A review of methods for capacity identification in Choquet integral based multi-attribute utility theory

Applications of the Kappalab R package

Michel Grabisch
Université de Paris I, LIP6
8 rue du Capitaine Scott, 75015 Paris, France
michel.grabisch@lip6.fr

Ivan Kojadinovic*
LINA CNRS FRE 2729, École Polytechnique de l'Université de Nantes
Rue Christian Pauc, 44306 Nantes, France
ivan.kojadinovic@polytech.univ-nantes.fr

Patrick Meyer
Applied Mathematics Unit, University of Luxembourg
162A, avenue de la Faïencerie, L-1511 Luxembourg, G.D. Luxembourg
patrick.meyer@uni.lu

February 7, 2007

Abstract

The application of multi-attribute utility theory whose aggregation process is based on the Choquet integral requires the prior identification of a capacity. The main approaches to capacity identification proposed in the literature are reviewed and their advantages and inconveniences are discussed. All the reviewed methods have been implemented within the Kappalab R package. Their application is illustrated on a detailed example.

Keywords: Multi-criteria decision aiding, multi-attribute utility theory, Choquet integral, free software.

1 Introduction

Let $X \subseteq X_1 \times \cdots \times X_n$, $n \geq 2$, be a set of objects of interest described by a set $N := \{1, \ldots, n\}$ of attributes. The aim of multi-attribute utility theory (MAUT) [1] is to model

*Corresponding author. Tel: + (33) 2 40 68 32 57. Fax: + (33) 02 40 68 32 32.

the preferences of the decision maker (DM), represented by a binary relation $\succeq$ on X, by means of an *overall utility function* $U : X \to \mathbb{R}$ such that,

$$x \succeq y \iff U(x) \geq U(y), \qquad \forall\, x, y \in X.$$

The function U is generally determined by means of an interactive and incremental process requiring from the DM that he/she expresses his/her preferences over a small subset of selected objects. The resulting global utility function can then be seen as providing a *numerical representation* of the preference relation $\succeq$ on X and can be used in applications as a *model* of the expertise of the DM.

The preference relation $\succeq$ is usually assumed to be complete and transitive. As far as the overall utility function is considered, the most frequently encountered model is the *additive value model* (see e.g. [2]). In this work, we consider the more general *transitive decomposable model* of Krantz et al. [3, 4] in which U is defined by

$$U(x) := F(u_1(x_1), \ldots, u_n(x_n)), \qquad \forall\, x = (x_1, \ldots, x_n) \in X, \tag{1}$$

where the functions $u_i : X_i \to \mathbb{R}$ are called the *utility functions* and $F : \mathbb{R}^n \to \mathbb{R}$, non-decreasing in its arguments, is sometimes called the *aggregation function*. As far as the utility functions are considered, for any $x \in X$, the quantity $u_i(x_i)$ can be interpreted as a measure of the satisfaction of the value x_i for the DM. From now on, the term *criterion* will be used to designate the association of an attribute $i \in N$ with the corresponding utility function u_i. For the previous decomposable model to hold, it is necessary that the preference relation is a *weakly separable* weak order (see e.g. [2]).

The exact form of the overall utility function U depends on the hypotheses on which the multi-criteria decision aiding (MCDA) problem is based. When *mutual preferential independence* (see e.g. [5]) among criteria can be assumed, it is frequent to consider that the function F is additive and takes the form of a weighted sum. The decomposable model given in Eq. (1) is then equivalent to the classical additive value model. The assumption of mutual preferential independence among criteria is however rarely verified in practice. In order to be able to take interaction phenomena among criteria into account, it has been proposed to substitute a monotone set function on N, called *capacity* [6] or *fuzzy measure* [7], to the weight vector involved in the calculation of weighted sums. Such an approach can be regarded as taking into account not only the importance of each criterion but also the importance of each subset of criteria. A natural extension of the weighted arithmetic mean in such a context is the *Choquet integral* with respect to (w.r.t.) the defined capacity [8, 9, 10].

The use of a Choquet integral as an aggregation function in Eq. (1) requires, as we shall see, the ability to compare the utility of an object according to different criteria. In other words, it is necessary that the utility functions be *commensurable*, i.e. $u_i(x) = u_j(x)$ if and only if, for the DM, the object x is satisfied to the same extent on criteria i and j; see e.g. [11] for a more complete discussion on commensurability. In the considered context, commensurable utility functions can be determined using the extension of the MACBETH methodology [12] proposed in [10]; see also [11, 13]. This task is not trivial and can take a large percentage of the time dedicated to a MCDA/preference modeling problem. We will not discuss it further here, as the topic of this paper is the *capacity identification problem*.

Assuming that the utility functions have been determined, the learning data from which the capacity is to be identified consists of the *initial preferences* of the DM: usually, a partial

weak order over a (small) subset of the set X of objects, a partial weak order over the set of criteria, intuitions about the importance of the criteria, etc. The precise form of these prior preferences will be discussed in Section 3.

The aim of this paper is to review the main approaches to capacity identification proposed in the literature that can be applied to Choquet integral based MAUT. As we shall see, most of the presented methods can be stated as optimization problems. They differ w.r.t. the objective function and the preferential information they require as input. For each of the reviewed methods, we point out its main advantages/disadvantages. The last part of this paper is devoted to a presentation of the discussed identification methods through the Kappalab toolbox [14]. Kappalab, which stands for "laboratory for capacities", is a package for the GNU R statistical system [15], which is a Matlab like free software environment for statistical computing and graphics. Kappalab contains high-level routines for capacity and non-additive integral manipulation on a finite setting which can be used in the framework of decision aiding or cooperative game theory. All the reviewed identification methods have been implemented within Kappalab and can be easily used through the high-level R language.

The paper is organized as follows. The second section is devoted to the presentation of the Choquet integral as an aggregation operator and to numerical indices that can be used to understand its behavior. In the third section, we discuss the form of the preferential information from which the capacity is to be identified. The next section contains a review of the main approaches to capacity identification existing in the literature. The application of the reviewed methods is illustrated in the last section through the use of the Kappalab R package.

In order to avoid a heavy notation, we will omit braces for singletons and pairs, e.g., by writing $\mu(i)$, $N \setminus ij$ instead of $\mu(\{i\})$, $N \setminus \{i,j\}$. Furthermore, cardinalities of subsets $S, T, \ldots$, will be denoted by the corresponding lower case letters $s, t, \ldots$ Also, as classically done, the asymmetric part of a binary relation $\succeq$ will be denoted by $\succ$ and its symmetric part by $\sim$. Finally, the power set of N will be denoted by $\mathcal{P}(N)$.

