Aggregation and Ranking on an Ordinal Scale Using Threshold Evidential Combination Rules

Alexander Lepskiy

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

The 11th International Conference on Information Technology and Quantitative Management (ITQM 2024),

August 23–25, 2024, Bucharest, Romania

Research Motivation

- We will consider the problem of aggregating information and ranking alternatives in the case of an ordinal scale. There are several aggregation methods in such a scale [Aleskerov 1999].
- It is desirable to consider the degree of inconsistency of information coming from different sources, the degree of its uncertainty, the reliability of these sources, when developing aggregation rules. These features are well modeled in the theory of evidence [Dempster 1967, Shafer 1976].
- An evidential approach to aggregating and ranking alternatives from several sources on an ordinal scale was proposed in a recent work by the author [Lepskiy 2023].
- As an extension of the previous approach, this study introduces a new class of aggregation rules (so-called threshold rules) that are more robust and stable.

Outline of Presentation

- Necessary Information from the Theory of Evidence;
- Threshold Functions and Operations of Evidence Theory
  - Similarity Coefficients;
  - Threshold Rules and Measures;
  - Averaging Operators;
- General Scheme of Evidence-Based Aggregation and Ranking in an Ordinal Scale;
  - Averaging Operators;
  - Ranking of Aggregate Estimates;
- Numerical Example;
- Summary and Conclusion.

Information from the Theory of Evidence

Let [Shafer 1976, Lepskiy 2023]:

- $T = \{t_1, ..., t_s\}$ be some finite set with a strict linear order of its elements $t_1 < \ldots < t_s$. We will consider on T ordered sequential sets of elements of the form $A = \{t_r, t_{r+1}, \ldots, t_{r+q}\} \subseteq T$.;
- $\left(2^T\right)_{od} \subseteq 2^T$ be the set of all such subsets of T;
- $m : \left(2^T\right)_{od} \to [0, 1]$, $\sum_{A \in \left(2^T\right)_{od}} m(A) = 1$, $m(\emptyset) = 0$ is the basic belief assignment (mass function);
- $\mathcal{A} = \{A\}$ is the set of all focal elements, i. e. $A \in \mathcal{A}$ if $m(A) > 0$;
- $F = (\mathcal{A}, m)$ is the body of evidence (BE);
- $F_A = (\{A\}, 1)$ is called **categorical** BE. In particular, $F_T = (\{T\}, 1)$ is called **vacuous** BE;
- an arbitrary BE $F = (\mathcal{A}, m)$ can be represented as $F = \sum_{A \in \mathcal{A}} m(A) F_A$;
- $F_A^\alpha = \alpha F_A + (1 - \alpha) F_T$, $\alpha \in (0, 1)$ is called **simple** BE.

We will use conjunctive rules to aggregate the vector scores represented by BEs $F_k = (\mathcal{A}_k, m_k)$, $k = 1, \ldots, n$. The non-normalized conjunctive rule has the form $F_\cap = \bigotimes_{k=1}^{n}{}_\cap F_k = (\mathcal{A} \cup \{\emptyset\}, m_\cap)$, where

$$m_\cap(A) = \sum_{B_1 \cap \ldots \cap B_n = A} m_1(B_1) \ldots m_n(B_n).$$

The value

$$Con(F_1, \ldots, F_n) = m_\cap(\emptyset) = \sum_{B_1 \cap \ldots \cap B_n = \emptyset} m_1(B_1) \ldots m_n(B_n)$$

characterizes the degree of inconsistency of sources of information presented by BEs, and is called a measure of conflict.

If $m_\cap(\emptyset)>0$, then it is necessary to redistribute the mass $m_\cap(\emptyset)$ among other focal elements. Next, we will use the classical Dempster rule $\otimes_D$, in which the redistribution of mass $m_\cap(\emptyset)$ is carried out uniformly over all focal elements: $m(A)=\frac{1}{1-m_\cap(\emptyset)}m_\cap(A)$, if $A\neq\emptyset$ and $m(\emptyset)=0$.

