Measuring the interactions among variables of functions over the unit hypercube

J.-L. Marichal and P. Mathonet
University of Luxembourg, Mathematics Research Unit, FSTC
{jean-luc.marichal,pierre.mathonet}[at]uni.lu

We consider a situation where a single dependent variable, call it y, is explained in terms of n independent variables $x_1, \ldots, x_n$ through an equation (a model) of the form $y = f(x_1, \ldots, x_n)$, where f is a real function of n variables.

Here we suppose that the function f describing the model is given and that we want to investigate its behavior through simple terms. For instance, suppose we want to measure the overall contribution (importance or influence) of each independent variable to the model. A natural approach to this problem consists in defining the overall importance of each variable as the coefficient of this variable in the least squares linear approximation of f. This approach was considered by Hammer and Holzman [6] for pseudo-Boolean functions and cooperative games $f\colon \{0,1\}^n \to \mathbb{R}$. They observed that the coefficient of each variable in the linear approximation is exactly the well-known Banzhaf power index [1] of the corresponding player in the game f.

In many practical situations, the information provided by the overall importance degree of each variable may be far insufficient due to the possible interactions among the variables. Then, a more flexible approach to investigate the behavior of f consists in measuring an overall importance degree for each combination (subset) of variables. Such a concept was first introduced in [7] for Boolean functions $f\colon \{0,1\}^n \to \{0,1\}$ (see also [2]), then in [8] for pseudo-Boolean functions and games $f\colon \{0,1\}^n \to \mathbb{R}$ (see also [9]), and in [4] for square integrable functions $f\colon [0,1]^n \to \mathbb{R}$.

In addition to these importance indexes, we can also measure directly the interaction degree among the variables by defining an overall interaction index for each combination of variables. This concept was introduced axiomatically in [5] (see also [3]) for games $f\colon \{0,1\}^n \to \mathbb{R}$. However, it has not yet been extended to real functions defined on $[0,1]^n$. We intend to fill this gap by defining and investigating an appropriate index to measure the interaction degree among variables of a given square integrable function $f\colon [0,1]^n \to \mathbb{R}$.

Our sources of inspiration to define such an index are actually threefold. In cooperative game theory, the Banzhaf interaction index can be obtained from a least squares approximation of the game under consideration. In analysis, a local interaction of variables in a function f can be obtained from the limited Taylor expansion of f. Finally, in statistics, two-way interactions appear as the coefficients of leading terms in quadratic models, three-way interactions appear as the coefficients of leading terms in cubic models, and so forth.

We then naturally consider the least squares approximation problem of a given square integrable function $f\colon [0,1]^n \to \mathbb{R}$ by a multilinear polynomial of a given degree. Then, given a subset $S \subseteq \{1, \ldots, n\}$, an index $\mathcal{I}(f, S)$ measuring the interaction among the variables $\{x_i : i \in S\}$ of f is defined as the coefficient of the monomial $\prod_{i \in S} x_i$ in the best approximation of f by a multilinear polynomial of degree at most $|S|$.

We show that this new index has many appealing properties, such as linearity, continuity, and symmetry. In particular, we show that, similarly to the Banzhaf interaction index introduced for games, the index $\mathcal{I}(f, S)$ can be interpreted in a sense as an expected value of the discrete derivative of f in the direction of S or as an expected value of the difference quotient of f in the direction of S. These latter results reveal a strong analogy between the interaction index and the overall importance index introduced by Grabisch and Labreuche [4].

We finally discuss various applications, including the computation of explicit expressions of the interaction index for certain classes of functions. We also define and investigate a normalized version of the interaction index to compare different functions in terms of interaction degrees of their variables and a coefficient of determination to measure the quality of multilinear approximations.

Références

[1] J. F. Banzhaf. Weighted voting doesn't work : A mathematical analysis. *Rutgers Law Review*, 19 :317–343, 1965.

[2] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson, and N. Linial. The influence of variables in product spaces. *Isr. J. Math.*, 77(1-2) :55–64, 1992.

[3] 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.

[4] M. Grabisch and C. Labreuche. How to improve acts : An alternative representation of the importance of criteria in MCDM. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 9(2) :145–157, 2001.

[5] M. Grabisch and M. Roubens. An axiomatic approach to the concept of interaction among players in cooperative games. *Int. J. Game Theory*, 28(4) :547–565, 1999.

[6] P. Hammer and R. Holzman. Approximations of pseudo-Boolean functions ; applications to game theory. *Z. Oper. Res.*, 36(1) :3–21, 1992.

[7] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In *Proc. 29th Annual Symposium on Foundations of Computational Science*, pages 68–80. Computer Society Press, 1988.

[8] 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.

[9] J.-L. Marichal, I. Kojadinovic, and K. Fujimoto. Axiomatic characterizations of generalized values. *Discrete Applied Mathematics*, 155(1) :26–43, 2007.