2 The Choquet integral as an aggregation operator

2.1 Capacities and Choquet integral

As mentioned in the introduction, *capacities* [6], also called *fuzzy measures* [7], can be regarded as generalizations of weighting vectors involved in the calculation of weighted sums.

Definition 2.1. *A capacity on N is a set function $\mu : \mathcal{P}(N) \to [0,1]$ satisfying the following conditions:*

(i) $\mu(\emptyset) = 0$, $\mu(N) = 1$,

(ii) for any $S, T \subseteq N$, $S \subseteq T \Rightarrow \mu(S) \leq \mu(T)$.

Furthermore, a capacity μ on N is said to be

- *additive* if $\mu(S \cup T) = \mu(S) + \mu(T)$ for all disjoint subsets $S, T \subseteq N$,
- *cardinality-based* if, for any $T \subseteq N$, $\mu(T)$ depends only on the cardinality of T.

Note that there is only one capacity on N that is both additive and cardinality-based. We call it *the uniform capacity* and denote it by μ^*. It is easy to verify that μ^* is given by

$$\mu^*(T) = t/n, \qquad \forall\, T \subseteq N.$$

In the framework of aggregation, for each subset of criteria $S \subseteq N$, the number $\mu(S)$ can be interpreted as the *weight* or the *importance* of S. The monotonicity of μ means that the weight of a subset of criteria cannot decrease when new criteria are added to it.

When using a capacity to model the importance of the subsets of criteria, a suitable aggregation operator that generalizes the weighted arithmetic mean is the Choquet integral [8, 9, 10].

Definition 2.2. *The Choquet integral of a function $x : N \to \mathbb{R}$ represented by the vector $(x_1, \ldots, x_n)$ w.r.t. a capacity μ on N is defined by*

$$C_\mu(x) := \sum_{i=1}^{n} x_{\sigma(i)}[\mu(A_{\sigma(i)}) - \mu(A_{\sigma(i+1)})],$$

where σ is a permutation on N such that $x_{\sigma(1)} \leq \cdots \leq x_{\sigma(n)}$. Also, $A_{\sigma(i)} := \{\sigma(i), \ldots, \sigma(n)\}$, for all $i \in \{1, \ldots, n\}$, and $A_{\sigma(n+1)} := \emptyset$.

Seen as an aggregation operator, the Choquet integral w.r.t. μ can be considered as taking into account interaction phenomena among criteria, that is, *complementarity* or *substitutivity* among elements of N modeled by μ [9].

The Choquet integral generalizes the weighted arithmetic mean in the sense that, as soon as the capacity is additive, which intuitively coincides with the independence of the criteria, it collapses into a weighted arithmetic mean.

An intuitive presentation of the Choquet integral is given in [16]. An axiomatic characterization of the Choquet integral as an aggregation operator can be found in [9]. Note that the first use of the Choquet integral in decision analysis is probably due to Schmeidler in the context of decision under uncertainty; see e.g. [17].

2.2 The Möbius transform of a capacity

Any set function $\nu : \mathcal{P}(N) \to \mathbb{R}$ can be uniquely expressed in terms of its *Möbius representation* [18] by

$$\nu(T) = \sum_{S \subseteq T} m_\nu(S), \qquad \forall\, T \subseteq N, \tag{2}$$

where the set function $m_\nu : \mathcal{P}(N) \to \mathbb{R}$ is called the *Möbius transform* or *Möbius representation* of ν and is given by

$$m_\nu(S) = \sum_{T \subseteq S} (-1)^{s-t} \nu(T), \qquad \forall\, S \subseteq N. \tag{3}$$

Of course, any set of 2^n coefficients $\{m(S)\}_{S\subseteq N}$ does not necessarily correspond to the Möbius transform of a capacity on N. The boundary and monotonicity conditions must be ensured [19], i.e., we must have

$$\begin{cases} m(\emptyset) = \emptyset, & \sum\limits_{T\subseteq N} m(T) = 1, \\ \sum\limits_{\substack{T\subseteq S \\ T\ni i}} m(T) \geq 0, & \forall\, S \subseteq N,\ \forall\, i \in S. \end{cases} \tag{4}$$

As shown in [19], in terms of the Möbius representation of a capacity μ on N, for any $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the Choquet integral of x w.r.t. μ is given by

$$C_{m_\mu}(x) = \sum_{T\subseteq N} m_\mu(T) \bigwedge_{i\in T} x_i, \tag{5}$$

where the symbol $\wedge$ denotes the minimum operator. The notation C_{m_μ}, which is equivalent to the notation C_μ, is used to emphasize the fact that the Choquet integral is here computed w.r.t. the Möbius transform of μ.

2.3 Analysis of the aggregation

The behavior of the Choquet integral as an aggregation operator is generally difficult to understand. For a better comprehension of the interaction phenomena modeled by the underlying capacity, several numerical indices can be computed [20, 21]. In the sequel, we present two of them in detail.

2.3.1 Importance index

The overall importance of a criterion $i \in N$ can be measured by means of its Shapley value [22], which is defined by

$$\phi_\mu(i) := \sum_{T\subseteq N\setminus i} \frac{(n-t-1)!\,t!}{n!}[\mu(T\cup i) - \mu(T)].$$

Having in mind that, for each subset of criteria $S \subseteq N$, $\mu(S)$ can be interpreted as the *importance* of S in the decision problem, the Shapley value of i can be thought of as an average value of the *marginal contribution* $\mu(T \cup i) - \mu(T)$ of criterion i to a subset T not containing it. A fundamental property is that the numbers $\phi_\mu(1), \ldots, \phi_\mu(n)$ form a probability distribution over N. In terms of the Möbius representation of μ, the Shapley value of i can be rewritten as

$$\phi_{m_\mu}(i) = \sum_{T\subseteq N\setminus i} \frac{1}{t+1} m_\mu(T\cup i). \tag{6}$$

2.3.2 Interaction index

In order to intuitively approach the concept of interaction, consider two criteria i and j such that $\mu(ij) > \mu(i) + \mu(j)$. Clearly, the previous inequality seems to model a *positive*

interaction or *complementary* effect between i and j. Similarly, the inequality $\mu(ij) < \mu(i) + \mu(j)$ suggests considering that i and j interact in a *negative* or *redundant* way. Finally, if $\mu(ij) = \mu(i) + \mu(j)$, it seems natural to consider that criteria i and j do not interact, i.e., that they have *independent* roles in the decision problem.

A coefficient measuring the interaction between i and j should therefore depend on the difference $\mu(ij) - [\mu(i) + \mu(j)]$. However, as discussed by Grabisch and Roubens [23], the intuitive concept of interaction requires a more elaborate definition. Clearly, one should not only compare $\mu(ij)$ and $\mu(i) + \mu(j)$ but also see what happens when i, j, and ij join other subsets. In other words, an index of interaction between i and j should take into account all the coefficients of the form $\mu(T \cup i)$, $\mu(T \cup j)$, and $\mu(T \cup ij)$, with $T \subseteq N \setminus ij$.

