Behavioral analysis of aggregation functions

Jean-Luc Marichal[1], Michel Grabisch[2], Radko Mesiar[3], and Endre Pap[4]

[1] University of Luxembourg, Luxembourg
[2] University of Panthéon-Sorbonne
[3] Slovak University of Technology, Slovakia
[4] University of Novi Sad, Serbia

Abstract

1 Introduction

Aggregation functions arise wherever aggregating information is important: applied and pure mathematics (probability, statistics, decision theory, functional equations), operations research, computer science, and many applied fields (economics and finance, pattern recognition and image processing, data fusion, etc.). For recent references, see Beliakov et al. [2] and Grabisch et al. [9].

Let $\mathbb{I}$ be a real interval, bounded or not. Given an aggregation function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$, it is often useful to define values or indices that offer a better understanding of the general behavior of F with respect to its variables. These indices may constitute a kind of identity card of F and enable one to classify the aggregation functions according to their behavioral properties.

For example, given an internal aggregation function A ("internal" means $\mathsf{Min} \leqslant \mathsf{A} \leqslant \mathsf{Max}$, where Min and Max are the minimum and maximum functions, respectively), it might be convenient to appraise the degree to which A is conjunctive, that is, close to Min. Similarly, it might be very instructive to know which variables, among $x_1, \ldots, x_n$, have the greatest influence on the output value $\mathsf{A}(\mathbf{x})$.

In this note, we present various indices, such as: andness and orness degrees of internal functions, idempotency degrees of conjunctive and disjunctive functions, importance and interaction indices, tolerance indices, and dispersion indices.

Sometimes different indices can be considered to measure the same behavior. In that case it is often needed to choose an appropriate index according to the nature of the underlying aggregation problem.

The material presented here is a summary of [9, Chapter 10], a forthcoming monograph on aggregation functions written by the authors.

2 Expected values and distribution functions

A very informative treatment of a given function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ consists in applying it to a random input vector and examining the behavior of the output signal by computing its distribution function. However, determining an explicit form of the distribution remains very difficult in general.

Instead, we can calculate the expected value or, more generally, the moments of the output variable and derive indices that would provide information on the location of the output values within the range of the function.

3 Importance indices

When using a given aggregation function A of n variables, one may wonder which are the most influential variables in the computation of $\mathsf{A}(\mathbf{x})$, if any. We may say that no such variable exists if A is symmetric. In case symmetry does not hold, e.g., for weighted aggregation functions and for integral-based ones, it is very instructive to know the level or percentage of contribution of each variable in the computation of the result. We call this level of contribution or influential power the *importance index*.

For weighted aggregation functions, a naive answer to the above question is to take as importance index simply the weight of each variable. A first simple reason to discard this idea is that the definition, meaning and normalization of weights differ from one aggregation function to another: just consider the weighted arithmetic mean, with weights in $[0,1]$ summing up to 1, and the weighted maximum, whose weights are in $[0,1]$ with no summation condition, but whose maximum value is 1. A second reason is that intuitively, a weight value, say 0.5, does not have the same effect in a weighted arithmetic mean as in a weighted geometric mean.

A natural approach when $\mathbb{I}$ is a bounded closed interval $[a,b]$ is to compute an average of the marginal contribution of variable x_i.

4 Interaction indices

Although the notion of importance index is useful to analyze a given aggregation function, the description it provides is still very primitive. Take for example the arithmetic mean, the minimum, and the maximum. Since they are symmetric, they have the same importance index for all coordinates, yet they are extremely different aggregation functions, because the minimum operator is a conjunctive aggregation function, the maximum operator is disjunctive, and the arithmetic mean is neither conjunctive nor disjunctive.

The question is how to quantify or describe the difference between these aggregation functions, which are indistinguishable by the importance index. Since andness and orness are reserved to internal aggregation functions, other indices have to be found.

The key of the problem lies in the interrelation between variables. The notion of importance index is based on the variation of the aggregated value vs. the total variation of a given variable, the others being fixed. We may consider the variation induced by the mixed variation of two variables, or more. This is expressed by the *second order (total) variation of* A *with respect to coordinates* i *and* j.

5 Tolerance indices

Some internal aggregation functions are more or less intolerant (respectively, tolerant) in the sense that they are bounded from above (respectively, below) by one of the input values or by a function of these values.

Here, we mainly deal with internal aggregation functions having veto and/or favor coordinates as well as k-conjunctive and k-disjunctive internal aggregation functions. Starting from the properties of these functions, we define indices that provide degrees to which an internal aggregation function is intolerant or tolerant.

6 Measures of arguments contribution and involvement

Given an aggregation function A, we consider in this final section the following two indices:

1. *The index of uniformity of arguments contribution*, which measures the uniformity of contribution of the n components of $\mathbf{x} \in \mathbb{I}^n$ in the computation of the aggregated value $\mathsf{A}(\mathbf{x})$.
2. *The index of arguments involvement*, which measures the proportion of arguments among $x_1, \ldots, x_n$ that are involved in the computation of the aggregated value $\mathsf{A}(\mathbf{x})$.

