A Dimensionality Reduction Approach for Qualitative Preference Aggregation

Quentin Brabant\(^1\), Miguel Couceiro\(^1\), Fabien Labernia\(^2\), Amedeo Napoli\(^1\)

\(^1\) LORIA (CNRS, Inria Nancy Grand Est - Université de Lorraine)
\{quentin.brabant, miguel.couceiro, amedeo.napoli\}@inria.fr
\(^2\) LAMSADE (CNRS - Université Paris-Dauphine)
fabien.labernia@dauphine.fr

1 Qualitative preference aggregation models

In this paper we briefly present a method for reducing the dimensionality of data in a qualitative preference aggregation framework. For a more complete description of this approach, see [4]. For an alternative approach based on rough sets theory, see [1].

We consider the following setting. \(X\) is a set of alternatives that are evaluated according to a set of criteria represented by there indices: \([n] = \{1, \ldots, n\}\). For an alternative \(x \in X\) we denote by \((x_1, \ldots, x_n) \in L^n\) the tuple of the evaluations of \(x\) in each crieterion. \(L\) is called the evaluation space, and is a distributive lattice for which we denote respectively by 0 and 1 the minimal and maximal element. We consider a binary preference relation \(\preceq\) between the alternatives that can be expressed in terms of a utility function:

\[
\forall x, y \in X : x \preceq y \Rightarrow U(x) \leq U(y),
\]

where \(U : X \rightarrow L\) associates a global evaluation on \(L\) to each alternative, and is obtained through the aggregation of the evaluations in criteria by a Sugeno integral \(S_\mu : L^n \rightarrow L\). In other words we have \(U(x) = S_\mu(x_1, \ldots, x_n)\). The Sugeno integral defined over distributive lattices [3], is expressed

\[
S_\mu(x_1, \ldots, x_n) = \bigvee_{I \subseteq [n]} \mu(I) \bigwedge_{i \in I} x_i,
\]

where \(\mu : 2^{[n]} \rightarrow L\) a capacity, that is to say a non-decreasing set function on \([n]\), with \(\mu(\emptyset) = 0\) and \(\mu([n]) = 1\). Capacities (and Sugeno integrals) are defined by a value on \(L\) for each subset of \([n]\), and therefore carry an intrinsic complexity, that grows exponentially with \(n\). We now consider a set \(\mathcal{D} = \{(x^1_1, y^1), \ldots, (x^m, y^m)\} \subseteq L^n \times L\), where each \(x^i = (x^i_1, \ldots, x^i_n) \in L^n\) is a tuple of evaluations in \(n\) criteria, and \(y^i\) is a utility value associated to \(x^j\). We want to learn a Sugeno integral \(S_\mu\) that generalizes these data. Ideally this function would be such that \(S_\mu(x^j) = y^j\) for any \(j \in \{1, \ldots, m\}\).

However, it is very common that no such function exists: in that case \(\mathcal{D}\) is said to be inconsistent, and we aim at learning a Sugeno integral that realizes the prediction of \(y^j\) for each element, with an error as low as possible. Because of the nature of capacities, this optimization problem is on \(2^n\) variables, and is therefore hard to solve when a high number of criteria is considered.
2 Dimensionnality reduction based on quality measure

By a quality measure over $\mathcal{D}$ we mean a degree with which $\mathcal{D}$ satisfies a certain hypothesis. In this presentation we consider two of such measures.

The first quality measure is the monotonicity degree, that is, the ratio of pairs $\{i, j\} \subseteq \{1, \ldots, m\}$ that satisfy the following condition

$$y^i > y^j \Rightarrow \exists k \in [n] : y^i_k > y^j_k.$$ 

This condition can be seen as a generalization of the Pareto condition to partially ordered evaluation spaces. The second quality measure is the compatibility degree, that is, the ratio of pairs satisfying the the condition

$$\exists S_{\mu} : [S_{\mu}(x^i) = y^i \text{ and } S_{\mu}(x^j) = y^j].$$ (1) 

This condition is justified by results from [2] that apply only when $L$ is totally ordered. Indeed it can be shown that $\mathcal{D}$ is consistent if and only if (2) is true for any pair from $\mathcal{D}$. Moreover, for a given pair this condition can be checked in a linear time w.r.t. $n$. Hence, provided that $L$ is totally ordered, the compatibility degree is both theoretically meaningful and practically interesting. If $L$ is not totally ordered, the monotonicity degree is the quality measure that makes sense.

The principle of the algorithm for dimensionality reduction that we propose is to iteratively remove a criterion, in order to minimize the decrease of the quality of the dataset at each step. Criteria are deleted until it is impossible to remove a criterion without decreasing the quality of the data below a certain ratio $\alpha$. This algorithm was tested on empirical data $^1$ and allowed a reduction of the number of criteria from 7 to 3. Aggregation models trained on original data and on data reduced to 3 criteria showed to have similar accuracy. On the other hand, models trained on data with only 2 criteria left had significantly worse accuracy, suggesting that a reduction to 3 criteria constitutes the best compromise between simplicity and accuracy for these data.

Future research work should include further empirical studies and should aim to determining a procedure for deciding the optimal value of $\alpha$, currently being set by hand.

References


$^1$Tripadvisor: http://sifaka.cs.uiuc.edu/~wang296/Data/index.html