Murofushi and Soneda [24] suggested to measure the average interaction between two criteria i and j by means of the following *interaction index*:

$$I_\mu(ij) := \sum_{T \subseteq N \setminus ij} \frac{(n-t-2)!\, t!}{(n-1)!} [\mu(T \cup ij) - \mu(T \cup i) - \mu(T \cup j) + \mu(T)].$$

Notice that, given a subset T not containing i and j, the expression

$$\mu(T \cup ij) - \mu(T \cup i) - \mu(T \cup j) + \mu(T)$$

can be regarded as the difference between the marginal contributions $\mu(T \cup ij) - \mu(T \cup i)$ and $\mu(T \cup j) - \mu(T)$. We call this expression the *marginal interaction between i and j in the presence of T*. Indeed, it seems natural to consider that if

$$\mu(T \cup ij) - \mu(T \cup i) > \mu(T \cup j) - \mu(T) \text{ (resp. } <),$$

i and j interact *positively* (resp. *negatively*) in the presence of T.

The quantity $I_\mu(ij)$ can therefore be interpreted as a measure of the *average* marginal interaction between i and j. An important property is that $I_\mu(ij) \in [-1, 1]$ for all $ij \subseteq N$, the value 1 (resp. -1) corresponding to maximum complementarity (resp. substitutivity) between i and j [25]. In terms of the Möbius representation of μ, $I_\mu(ij)$ can be rewritten as

$$I_{m_\mu}(ij) = \sum_{T \subseteq N \setminus ij} \frac{1}{t+1} m_\mu(T \cup ij). \tag{7}$$

2.4 The concept of k-additivity

From the results presented in Sections 2.1 and 2.2, one can see that a capacity μ on N is completely defined by the knowledge of $2^n - 2$ coefficients, for instance $\{\mu(S)\}_{\emptyset \neq S \subsetneq N}$ or $\{m_\mu(S)\}_{\emptyset \neq S \subsetneq N}$. Such a complexity may be prohibitive in certain applications. The fundamental notion of *k-additivity* proposed by Grabisch [25] enables to find a trade-off between the complexity of the capacity and its modeling ability.

Definition 2.3. *Let $k \in \{1, \ldots, n\}$. A capacity μ on N is said to be* k-additive *if its Möbius representation satisfies $m_\mu(T) = 0$ for all $T \subseteq N$ such that $t > k$ and there exists at least one subset T of cardinality k such that $m_\mu(T) \neq 0$.*

As one can easily check, the notion of 1-additivity coincides with that of additivity. Notice that, in this case, it follows from Eq. (7) that the interaction index is zero for any pair of criteria, which is in accordance with the intuition that an additive capacity cannot model interaction. More generally, it can be shown that a k-additive capacity, $k \in \{2, \ldots, n\}$, can model interaction among at most k criteria; see e.g. [26].

Let $k \in \{1, \ldots, n\}$ and let μ be a k-additive capacity on N. From Eq. (2), we immediately have that

$$\mu(S) = \sum_{\substack{\emptyset \neq T \subseteq S \\ t \leq k}} m_\mu(T), \qquad \forall\, S \subseteq N,$$

which confirms that a k-additive capacity ($k < n$) is completely defined by the knowledge of $\sum_{l=1}^{k} \binom{n}{l}$ coefficients.

3 The identification problem

Assuming that the utility functions have been determined (see [10, 13]), the next step consists in identifying a capacity, if it exists, such that the Choquet integral w.r.t. this capacity numerically represents the preferences of the DM (see Eq. (1)). When dealing with a real-world preference modeling problem, the DM is generally asked to reason on a finite and usually small subset $\mathcal{O}$ of the set X of objects of interest. The set $\mathcal{O}$ is usually composed either of physically available objects (at hand) or of selected, potentially fictitious objects the reasoning on which may prove particularly useful for modeling the preferences of the DM. Using the terminology of statistical learning, the set $\mathcal{O}$ could be seen as a *learning set*. Its usually small cardinality (rarely more than 20 objects) is due to the fact that the expression of the preferences of the DM is generally a very time-consuming process. See for instance [5, Chap. 3] for a more complete discussion about the subset $\mathcal{O}$.

Once an appropriate subset $\mathcal{O}$ has been determined, the DM is asked to express his *initial preferences*. The use of the adjective *initial* will become clear at the end of this section. These preferences, from which the capacity is to be determined, can take the form of:

- a partial weak order $\succeq_{\mathcal{O}}$ over $\mathcal{O}$ (ranking of the available objects);
- a partial weak order $\succeq_N$ over N (ranking of the importance of the criteria);
- a partial weak order $\succeq_P$ on the set of pairs of criteria (ranking of interactions);
- etc.

In the context of MAUT based on the Choquet integral, it seems natural to translate some of the above prior information as follows:

- $x \succ_{\mathcal{O}} x'$ can be translated as $C_\mu(u(x)) - C_\mu(u(x')) \geq \delta_C$;
- $x \sim_{\mathcal{O}} x'$ can be translated as $-\delta_C \leq C_\mu(u(x)) - C_\mu(u(x')) \leq \delta_C$;
- $i \succ_N j$ can be translated as $\phi_\mu(i) - \phi_\mu(j) \geq \delta_{Sh}$;

- $i \sim_N j$ can be translated as $-\delta_{Sh} \leq \phi_\mu(i) - \phi_\mu(j) \leq \delta_{Sh}$;
- $ij \succ_P kl$ can be translated as $I_\mu(ij) - I_\mu(kl) \geq \delta_I$;
- $ij \sim_P kl$ can be translated as $-\delta_I \leq I_\mu(ij) - I_\mu(kl) \leq \delta_I$;

where $u(x) := (u_1(x_1), \ldots, u_n(x_n))$ for all $x \in X$, and δ_C, δ_{Sh} and δ_I are nonnegative indifference thresholds to be defined by the DM. In other terms, the partial weak orders $\succeq_{\mathcal{O}}$, $\succeq_N$, $\succeq_P$ previously mentioned are translated into partial semiorders with fixed indifference thresholds. Note that in practice a constraint of the form $I_\mu(ij) - I_\mu(kl) \geq \delta_I$ is generally accompanied by either the constraint $I_\mu(ij) \leq 0$ or the constraint $0 \leq I_\mu(kl)$.

