Multilinear model: New issues in capacity identification

Guilherme D. Pelegrina[1] and **Michel Grabisch**[2] and **Leonardo T. Duarte**[3] and **João M. T. Romano**[4]

Abstract. In the literature, few research has been done with respect to the problem of capacity identification in the multilinear model. In this paper, we provide a preliminary investigation of a supervised approach and also future perspectives to deal with this problem.

1 Introduction

A typical problem in Multiattribute Utility Theory (MAUT) [6] is to model preference relations among a set of alternatives $\mathbf{x}_j = (x_{1,j}, x_{2,j}, \ldots, x_{m,j})$, $j = 1, \ldots, n$, evaluated according to a set C of m criteria. Suppose that the evaluation of each alternative $\mathbf{x}_j$ with respect to the criterion i is given by the value function $u_i(x_{i,j})$. Moreover, consider that the global evaluation of alternative $\mathbf{x}_j$ is given by the overall value function $U(\mathbf{x}_j) = F(u(\mathbf{x}_j))$, where $u(\mathbf{x}_j) = (u_1(x_{1,j}), \ldots, u_m(x_{m,j}))$ and $F(\cdot)$ is an aggregation function [5]. Therefore, the aim in MAUT is to find a function $F(\cdot)$ that represents the preference relation $\succeq$ of the decision maker:

$$\mathbf{x}_j \succeq \mathbf{x}_{j'} \Leftrightarrow U(\mathbf{x}_j) \geq U(\mathbf{x}_{j'}), \ \ \forall j, j' = 1, \ldots, n, \tag{1}$$

where $\succeq$ indicates that $\mathbf{x}_j$ is at least as good as $\mathbf{x}_{j'}$. From now on, for simplicity of notation, we refer to $u_i(x_{i,j})$ as $u_{i,j}$.

The definition of $F(\cdot)$ depends on the hypotheses about the decision problem. For instance, if one assumes a mutual preferential independence [6] among criteria, one may use an additive function, such as the Weighted Arithmetic Mean (WAM) [5], given by

$$F_{WAM}(u(\mathbf{x}_j)) = \sum_{i=1}^{m} w_i u_{i,j}, \tag{2}$$

where the weight w_i ($w_i \geq 0$, $\sum_{i=1}^{m} w_i = 1$) represents the importance of criterion i in the decision problem. However, the independence assumption does not hold in several practical situations. Therefore, it becomes necessary to apply an aggregation function that takes into account the interaction among criteria, such as the Choquet integral [2] and the multilinear model [8]. In order to perform the desired interaction, both approaches use a set function [5] (more precisely, a capacity [2]), which is generally determined through a capacity identification method.

Several optimization approaches in the literature deal with the capacity identification problem in the context of Choquet integral [4]. However, one does not find the same references for the multilinear model. Therefore, this paper provides a preliminary discussion on the capacity identification problem with respect to this method. More precisely, we formulate the addressed optimization problem in a supervised fashion and perform numerical experiments to verify our proposal. Moreover, we discuss our ongoing works, which comprise the use of regularization terms and a non-supervised approach.

[1] FEEC, University of Campinas; CES, University of Paris I, email: pelegrina@decom.fee.unicamp.br
[2] PSE, University of Paris I, michel.grabisch@univ-paris1.fr
[3] FCA, University of Campinas, email: leonardo.duarte@fca.unicamp.br
[4] FEEC, University of Campinas, email: romano@dmo.fee.unicamp.br

2 The multilinear model

The multilinear model [8] is an approach that uses a polynomial aggregation of the criteria evaluations. It is defined as follows:

$$F_{ML}(\mathbf{x}_j) = \sum_{A \subseteq C} \mu(A) \prod_{i \in A} u_{i,j} \prod_{i \in \overline{A}} (1 - u_{i,j}), \tag{3}$$

where $u_{i,j} \in [0, 1]$, $\overline{A}$ is the complement set of A and $\mu : 2^C \to \Re$ is a set function satisfying the axioms (i) $\mu(\emptyset) = 0$, (ii) $\mu(C) = 1$ and (iii) if $A \subseteq B \subseteq C$, $\mu(A) \leq \mu(B) \leq \mu(C)$, i.e. μ is a capacity [5].

In order to illustrate the application of this method, consider a decision problem with 4 alternatives and 2 criteria, both with equal importance and to be maximized. Figure 1 illustrates this scenario. It is clear that $\mathbf{x}_1$ and $\mathbf{x}_4$ are the worst and the best alternatives, respectively. Therefore, one may consider that $F(u(\mathbf{x}_1)) = 0$ and $F(u(\mathbf{x}_4)) = 1$ for any $F(\cdot)$. Moreover, due to monotonicity, $F(u(\mathbf{x}_1)) \leq F(u(\mathbf{x}_j))$ and $F(u(\mathbf{x}_4)) \geq F(u(\mathbf{x}_j))$ for all j. In this situation, different decision makers may have the following opinions:

- **Opinion 1:** Since both $\mathbf{x}_2$ and $\mathbf{x}_3$ have an evaluation equals to 0, they are as bad as $\mathbf{x}_1$, i.e. $F(u(\mathbf{x}_j)) = 0, \forall j = 1, 2, 3$. Therefore, an alternative $\mathbf{x}_j$ is satisfactory if both evaluations are satisfactory (positive interaction among criteria).
- **Opinion 2:** Since both $\mathbf{x}_2$ and $\mathbf{x}_3$ have an evaluation equals to 1, they are as good as $\mathbf{x}_4$, i.e. $F(u(\mathbf{x}_j)) = 0, \forall j = 2, 3, 4$. Therefore, in order to an alternative $\mathbf{x}_j$ be satisfactory, it is sufficient that one evaluation be satisfactory (negative interaction among criteria).
- **Opinion 3:** Since both $\mathbf{x}_2$ and $\mathbf{x}_3$ satisfy one criterion, they are better than $\mathbf{x}_1$, but worse than $\mathbf{x}_4$. In this case, there are no interaction among criteria and the importance of an alternative is approximately the sum of the importance of each criterion.

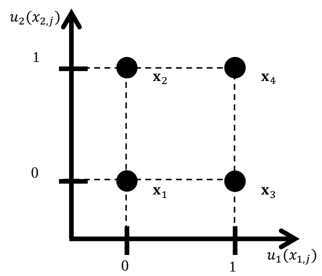

Figure 1. Evaluations of alternatives.

Clearly, WAM satisfies only opinion 3 (with, $w_1 = w_2 = 0.5$). However, if we consider the multilinear model (with $\mu(\{1, 2\}) = 1$), all the opinions may be satisfied. For instance, one satisfies opinions 1, 2 and 3 with $\mu(\{1\}) = \mu(\{2\})$ equal to 0, 1 and 0.5, respectively.

3 Capacity identification - Supervised approach

Section 2 presented an example with only 2 unknown capacity coefficients[5], which could be easily identified. However, in scenarios with more criteria, this task becomes difficult. Therefore, an optimization procedure is generally used to perform capacity identification.

In this section, we formulate the optimization problem. For instance, consider a supervised approach whose aim is to minimize the mean squared error between the obtained evaluations $F_{ML}(u(\mathbf{x}_j))$ and the desired ones $y(u(\mathbf{x}_j))$, given by

$$\sum_{j=1}^{n} \left(F_{ML}(u(\mathbf{x}_j)) - y(u(\mathbf{x}_j))\right)^2 . \tag{4}$$

One may represent this cost function in terms of vectors and matrices. In this case, consider the set of capacities $\mu = [\mu(\emptyset), \mu(\{1\}), \ldots, \mu(\{1, \ldots\}), \ldots, \mu(C)]^T$ in lexicographic representation, the set of desired evaluation $\mathbf{y} = [y(u(\mathbf{x}_1)), y(u(\mathbf{x}_2)), \ldots, y(u(\mathbf{x}_n))]^T$ and the following matrix:

$$\mathbf{P} = \begin{bmatrix} \prod\limits_{i\in C}(1-u_{i,1}) & \cdots & \prod\limits_{i\in C}(1-u_{i,n}) \\ \prod\limits_{i\in\{1\}} u_{i,1} \prod\limits_{i\in\overline{\{1\}}}(1-u_{i,1}) & \cdots & \prod\limits_{i\in\{1\}} u_{i,n} \prod\limits_{i\in\overline{\{1\}}}(1-u_{i,n}) \\ \vdots & \ddots & \vdots \\ \prod\limits_{i\in C} u_{i,1} & \cdots & \prod\limits_{i\in C} u_{i,n} \end{bmatrix} .$$

The optimization problem, in a quadratic form, can be represented by

$$\begin{aligned} &\min_{\mu} \frac{1}{2}\mu^T\mathbf{Q}\mu + \mathbf{v}^T\mu + \mathbf{y}^T\mathbf{y} \equiv \min_{\mu} \frac{1}{2}\mu^T\mathbf{Q}\mu + \mathbf{v}^T\mu \\ &\qquad\qquad \mathit{Subject\ to\ constraints\ on}\ \mu \end{aligned} \tag{5}$$

where $\mathbf{Q} = 2\mathbf{P}\mathbf{P}^T$ and $\mathbf{v} = -2\mathbf{P}\mathbf{y}$. The constraints on μ guarantees that the axioms of the capacity be satisfied (see Section 2).

4 Numerical experiment

This section presents a numerical experiment that verifies the application of the quadratic model to retrieve the capacity coefficients. For instance, we consider decision problems with $m = 3, 4, 5$ or 6 criteria and with a number of training data varying from 1 to 120. As performance index, we calculate the mean squared error

$$\epsilon = \frac{1}{2^m - 2} \sum_{A\in C, A\neq\emptyset, A\neq C} \left(\mu(A) - \hat{\mu}(A)\right)^2 , \tag{6}$$

where μ is the capacity used to obtain $y(u(\mathbf{x}_j))$ and $\hat{\mu}$ is the retrieved one. Figure 2 presents the obtained results (averaged over 50 simulations). It is worth mentioned that we randomly generated both μ and $u(\mathbf{x}_j)$.