The BE $F=(\mathcal{A},m)$ can be transformed into the **pignistic probability** Bet_F:

$$Bet_F(t_i)=\frac{1}{1-m(\emptyset)}\sum_{A\in\mathcal{A},\, t_i\in A}\frac{m(A)}{|A|},\quad i=1,\ldots,s.$$

Threshold Operations of Evidence Theory. Similarity coefficients

Since conjunctive rules are based on counting the masses of intersecting focal elements, the main idea of using threshold rules is to take into account only "significant" such intersections in operations.

We will use the n-ary similarity coefficient $S(B_1, \dots, B_n)$ to measure the degree of intersection. It must satisfy the conditions:

1. $0 \leq S(B_1, \dots, B_n) \leq 1$;
2. $S(B_1, \dots, B_n) = 0 \Leftrightarrow B_1 \cap \dots \cap B_n = \emptyset$;
3. $S(B_1, \dots, B_n) = 1 \Leftrightarrow B_1 = \dots = B_n$.

Examples of such indices are:

$$J(B_1, \dots, B_n) = \frac{|B_1 \cap \dots \cap B_n|}{|B_1 \cup \dots \cup B_n|}, \; D(B_1, \dots, B_n) = \frac{2^{|B_1 \cap \dots \cap B_n|} - 1}{2^{|B_1 \cup \dots \cup B_n|} - 1}. \quad (1)$$

We will consider only those focal elements "significant" for which $S(B_1, \dots, B_n) > h$ is true for some threshold value $h \in [0, 1)$.

Threshold Rules and Measures

Then, for example, Dempster's threshold rule will have the form (notation: $F^{(h)} = \bigotimes_{k=1}^{n}{}_h F_k$):

$$m_h(A) = \frac{1}{K_h} \sum_{\substack{B_1 \cap \ldots \cap B_n = A, \\ S(B_1,\ldots,B_n) > h}} m_1(B_1) \ldots m_n(B_n), \quad m_h(\emptyset) = 0, \tag{2}$$

where $K_h = \sum_{S(B_1,\ldots,B_n)>h} m_1(B_1) \ldots m_n(B_n)$.

The value

$$Con_h(F_1, \ldots, F_n) = 1 - K_h = \sum_{S(B_1,\ldots,B_n) \le h} m_1(B_1) \ldots m_n(B_n) \tag{3}$$

can be considered as a threshold measure of conflict between BEs.

Averaging Operators

When we make a decision based on information obtained using threshold rules, we can either choose some specific threshold $h \in [0, 1)$ or we can use integral characteristics.

If g_h is some point characteristic integrable with respect to $h \in [0, 1)$, then the linear averaging operator can be used

$$Av_w(g.) = \int\limits_0^1 w(h) g_h \, dh, \tag{4}$$

where $w(h)$ is a non-negative integrable weight function that regulates the importance priorities of the values $h \in [0, 1)$ and satisfies the condition $\int\limits_0^1 w(h) dh = 1$.

General Scheme of Evidence-Based Ranking

Suppose that information from n experts about l alternatives $\{a_1, \ldots, a_l\}$ is given in the form of simple BEs

$$F_{jk} = \alpha_{jk} F_{A_{jk}} + (1 - \alpha_{jk})\, F_T, \quad k = 1, \ldots, n, \quad j = 1, \ldots, l$$

on a rank scale $T = \{t_1, \ldots, t_s\}$, $t_1 < \ldots < t_s$. The BE F_{jk} informs that the assessment of the j-th alternative belongs to the interval $A_{jk} \subseteq T$ in the opinion of the k-th expert with a degree of confidence $\alpha_{jk} \in [0, 1]$.

Let us perform the following steps for an evidential threshold ranking of these alternatives.