Most of the identification methods proposed in the literature can be stated under the form of an optimization problem:

$$\text{min or max } \mathcal{F}(\ldots)$$
$$\text{subject to } \begin{cases} \nu(S \cup i) - \nu(S) \geq 0, \forall\, i \in N, \forall\, S \subseteq N \setminus i, \\ \nu(N) = 1, \\ C_\nu(u(x)) - C_\nu(u(x')) \geq \delta_C, \\ \vdots \\ \phi_\nu(i) - \phi_\nu(j) \geq \delta_{Sh}, \\ \vdots \\ I_\nu(ij) - I_\nu(kl) \geq \delta_I, \\ I_\nu(kl) \geq 0 \\ \vdots \end{cases}$$

where ν is a *game* on N, i.e. a set function $\nu : \mathcal{P}(N) \to \mathbb{R}$ such that $\nu(\emptyset) = 0$ and $\mathcal{F}$ is an objective function that differs w.r.t. to the chosen identification method.

A solution to the above problem is a general capacity defined by $2^n - 1$ coefficients. The number of variables involved in it increases exponentially with n and so will the computational time. For large problems, both for computational and simplicity reasons, it may be preferable to restrict the set of possible solutions to k-additive capacities, $k \in \{1, \ldots, n\}$, typically $k = 2$ or 3. The idea is here simply to rewrite the above optimization problem in terms of the Möbius transform of a k-additive game using Eqs. (3), (5), (6) and (7), which will decrease the number of variables from $2^n - 1$ to $\sum_{l=1}^{k} \binom{n}{l}$ as one can see from Table 1. We obtain

$$\text{min or max } \mathcal{F}(\ldots)$$
$$\text{subject to } \begin{cases} \sum\limits_{\substack{T \subseteq S \\ t \leq k-1}} m_\nu(T \cup i) \geq 0,\ \forall i \in N,\ \forall S \subseteq N \setminus i, \\ \sum\limits_{\substack{T \subseteq N \\ 0 < t \leq k}} m_\nu(T) = 1, \\ C_{m_\nu}(u(x)) - C_{m_\nu}(u(x')) \geq \delta_C, \\ \vdots \\ \phi_{m_\nu}(i) - \phi_{m_\nu}(j) \geq \delta_{Sh}, \\ \vdots \end{cases}$$

where m_ν is the Möbius representation of a k-additive game ν on N.

Of course, the above optimization problem may be infeasible if the constraints are inconsistent. Such a situation can arise for two main reasons :

Table 1: Influence of the order of k-additivity and the number n of criteria on the number of variables.

k / n	2	3	4	5	6	7	8	9	10
1-add	2	3	4	5	6	7	8	9	10
2-add	3	6	10	15	21	28	36	45	55
3-add	-	7	14	25	41	63	92	129	175
n-add	3	7	15	31	63	127	255	511	1023

- The preferential information provided by the DM is contradictory or violates natural axioms underlying most decision-making procedures such as compatibility with dominance, transitivity of strict preferences, etc.
- The number of parameters of the model, i.e. the number of coefficients of the Möbius transform, is too small to have all the constraints satisfied. In this case, in order to increase the number of free parameters, and therefore to be more likely to be able to model the DM's initial preferences, the approach usually consists in incrementing the order of k-additivity. It may happen however that even with a general (n-additive) capacity, the constraints imposed by the DM, still being in accordance with the previously mentioned natural axioms, cannot be satisfied. In such a case, some more specific axioms underlying the Choquet integral model are violated (see e.g. [27]) and the Choquet integral cannot be considered as sufficiently flexible for modeling the initial preferences of the DM.

It is important to note that, in practice, finding a solution to the above optimization problem does not necessarily end the identification process. Indeed, the obtained Choquet integral-based model is then usually analyzed by means of the indices presented in Section 2.3. If the results are not completely in accordance with the DM's reasoning, his *initial* preferences are enriched by additional constraints and a new identification is performed. Such an incremental process is carried out until a satisfactory model is found.

4 Review of the main approaches to capacity identification

As discussed in Section 3, most of the methods for capacity identification proposed in the literature can be stated as optimization problems. They differ according to their objective function and the preferential information they require as input. After presenting two approaches that could be seen as generalizations of multiple linear regression, we review methods based on *maximum split*, *minimum variance* and *minimum distance* identification principles. We end this section by describing a hybrid method enabling to provide an approximate solution if there are no capacities compatible with the DM's preferences.

4.1 Least-squares based approaches

Historically, the first approach that has been proposed can be regarded as a generalization of classical multiple linear regression [28]. It requires the additional knowledge of the desired overall evaluations $y(x)$ of the available objects $x \in \mathcal{O}$. The objective function is defined as

$$\mathcal{F}_{LS}(m_\nu) := \sum_{x \in \mathcal{O}} \left[C_{m_\nu}(u(x)) - y(x)\right]^2,$$

where $u(x) := (u_1(x_1), \ldots, u_n(x_n))$ for all $x \in X$. The aim is to minimize the average quadratic distance between the overall utilities $\{C_{m_\nu}(u(x))\}_{x \in \mathcal{O}}$ computed by means of the Choquet integral and the desired overall scores $\{y(x)\}_{x \in \mathcal{O}}$ provided by the DM. The optimization problem takes therefore the form of a quadratic program, not necessarily strictly convex [29], which implies that the solution, if it exists, is not necessarily unique (this aspect is investigated in detail in [30]). In order to avoid the use of quadratic solvers, heuristic suboptimal versions of this approach have been proposed by Ishii and Sugeno [31], Mori and Murofushi [28] and Grabisch [32]. Let us detail this latter approach.

Called *Heuristic Least Mean Squares* (HLMS), it is based on a gradient approach starting from a DM defined capacity μ that we shall call the *initial capacity*. This capacity is typically an additive capacity representing the DM's prior idea of what the aggregation function F in Eq. (1) should be. In the absence of clear requirements on the aggregation function, a very natural choice for μ is the uniform capacity μ^* since the Choquet integral w.r.t. that capacity is nothing else than the arithmetic mean. Once the initial capacity μ has been chosen, for each object $x \in \mathcal{O}$, the gradient modifies only the coefficients of μ involved in the computation of $C_\mu(u(x))$ (without violating monotonicity constraints). When all the available objects have been used, unmodified coefficients of μ are modified towards the average value of neighboring coefficients. This forms one iteration, and the process is restarted until a stopping criterion is satisfied. The advantage over the optimal quadratic approach is that only the vector of the coefficients of μ has to be stored, while in the latter, a squared matrix of same dimensions has to be stored. Also, as we shall see in Subsection 5.2, this heuristic approach tends to provide less extreme solutions than the optimal approach. However, unlike for the optimal quadratic method, it is not possible to require that the solution be k-additive, $k < n$.