One may note that, for $m = 3, 4, 5$ and 6 criteria, $\epsilon \approx 0$ when the number of learning data is greater than 6, 15, 43 and 110, respectively. If one has a lower number of training data, the problem is ill-posed, and the retrieved capacity are not necessarily the desired one.

[5] Note that the number of unknown capacity coefficients is equal to $2^m - 2$, since $\mu(\emptyset) = 0$ and $\mu(C) = 1$.

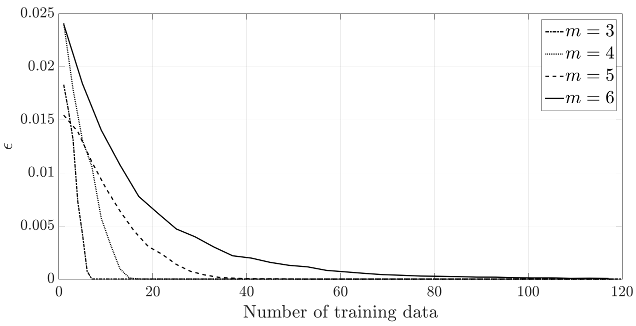

Figure 2. Mean squared error for different number of training data.

5 A supervised approach with regularization

Since (5) may be ill-posed, one may use a regularization term to perform capacity identification. For instance, [1] and [7] apply the ℓ_1-norm to retrieve the Choquet integral parameters. Similarly to [7], our future work consists in the application of this regularization term in the Banzhaf interaction index $\mathbf{I}_B = [I_B(\emptyset), I_B(\{1\}), \ldots, I_B(\{1, \ldots\}), \ldots, I_B(C)]^T$ [5], which leads to the simplest multilinear model whose capacity tends to be additive. Therefore, one may represent the optimization problem by

$$\begin{aligned} &\min_{\mathbf{I}_B} \frac{1}{2}\mathbf{I}_B^T\mathbf{Q}'\mathbf{I}_B + \mathbf{v}'^T\mathbf{I}_B + \lambda \left\|\mathbf{I}_B\right\|_1 \\ &\qquad \mathit{Subject\ to\ constraints\ on}\ \mathbf{I}_B, \end{aligned} \tag{7}$$

where λ is a constant, $\mathbf{Q}' = \mathbf{S}^T\mathbf{Q}\mathbf{S}$, $\mathbf{v}' = \mathbf{S}^T\mathbf{v}$ and $\mathbf{S}$ is the transformation matrix from $\mathbf{I}_B$ to μ, i.e. $\mu = \mathbf{S}\mathbf{I}_B$.

6 Conclusions and future perspectives

This paper verifies the application of the multilinear model to perform interaction among criteria and investigates the capacity identification problem. Our preliminary results indicates that, if the number of training data is not enough, one should consider more information about the decision problem to retrieve the correct capacity. Future works comprise the application of regularization terms in the optimization problem, as described in Section 5. Moreover, we also aim at developing a non-supervised method that is able to estimate the capacity of a multilinear model without any training data. In this context, our ongoing work consists the application of an approach inspired by [3], which takes into account the covariance matrix of the decision data.

Acknowledgements. The authors would like to thank the FAPESP (Processes n. 2016/21571-4 and 2017/23879-9) and the CNPq (Process n. 311357/2017-2) for the financial support.

REFERENCES

[1] T. A. Adeyeba, D. T. Anderson, and T. C. Havens, 'Insights and characterization of l_1-norm based sparsity learning of a lexicographically encoded capacity vector for the Choquet integral', in *2015 IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE)*. IEEE, (2015).

[2] G. Choquet, 'Theory of capacities', *Annales de l'Institut Fourier*, **5**, 131–295, (1954).

[3] L. T. Duarte, 'A novel multicriteria decision aiding method based on unsupervised aggregation via the Choquet integral', *IEEE Transactions on Engineering Management*, **65**(2), 293–302, (2018).

[4] M. Grabisch, I. Kojadinovic, and P. Meyer, 'A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package', *European Journal of Operational Research*, **186**, 766–785, (2008).

[5] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, *Aggregation functions*, Cambridge University Press, New York, 2009.

[6] R. L. Keeney and H. Raiffa, *Decisions with multiple objectives*, Wiley, New Tork, 1976.

[7] H. E. Oliveira, J. M. T. Romano, and L. T. Duarte, 'Identificação dos parâmetros da integral de Choquet via uma abordagem baseada em processamento de sinais esparsos', in *XXXV Simp. Bras. de Telecomunicações e Processamento de Sinais (SBrT)*, pp. 692–696, São Pedro, Brazil, (2017).

[8] G. Owen, 'Multilinear extensions of games', *Management Science*, **Part 2, 18**(5), 64–79, (1972).