1. We aggregate the BEs F_{jk} from all n experts and for each j-th alternative using Dempster's threshold aggregation rule (1)-(3). We obtain BEs $F_j^{(h)} = \bigotimes_{k=1}^{n}{}_h F_{jk}$, $j = 1, \ldots, l$, $h \in [0, 1)$.
2. Next, we apply the averaging functional (4) to the aggregated estimates $F_j^{(h)}$. We obtain the BE $F_j = Av_w(F_j^{(\cdot)})$, $j = 1, \ldots, l$.
3. Let us find the pignistic probabilities Bet_{F_j}, $j = 1, \ldots, l$.
4. Finally, we will perform the final ranking of alternatives $\{a_1, \ldots, a_l\}$ using some ranking function $g(a_j)$, which determines the rank of the alternative:

$$a_j \succ a_r \;\; (a_j \sim a_r), \;\; \text{if} \;\; g(a_j) > g(a_r) \;\; (g(a_j) = g(a_r)).$$

The following functions are examples of rankings based on pignistic probabilities:

- $g(a_j) = \text{mode } Bet_{F_j^{(h)}} = \arg\max_{t_i} Bet_{F_j^{(h)}}(t_i)$ for some specific threshold value $h \in [0, 1)$;
- $g(a_j) = \arg\max_{t_i} Av_w(Bet_{F_j^{(\cdot)}})(t_i)$.
- $g(a_j) = \text{med } Bet_{F_j^{(h)}}$ is the median of the pignistic probability Bet_{F_j}.

Numerical Example

Let us illustrate the proposed scheme for aggregating expert assessments and ranking on the data of reviewing articles in the EasyChair conference management system (`https://easychair.org`).

This system uses a seven-rank ordinal rating scale $T = \{t_1, \dots, t_7\}$. These ranks t_i, $i = 1, \dots, 7$ meet the recommendations (in ascending order) "strong reject", "reject", "weak reject", "borderline paper", "weak accept", "accept", "strong accept". In addition, the reviewer gives an assessment on a five-rank scale (c_1 – "none", c_2 – "low", c_3 – "medium", c_4 – "high", c_5 – "expert") about the degree of confidence in the correctness of his decision: $c_1 < \dots < c_5$. For simplicity, we will assume that the degrees of confidence are given on a numerical scale by the formula $c_p = 0.2p$, $p = 1, \dots, 5$.

Point data $\left\{\left(x_k^{(j)}, c_k^{(j)}\right)\right\}_{k=1}^{3}$ of $n = 3$ reviewers regarding $l = 4$ papers $\{a_1, \ldots, a_4\}$ are presented in Table (k is the index of the reviewer, j is the index of the article), where $x_k^{(j)} \in T$.

	paper a_1	paper a_2	paper a_3	paper a_4
reviewer 1	$(t_6, 0.8)$	$(t_6, 1)$	$(t_6, 0.8)$	$(t_4, 0.6)$
reviewer 2	$(t_5, 1)$	$(t_5, 0.4)$	$(t_6, 0.6)$	$(t_5, 0.4)$
reviewer 3	$(t_4, 0.8)$	$(t_5, 0.6)$	$(t_3, 0.6)$	$(t_6, 1)$

Let us apply the procedure of blurring point estimates $x_k^{(j)}$ taking into account the degrees of confidence $c_k^{(j)}$ (see [Lepskiy 2023]). As a result, we get simple BEs

$$F_{jk} = c_k^{(j)} F_{A\left(x_k^{(j)}, c_k^{(j)}\right)} + \left(1 - c_k^{(j)}\right) F_T.$$

Table of focal elements and their masses $\left(A\left(x_k^{(j)}, c_k^{(j)}\right), c_k^{(j)}\right)$.

	paper a_1	paper a_2	paper a_3	paper a_4
reviewer 1	$(\{t_5, t_6\}, 0.8)$	$(\{t_6\}, 0.95)$	$(\{t_5, t_6\}, 0.8)$	$(\{t_3, t_4, t_5\}, 0.6)$
reviewer 2	$(\{t_5\}, 0.95)$	$(\{t_3, t_4, t_5, t_6\}, 0.4)$	$(\{t_5, t_6, t_7\}, 0.6)$	$(\{t_3, t_4, t_5, t_6\}, 0.4)$
reviewer 3	$(\{t_4\}, 0.8)$	$(\{t_4, t_5, t_6\}, 0.6)$	$(\{t_2, t_3, t_4\}, 0.6)$	$(\{t_6\}, 0.95)$
$Con^{(j)}$	0.792	0	0.552	0.57