In the context of MAUT based on the Choquet integral, the main inconvenience of these approaches is that they require the knowledge of the desired overall utilities $\{y(x)\}_{x \in \mathcal{O}}$, which cannot always be obtained from the DM.

4.2 A maximum split approach

An approach based on linear programming was proposed by Marichal and Roubens [33]. The proposed identification method can be stated as follows:

$$\max \mathcal{F}_{LP}(\varepsilon) := \varepsilon$$
$$\text{subject to } \begin{cases} \sum\limits_{\substack{T\subseteq S \\ t\leq k-1}} m_\nu(T \cup i) \geq 0,\ \forall i \in N,\ \forall S \subseteq N \setminus i, \\ \sum\limits_{\substack{T\subseteq N \\ 0<t\leq k}} m_\nu(T) = 1, \\ C_{m_\nu}(u(x)) - C_{m_\nu}(u(x')) \geq \delta_C + \varepsilon, \\ \vdots \end{cases}$$

Roughly speaking, the idea of the proposed approach is to maximize the minimal difference between the overall utilities of objects that have been ranked by the DM through the partial weak order $\succeq_{\mathcal{O}}$ (hence the name *maximum split*). Indeed, if the DM states that $x \succ_{\mathcal{O}} x'$, he may want the overall utilities to reflect this difference in the most significant way.

The main advantage of this approach is its simplicity. However, as the least squares based approach presented in the previous subsection, this identification method does not necessarily lead to a unique solution, if any. Furthermore, as it will be illustrated in Section 5, the provided solution can sometimes be considered as too extreme, since it corresponds to a capacity that maximizes the difference between overall utilities.

4.3 Minimum variance and minimum distance approaches

The idea of the minimum variance method [34] is to favor the "least specific" capacity, if any, compatible with the initial preferences of the DM. The objective function is defined as the *variance* of the capacity, i.e.

$$\mathcal{F}_{MV}(m_\nu) := \frac{1}{n} \sum_{i\in N} \sum_{S\subseteq N\setminus i} \gamma_s(n) \left(\sum_{T\subseteq S} m_\mu(T \cup i) - \frac{1}{n} \right)^2.$$

As shown in [34], minimizing this variance is equivalent to maximizing the extended Havrda and Charvat entropy of order 2. This method can therefore be equivalently regarded as a maximum entropy approach. The optimization problem takes the form of the following strictly convex quadratic program:

$$\min \mathcal{F}_{MV}(m_\nu)$$
$$\text{subject to } \begin{cases} \sum\limits_{\substack{T\subseteq S \\ t\leq k-1}} m_\nu(T \cup i) \geq 0,\ \forall i \in N,\ \forall S \subseteq N \setminus i, \\ \sum\limits_{\substack{T\subseteq N \\ 0<t\leq k}} m_\nu(T) = 1, \\ C_{m_\nu}(u(x)) - C_{m_\nu}(u(x')) \geq \delta_C, \\ \vdots \end{cases}$$

As discussed in [34, 35], the Choquet integral w.r.t. the minimum variance capacity compatible with the initial preferences of the DM, if it exists, is the one that will exploit the most on average its arguments.

One of the advantages of this approach is that it leads to a unique solution, if any, because of the strict convexity of the objective function. Also, in the case of "poor" initial preferences involving a small number of constraints, this unique solution will not exhibit too

specific behaviors characterized for instance by very high positive or negative interaction indices or a very uneven Shapley value.

A generalization of this approach [36] consists in finding, if it exits, the closest capacity to a capacity defined by the DM and compatible with his/her initial preferences. As already discussed in Subsection 4.1, this *initial capacity* is typically an additive capacity representing the DM's prior idea of what the aggregation function should be. In the absence of clear requirements a very natural choice for μ is the uniform capacity μ^*. In order to practically implement such a minimum distance principle, in [36], three quadratic distances have been studied. In the sequel, we shall restrict ourselves to the following one defined, for any two games μ, ν on N by

$$d^2(m_\mu, m_\nu) := \int_{[0,1]^n} [C_{m_\nu}(x) - C_{m_\mu}(x)]^2 dx. \tag{8}$$

This quadratic distance, thoroughly studied in [37, Chap. 7] in the context of the extension of pseudo-Boolean functions, can be interpreted as the expected quadratic difference between overall utilities computed by C_{m_μ} and C_{m_ν} assuming that the vectors of partial utilities are uniformly distributed in $[0,1]^n$.

In the absence of clear requirements on the aggregation function, a natural objective function for the above discussed minimum distance principle is thus given by

$$\mathcal{F}_{MD}(m_\nu) := \int_{[0,1]^n} \left[C_{m_\nu}(x) - \frac{1}{n} \sum_{i=1}^{n} x_i \right]^2 dx.$$

The resulting optimization problem is again a strictly convex quadratic program.

4.4 A less constrained approach

This approach, which is due to Meyer and Roubens [38], can be seen as a generalization of the least squares methods described in Subsection 4.1. The minimal preferential information which has to be provided by the DM is a weak order over the available objects. The objective function, depending on more variables than the least squares methods described earlier, is defined as

$$\mathcal{F}_{GLS}(m_\nu, y) := \sum_{x \in \mathcal{O}} \left[C_{m_\nu}(u(x)) - y(x) \right]^2,$$

where $y = \{y(x)\}_{x \in \mathcal{O}}$ are additional variables of the quadratic program[1] representing overall unknown evaluations of the objects that must verify the weak order imposed by the DM.

The optimization problem can be written as the following convex quadratic program :

$$\min \mathcal{F}_{GLS}(m_\nu, y)$$
$$\text{subject to } \begin{cases} \sum\limits_{\substack{T \subseteq S \\ t \leq k-1}} m_\nu(T \cup i) \geq 0, \ \forall i \in N, \ \forall S \subseteq N \setminus i, \\ \sum\limits_{\substack{T \subseteq N \\ 0 < t \leq k}} m_\nu(T) = 1, \\ y(x) - y(x') \geq \delta_y, \\ \vdots \end{cases}$$

[1]The acronym *GLS* stands for "generalized least squares".

where δ_y is a indifference threshold, playing a similar role as δ_C, i.e., it can be interpreted as the desired minimal difference between the overall utilities of two objects which are considered as significantly different by the DM. A solution of the quadratic program consists of the Möbius representation m_ν of the capacity and the overall evaluations $y = \{y(x)\}_{x\in\mathcal{O}}$.

