, 0.11, 0.121, 0.11, 0.11, 0.11)\,.$$

Positions of parties and their influence on voting results in Germany

The proposed approach based on clustering a body of evidence can reflect the conflict among pieces of information and in a view of the our example can reveal the degree of the heterogeneity of society w.r.t. the list of questions.
For this purpose, we have applied Algorithm 2 to our belief function with the following parameters: the conflict measure U_C is defined as in Theorem 1 with Φ from Example 3, i.e. $g(t) = t(1-t)$. We found that the value of $U_C(Bel)$ is achieved on a probability measure $P_0 \in \mathbf{P}(Bel)$, whose values are defined by a vector

$$(0.014, 0.345, 0.005, 0.028, 0.54, 0.003, 0.055, 0.009),$$

Thus, vouters have shown the highest consolidation in question 5, the rest of them in question 2, and etc.

Characteristics of clusters

Contour functions of clusters

$$Pl_i(\{x\}) = \sum_{x \in A | A \in \mathcal{A}_i} m(A),$$

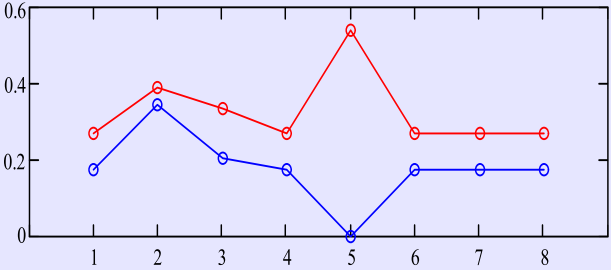

Fig. Contour functions of clusters: the red line is for cluster 5; the blue line is for cluster 2

Characteristics of clusters

The Shapley values

$$V_i(x) = \frac{1}{a_i} \sum_{x \in A | A \in \mathcal{A}_i} m(A)/|A|,$$

where $x \in X$, $a_i = \sum_{x \in A | A \in \mathcal{A}_i} m(A)$, $i = 2, 5$.

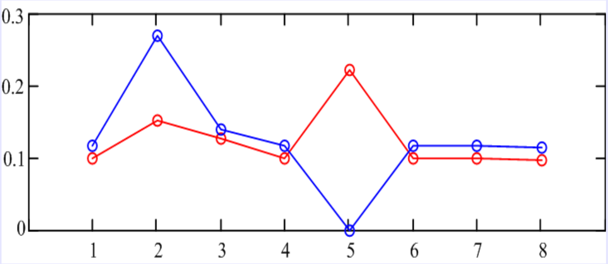

Fig. Shapley values of clusters: the red line is for cluster 5; the blue line is for cluster 2

Conclusions about clusterization

The most of focal elements are pinned to the questions 5 or 2. Question 5 is about increasing of the retirement age, and question 2 is about subsidies to families whose children study in non-state institutions. About 3.5 millions of vouters proposed to include question 5 to the list. Question 2 has been supported by 1 millions of vouters, but its contribution to importance in groups of questions was sufficently high to form the representative cluster.

Fig. 2 shows that the contribution of question 2 in cluster 2 is higher than the contribution of question 5 to cluster 5 according to the Shapley values. This reflects the fact that the question 5 is important for a larger number of voters than the question 2. Therefore, it is included in a larger number of significant coalitions of questions.

Thanks for your attention

brone@mail.ru
alex.lepskiy@gmail.com