Let us aggregate the BEs $F_j^{(h)} = \bigotimes_{k=1}^{n}{}_h F_{jk}$ using Dempster's threshold rule (2) for each j-th article, $j = 1, \ldots, l$. Next, we apply the averaging functional (4) with weight w to the aggregated estimates $F_j^{(h)}$:
$F_j = Av_w(F_j^{(\cdot)})$, $j = 1, \ldots, l$. Let us find the pignistic probabilities of these aggregated estimates Bet_{F_j}. The ranks of articles obtained using this procedure for some different averaging weights w and the **median ranking** rule are given in Table. The last two lines show the ranks for $h = 0$ and $h = 0.2$ respectively.

	a_1	a_2	a_3	a_4
$w = 1$	t_5	t_5	t_5	t_5
$w = 2(1-h)$	t_5	t_6	t_5	t_5
$w = \max\{6 - 18h, 0\}$	t_5	t_6	t_5	t_6
$w = 3(1-h)^2$	t_5	t_6	t_5	t_5
$h = 0$	t_5	t_6	t_5	t_6
$h = 0.2$	t_5	t_6	t_5	t_4

Next Table shows the ranks of articles found by **maximizing pignistic probabilities**: a) $Bet_{F_j^{(h)}}(t_i)$ with $h = 0$ and $h = 0.2$; b) $Av_w(Bet_{F_j^{(\cdot)}})(t_i)$ for all weights from previous Table.

	a_1	a_2	a_3	a_4
$h = 0$	t_5	t_6	t_5, t_6	t_6
$h = 0.2$	t_5, t_6	t_6	t_5, t_6	t_3, t_4, t_5
$Av_w(Bet_{F_j^{(\cdot)}})(t_i)$	t_5	t_6	t_5, t_6	t_6

The results obtained by maximizing the averaged pignistic probabilities are **more stable** (they depend little on the weights, the threshold-free result coincides with the average weighted result), but are **less sensitive** (articles often have multiple ranks) compared to the median ranking.

The ranks of articles $\{a_1, a_2, a_3\}$ do not depend on the threshold procedures or the averaging weight for median ranking.

Only the rank of the article a_4 changes. It drops from t_6 to (on average) t_4 when aggregated with $h = 0.2$.

Let us present for comparison the result of ranking using **linear convolution**

$$C(a_j) = \sum_{k=1}^{3} c_k^{(j)} x_k^{(j)}, \ \ j = 1, 2, 3, 4$$

(we will assume that reviewers' ratings $x_k^{(j)}$ are given on a numerical scale $t_i = i$, $i = 1, \dots, 7$, and $c_k^{(j)}$ are weights reviewers' decisions).

We get for our data that $C(a_1) = 13$, $C(a_2) = 11$, $C(a_3) = 10.2$, $C(a_4) = 10.4$. This corresponds to the ranking

$$a_1 \succ a_2 \succ a_4 \succ a_3,$$

which differs significantly from the results of the evidential rankings discussed above.

Summary and Conclusion

- threshold aggregation rules within the framework of evidence theory are introduced in this work. In these rules, only significant focal elements relative to some measure of similarity and a specified threshold are taken into account when aggregating bodies of evidence;
- the use of threshold rules in the ranking problem (especially after applying averaging procedures) improves the robustness of the results to small changes in the source data and/or the blurring procedure used, improves the sensitivity of the method, and allows us to vary the degree of significance of the focal elements that participate in the aggregation procedure;
- the presence of additional aggregation attributes in the form of a similarity measure and a threshold makes it possible to manage the conflict of aggregated bodies of evidence, formulate problems of optimally finding a threshold, etc.

References

- F. Aleskerov, Arrovian Aggregation Models, Dordrecht: Kluwer, 1999.
- A. Dempster, Upper and lower probabilities induced by multivalued mapping, Ann. Math. Statist. 38 (1967) 325–339.
- G. Shafer, A Mathematical Theory of Evidence, Princeton Univ. Press, 1976.
- A. Lepskiy, Evidence-based aggregation and ranking in an ordinal scale, Procedia Computer Science 221C (2023) 1066–1073.

Thanks for you attention

alepskiy@hse.ru
https://www.hse.ru/en/org/persons/10586209