Let us first intuitively explain the main idea of the approach. As discussed at the end of Section 3, assuming that the constraints imposed by the DM are not contradictory and do not question natural multicriteria decision axioms such as compatibility with dominance, it may still happen that the number of parameters (following from the chosen order of k-additivity) is too small so that these constraints can be satisfied. A first possibility consists in increasing k, if possible. A second solution consists in relaxing some of the constraints by translating the desired weak order over the available objects by means of conditions on the unknown overall evaluations $\{y(x)\}_{x\in\mathcal{O}}$. In that case, the Möbius transform m_ν is less constrained and there may exist a solution.

The role of the objective function is to minimize the quadratic difference between the numerical representation $\{y(x)\}_{x\in\mathcal{O}}$ of the weak order imposed by the DM and the overall utilities computed by means of the Choquet integral. If the objective function is zero, then for each object x, its overall evaluation $y(x)$ equals its aggregated overall utility $C_{m_\nu}(u(x))$. In that case, the weak order obtained by ordering the objects according to their aggregated evaluations is consistent with the weak order imposed by the DM and the threshold δ_y is not violated. Two possibilities arise: either there is a unique solution or there exists an infinity of solutions to the problem. In the second case, the solution is chosen by the solver and its characteristics are difficult to predict. If the objective function is strictly positive, then the aggregated overall evaluations $\{C_{m_\nu}(u(x))\}_{x\in\mathcal{O}}$ do not exactly match the overall unknown evaluations $\{y(x)\}_{x\in\mathcal{O}}$ numerically representing the weak order imposed by the DM. Two possibilities arise: either the weak order induced by the aggregated overall evaluations $\{C_{m_\nu}(u(x)\}_{x\in\mathcal{O}}$ corresponds to $\succeq_\mathcal{O}$ but $x \succ_\mathcal{O} x'$ does not necessarily imply $C_{m_\nu}(u(x)) \geq C_{m_\nu}(u(x')) + \delta_y$ (see Subsection 5.6), or the weak order induced by the $\{C_{m_\nu}(u(x)\}_{x\in\mathcal{O}}$ does not correspond to $\succeq_\mathcal{O}$. In the former case, the solution does not respect the DM's choice for the minimal threshold δ_y. In the latter case the weak order is violated "on average" which might not be very satisfactory.

The advantage of this approach is that it may provide a solution even if the weak order over the available objects is incompatible with a Choquet integral model because some specific axioms are violated (see e.g. [27]), if the indifference threshold δ_y is too large, or if some of the constraints on the criteria are not compatible with a representation of the weak order by a Choquet integral. It is then on to the DM to decide if the result is satisfactory or not. Nevertheless, as already explained, this approach should be used with care when the objective function is zero, since then, it simply amounts to letting the quadratic solver choose a feasible solution whose characteristics are difficult to predict.

4.5 Practical implementation

The discussed identification methods have been implemented within the Kappalab package [14] for the GNU R statistical system [15]. The package is distributed as free software and can be downloaded from the *Comprehensive R Archive Network* (`http://cran.r-project.org`) or from `http://www.polytech.univ-nantes.fr/kappalab`. To solve lin-

ear programs, the LpSolve R package [39] is used; strictly convex quadratic programs are solved using the Quadprog R package [40]; finally, not necessarily strictly convex quadratic programs are solved using the `ipop` routine of the Kernlab R package [41].

As far as the maximum number of criteria is considered, Kappalab allows to work comfortably with up to $n = 10$ criteria if n-additive capacities are considered and with up to $n = 32$ criteria if 2 or 3-additive capacities are considered.

5 Applications of the Kappalab R package

5.1 Problem description

We consider an extended version of the fictitious problem presented in [34] concerning the evaluation of students in an institute training econometricians. The students are evaluated w.r.t. five subjects: statistics (S), probability (P), economics (E), management (M) and English (En). The utilities of seven students a, b, c, d, e, f, g on a $[0, 20]$ scale are given in Table 2.

Table 2: Partial evaluations of the seven students.

Student	S	P	E	M	En	Mean
a	18	11	11	11	18	13.80
b	18	11	18	11	11	13.80
c	11	11	18	11	18	13.80
d	18	18	11	11	11	13.80
e	11	11	18	18	11	13.80
f	11	11	18	11	11	12.40
g	11	11	11	11	18	12.40

Assume that the institute is slightly more oriented towards statistics and probability and suppose that the DM considers that there are 3 groups of subjects: statistics and probability, economics and management, and English. Furthermore, he/she considers that within the two first groups, subjects are somewhat substitutive, i.e. they overlap to a certain extent. Finally, if a student is good in statistics or probability (resp. bad in statistics and probability), it is better that he/she is good in English (resp. economics or management) rather than in economics or management (resp. English). This reasoning, applied to the profiles of Table 2, leads to the following ranking:

$$a \succ_{\mathcal{O}} b \succ_{\mathcal{O}} c \succ_{\mathcal{O}} d \succ_{\mathcal{O}} e \succ_{\mathcal{O}} f \succ_{\mathcal{O}} g. \tag{9}$$

We shall further assume that the DM considers that two students are significantly different if their overall utilities differ by at least half a unit.

By considering students a and b, and f and g, it is easy to see that the criteria do not satisfy mutual preferential independence, which implies that there is no additive model that can numerically represent the above weak order.

In order to use the identification methods reviewed in the previous section and implemented in Kappalab, we first create 7 R vectors representing the students:

```
> a <- c(18,11,11,11,18)
> b <- c(18,11,18,11,11)
> c <- c(11,11,18,11,18)
> d <- c(18,18,11,11,11)
> e <- c(11,11,18,18,11)
> f <- c(11,11,18,11,11)
> g <- c(11,11,11,11,18)
```

The symbol `>` represents the prompt in the R shell, the symbol `<-` the assignment operator, and `c` is the R function for vector creation.

5.2 The least squares approaches

In order to apply the least squares approaches presented in Subsection 4.1, the 7 vectors previously defined and representing the students need first to be concatenated into a 7 row matrix, called `C` here, using the `rbind` ("row bind") matrix creation function:

```
> C <- rbind(a,b,c,d,e,f,g)
```

Then, the DM needs to provide overall utilities for the seven students. Although it is unrealistic to consider that this information can always be given, we assume in this subsection that the DM is able to provide it. He respectively assigns 15, 14.5, 14, 13.5, 13, 12.5 and 12 to a, b, c, d, e, f and g. These desired overall utilities are encoded into a 7 element R vector:

```
> overall <- c(15,14.5,14,13.5,13,12.5,12)
```

The least squares identification routine based on quadratic programming (providing an optimal but not necessarily unique solution) can then be called by typing:

```
> ls <- least.squares.capa.ident(5,2,C,overall)
```

in the R terminal. The first argument sets the number of criteria, the second fixes the desired order of k-additivity, and the two last represent the matrix containing the partial utilities and the vector containing the desired overall utilities respectively. The result is stored in an R list object, called here `ls`, containing all the relevant information for analyzing the results.