References

1. J. F. Banzhaf. Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, 19:317–343, 1965.
2. G. Beliakov, A. Pradera, and T. Calvo. *Aggregation Functions: A Guide for Practitioners*. Studies in Fuziness and Soft Computing. Springer, Berlin, 2007.
3. J. J. Dujmović. Seven flavors of andness/orness. In B. De Baets, J. Fodor, and D. Radojević, editors, *Proc. Eurofuse 2005*, pages 81–92, Belgrade, 2005.
4. J. M. Fernández Salido and S. Murakami. Extending Yager's orness concept for the OWA aggregators to other mean operators. *Fuzzy Sets and Systems*, 139(3):515–542, 2003.
5. K. Fujimoto, I. Kojadinovic, and J.-L. Marichal. Axiomatic characterizations of probabilistic and cardinal-probabilistic interaction indices. *Games Econom. Behav.*, 55(1):72–99, 2006.
6. M. Grabisch. Alternative representations of discrete fuzzy measures for decision making. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 5(5):587–607, 1997.
7. M. Grabisch and C. Labreuche. How to improve acts: an alternative representation of the importance of criteria in MCDM. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 9(2):145–157, 2001.
8. M. Grabisch, C. Labreuche, and J.-C. Vansnick. On the extension of pseudo-Boolean functions for the aggregation of interacting criteria. *European J. Oper. Res.*, 148(1):28–47, 2003.
9. M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap. *Aggregation functions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, UK, to appear in 2009.
10. M. Grabisch, J.-L. Marichal, and M. Roubens. Equivalent representations of set functions. *Mathematics of Operations Research*, 25(2):157–178, 2000.
11. M. Grabisch and M. Roubens. An axiomatic approach to the concept of interaction among players in cooperative games. *Internat. J. Game Theory*, 28(4):547–565, 1999.

12. I. Kojadinovic, J.-L. Marichal, and M. Roubens. An axiomatic approach to the definition of the entropy of a discrete Choquet capacity. *Information Sciences*, 172:131–153, 2005.
13. A. Kolesárová. Revision of parametric evaluation of aggregation functions. In *Proc. 13th Zittau Fuzzy Colloquium*, pages 202–211, Zittau, Germany, September 2006.
14. C. Labreuche. Determination of the criteria to be improved first in order to improve as much as possible the overall evaluation. In *10th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2004)*, pages 609–616, Perugia, Italy, July 2004.
15. J.-L. Marichal. Behavioral analysis of aggregation in multicriteria decision aid. In *Preferences and decisions under incomplete knowledge*, pages 153–178. Physica, Heidelberg, 2000.
16. J.-L. Marichal. The influence of variables on pseudo-Boolean functions with applications to game theory and multicriteria decision making. *Discrete Appl. Math.*, 107(1-3):139–164, 2000.
17. J.-L. Marichal. Entropy of discrete Choquet capacities. *European J. Oper. Res.*, 137(3):612–624, 2002.
18. J.-L. Marichal. Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. *European J. Oper. Res.*, 155(3):771–791, 2004.
19. J.-L. Marichal. Cumulative distribution functions and moments of lattice polynomials. *Statistics & Probability Letters*, 76(12):1273–1279, 2006.
20. J.-L. Marichal. k-intolerant capacities and Choquet integrals. *European J. Oper. Res.*, 177(3):1453–1468, 2007.
21. J.-L. Marichal. Multivariate integration of functions depending explicitly on the minimum and the maximum of the variables. *J. Math. Anal. Appl.*, 341(1):200–210, 2008.
22. J.-L. Marichal. Weighted lattice polynomials of independent random variables. *Discrete Applied Mathematics*, 156(5):685–694, 2008.
23. J.-L. Marichal and I. Kojadinovic. Distribution functions of linear combinations of lattice polynomials from the uniform distribution. *Statistics & Probability Letters*, 78(8):985–991, 2008.
24. J.-L. Marichal, I. Kojadinovic, and K. Fujimoto. Axiomatic characterizations of generalized values. *Discrete Applied Mathematics*, 155(1):26–43, 2007.
25. J.-L. Marichal and P. Mathonet. Approximations of Lovász extensions and their induced interaction index. *Discrete Appl. Math.*, 156(1):11–24, 2008.
26. J.-L. Marichal and M. Roubens. Entropy of discrete fuzzy measures. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 8(6):625–640, 2000.
27. T. Murofushi. A technique for reading fuzzy measures (I): the Shapley value with respect to a fuzzy measure. In *2nd Fuzzy Workshop*, pages 39–48, Nagaoka, Japan, October 1992. In Japanese.
28. T. Murofushi and M. Sugeno. Some quantities represented by the Choquet integral. *Fuzzy Sets and Systems*, 56(2):229–235, 1993.
29. G. Owen. Multilinear extensions of games. *Management Sci.*, 18:P64–P79, 1972.
30. L. S. Shapley. A value for n-person games. In *Contributions to the theory of games, vol. 2*, Annals of Mathematics Studies, no. 28, pages 307–317. Princeton University Press, Princeton, N. J., 1953.
31. R. J. Weber. Probabilistic values for games. In A. E. Roth, editor, *The Shapley value*, pages 101–119. Cambridge Univ. Press, Cambridge, 1988.
32. R. R. Yager. A class of fuzzy measures generated from a Dempster-Shafer belief structure. *Int. J. Intell. Syst.*, 14(12):1239–1247, 1999.
33. R. R. Yager. On the entropy of fuzzy measures. *IEEE Trans. on Fuzzy Systems*, 8:453–461, 2000.