The solution, a 2-additive capacity given under the form of its Möbius representation, can be obtained by typing:

```
> m <- ls$solution
```

and visualized by entering `m` in the R terminal:

```
> m
                Mobius.capacity
{}              0.000000
{1}             0.311650
{2}             0.176033
...             ...
{4,5}           0.001752
```

As discussed in Subsection 4.1, for the considered example, the obtained solution is probably not unique [30].

The Choquet integral for instance of a w.r.t. the solution can be obtained by typing:

```
> Choquet.integral(m,a)
[1] 15
```

To use the least squares identification routine implementing the heuristic approach proposed in [32], we first need to create the *initial capacity* as discussed in Subsection 4.1. Here, in the absence of clear requirements on the form of the Choquet integral, we take the uniform capacity on the set of criteria (subjects):

```
> mu.unif <- as.capacity(uniform.capacity(5))
```

The heuristic least squares identification routine can then be called by typing:

```
> hls <- heuristic.ls.capa.ident(5,mu.unif,C,overall,alpha=0.05)
```

The first argument sets the number of criteria, the second contains the initial capacity, the third represents the matrix of the partial evaluations, the fourth the vector containing the desired overall utilities, and the last the value of the parameter controlling the gradient descent.

The overall utilities computed using the Choquet integral w.r.t. the two obtained solutions are given in the last two columns of the table below, the sixth column containing the desired overall evaluations:

```
   S  P  E  M En Given Mean   LS  HLS
a 18 11 11 11 18  15.0 13.8 15.0 15.0
b 18 11 18 11 11  14.5 13.8 14.5 14.5
c 11 11 18 11 18  14.0 13.8 14.0 14.0
d 18 18 11 11 11  13.5 13.8 13.5 13.5
e 11 11 18 18 11  13.0 13.8 13.0 13.0
f 11 11 18 11 11  12.5 12.4 12.5 12.5
g 11 11 11 11 18  12.0 12.4 12.0 12.0
```

As one can see, both the optimal and the heuristic approaches enable to recover the overall utilities provided by the DM. Recall however that the solution returned by the optimal quadratic method is 2-additive whereas that returned by the heuristic method is 5-additive.

The Shapley values of the solutions can be computed by means of the `Shapley.value` function taking as argument a capacity and are given in the following table :

```
       S    P    E    M   En
LS  0.29 0.14 0.21 0.13 0.24
HLS 0.24 0.18 0.20 0.16 0.21
```

As one could have expected, the Shapley value of the solution obtained by the heuristic approach is more even than that returned by the optimal quadratic method.

5.3 The LP, minimum variance and minimum distance approaches

As discussed earlier, the least squares approaches applied in the previous subsection are not well adapted to MAUT since they rely on information that a DM cannot always provide. The LP, the minimum variance and the minimum distance approaches require only a partial weak order over the available objects, such as the one provided by the DM in Eq. (9). This weak order is naturally translated as

$$C_{m_\nu}(a) > C_{m_\nu}(b) > C_{m_\nu}(c) > C_{m_\nu}(d) > C_{m_\nu}(e) > C_{m_\nu}(f) > C_{m_\nu}(g).$$

with indifference threshold $\delta_C = 0.5$.

Practically, the preference threshold is stored in an R variable:

```
> delta.C <- 0.5
```

and the weak order over the students is encoded into a 6 row R matrix:

```
> Acp <- rbind(c(a,b,delta.C), c(b,c,delta.C), c(c,d,delta.C),
               c(d,e,delta.C), c(e,f,delta.C), c(f,g,delta.C))
```

each row containing a constraint of the form $C_{m_\nu}(u(x)) \geq C_{m_\nu}(u(y)) + \delta_C$.

The LP approach is then invoked by typing:

```
> lp <- lin.prog.capa.ident(5,2,A.Choquet.preorder = Acp)
```

The first argument fixes the number of criteria, the second sets the desired order of k-additivity for the solution, and the last contains the partial weak order provided by the DM. All the relevant information to analyze the solution is stored in the R object `lp`.

The minimum variance approach is called similarly:

```
> mv <- mini.var.capa.ident(5,2,A.Choquet.preorder = Acp)
```

To use the minimum distance approach, we first need to create the *initial capacity*. In the absence of clear requirements from the DM, we choose the uniform capacity on the set of criteria (subjects), which can be created by entering:

```
> m.mu <- additive.capacity(c(0.2,0.2,0.2,0.2,0.2))
```

The closest capacity to the uniform capacity compatible with the initial preferences of the DM is then obtained by typing:

```
> md <- mini.dist.capa.ident(m.mu,2,"global.scores",
                             A.Choquet.preorder = Acp)
```

The second argument sets the desired order of k-additivity for the solution, while the third one indicates which of the 3 available quadratic distances between capacities should be used [36]. The character string `"global.scores"` refers to the distance given in Eq. (8).

The overall utilities computed using the Choquet integral w.r.t. the 2-additive solutions are given in the following table:

```
   S  P  E  M En Mean    LP    MV    MD
a 18 11 11 11 18 13.8 18.00 15.25 14.95
b 18 11 18 11 11 13.8 17.36 14.75 14.45
c 11 11 18 11 18 13.8 16.73 14.25 13.95
d 18 18 11 11 11 13.8 16.09 13.75 13.45
e 11 11 18 18 11 13.8 15.45 13.25 12.95
f 11 11 18 11 11 12.4 14.82 12.75 12.45
g 11 11 11 11 18 12.4 14.18 12.25 11.95
```

Note that, as expected, the LP approach leads to more dispersed utilities, which reach the maximum value (18) that a Choquet integral can take for the seven students. Note also that, for the minimum variance and the minimum distance approaches, the differences between the overall utilities of two consecutive students in the weak order provided by the DM equal exactly δ_C. This last observation follows from the fact that, in this example, the aim of both methods is roughly to find the Choquet integral that is the closest to the simple arithmetic mean while being in accordance with the preferential information provided by the DM.

The Shapley values of the 2-additive solutions are:

```
      S    P    E    M   En
LP 0.45 0.00 0.27 0.05 0.23
MV 0.27 0.16 0.21 0.14 0.22
MD 0.24 0.18 0.20 0.16 0.22
```

As one can see, all three solutions designate statistics (S) as the most important criterion. Note that the LP solution is very extreme, since the overall importance of probability (P) and management (M) is very small and that of S is close to one half.

However, the overall importances of the criteria are not in accordance with the orientation of the institution. Indeed, one would have expected to obtain that statistics (S) and probability (P), and economics (E) and management (M), have the same importances. This is due to the fact that until now, the preferential information which we used was limited to a small number of students which were ranked by the DM. In order to build a more accurate model, we can impose additional constraints as we shall see in the next subsection. This clearly justifies the use of a progressive interactive approach to model the DM's preferences in MAUT.

5.4 Additional constraints on the Shapley value

As discussed in the previous subsection, assume now that by considering the Shapley values of the 2-additive solutions obtained above, the DM explicitly requires that statistics (S) and probability (P), and economics (E) and management (M), have the same overall importances, i.e. $S \sim_N P$ and $E \sim_N M$.

These additional constraints are translated as $-\delta_\phi \leq \phi_{m_\nu}(S) - \phi_{m_\nu}(P) \leq \delta_\phi$ and $-\delta_\phi \leq \phi_{m_\nu}(E) - \phi_{m_\nu}(M) \leq \delta_\phi$, where the indifference threshold δ_ϕ is supposed to have been set to 0.01 by the DM. To encode them, an R variable representing the indifference threshold is first created:

```
> delta.phi <- 0.01
```

The inequalities discussed above are then encoded into a 4 row R matrix:

```
> Asp <- rbind(c(1,2,-delta.phi), c(2,1,-delta.phi),
               c(3,4,-delta.phi), c(4,3,-delta.phi))
```

each row corresponding to a constraint of the form $\phi_{m_\nu}(i) - \phi_{m_\nu}(j) \geq c$, $c \in [0, 1]$.

The LP approach is then invoked by typing

```
> lp2 <- lin.prog.capa.ident(5,2,A.Choquet.preorder = Acp,
                             A.Shapley.preorder = Asp)
```

into the R terminal. The minimum variance and minimum distance routines are called similarly.

The Shapley values of the 2-additive solutions are:

```
      S    P    E    M   En
LP 0.23 0.23 0.18 0.18 0.18
MV 0.22 0.21 0.18 0.17 0.22
MD 0.22 0.21 0.18 0.17 0.22
```

As expected, the solutions satisfy the constraints additionally imposed by the DM.

The overall utilities computed using the Choquet integral w.r.t. the 2-additive solutions are given in the following table:

```
   S  P  E  M En Mean    LP    MV    MD
a 18 11 11 11 18 13.8 16.03 15.12 14.84
b 18 11 18 11 11 13.8 15.52 14.62 14.34
c 11 11 18 11 18 13.8 15.01 14.12 13.84
d 18 18 11 11 11 13.8 14.50 13.62 13.34
e 11 11 18 18 11 13.8 13.99 13.12 12.84
f 18 18 11 11 11 12.4 13.48 12.62 12.34
g 11 11 18 11 11 12.4 12.97 12.12 11.84
```

This time, the three approaches give more similar overall utilities. This was to be expected as the problem is more constrained.

The interaction indices of the 2-additive capacities obtained by means of the LP, minimum variance and minimum distance approaches are respectively given in the three tables below:

```
[LP]   S     P     E     M    En
S     NA -0.27 -0.17  0.00 -0.03
P  -0.27    NA  0.00  0.16 -0.04
E  -0.17  0.00    NA -0.12 -0.06
M   0.00  0.16 -0.12    NA -0.07
En -0.03 -0.04 -0.06 -0.07    NA
```

```
[MV]   S     P     E     M    En
S     NA -0.21 -0.05 -0.06  0.10
P  -0.21    NA  0.01  0.15 -0.03
E  -0.05  0.01    NA -0.12  0.05
M  -0.06  0.15 -0.12    NA -0.01
En  0.10 -0.03  0.05 -0.01    NA
```

```
[MD]   S     P     E     M    En
S     NA -0.21 -0.04 -0.07  0.10
P  -0.21    NA  0.03  0.18 -0.01
E  -0.04  0.03    NA -0.10  0.09
M  -0.07  0.18 -0.10    NA  0.00
En  0.10 -0.01  0.09  0.00    NA
```

As one can notice, statistics (S) negatively interacts with almost all the subjects, which again is not in accordance with the orientation of the institution. Indeed, one would expect statistics (S) to be complementary with all subjects except probability (P). Once more, in the perspective of a progressive and interactive approach, this can be corrected by imposing additional constraints on the interaction indices as we shall see in the next subsection.

5.5 Additional constraints on the interaction indices

Assume finally that, in order to be in accordance with the orientation of the institution, the DM imposes that subjects within the same group[2] have to interact in a substitutive way, whereas two subjects from different groups have to interact in a complementary way. This additional preferential information is translated by means of the following constraints:

P	E	M	En	
$-1 \leq I_{m_\nu}(SP) \leq -\delta_I$	$\delta_I \leq I_{m_\nu}(SE) \leq 1$	$\delta_I \leq I_{m_\nu}(SM) \leq 1$	$\delta_I \leq I_{m_\nu}(SEn) \leq 1$	S
	$\delta_I \leq I_{m_\nu}(PE) \leq 1$	$\delta_I \leq I_{m_\nu}(PM) \leq 1$	$\delta_I \leq I_{m_\nu}(PEn) \leq 1$	P
		$-1 \leq I_{m_\nu}(EM) \leq -\delta_I$	$\delta_I \leq I_{m_\nu}(EEn) \leq 1$	E
			$\delta_I \leq I_{m_\nu}(MEn) \leq 1$	M

where δ_I, supposed set to 0.05, is a DM defined threshold to be interpreted as the minimal absolute value of an interaction index to be considered as significantly different from zero.

To encode this additional preferential information, an R variable representing the threshold is first created:

```
> delta.I <- 0.05
```

[2]Recall that the three groups of subjects are $\{S, P\}$, $\{E, M\}$, and $\{En\}$.

The constraints discussed above are then encoded into a 10 row R matrix:

```
>  Aii <- rbind(c(1,2,-1,-delta.I), c(1,3,delta.I,1), c(1,4,delta.I,1),
                c(1,5,delta.I,1), c(2,3,delta.I,1), c(2,4,delta.I,1),
                c(2,5,delta.I,1), c(3,4,-1,-delta.I), c(3,5,delta.I,1),
                c(4,5,delta.I,1))
```

each row corresponding to a constraint of the form $a \leq I_{m_\nu}(ij) \leq b$, $a, b \in [-1, 1]$.

There are no 2-additive capacities compatible with these additional constraints. The order of k-additivity is then incremented and the LP approach is invoked by typing:

```
> lp3 <- lin.prog.capa.ident(5,3, A.Choquet.preorder = Acp,
                               A.Shapley.preorder = Asp,
                               A.interaction.interval = Aii)
```

The minimum variance and minimum distance routines are called similarly.

The Shapley values and the interaction indices of the three 3-additive solutions are given in the four following tables:

	S	P	E	M	En
LP	0.23	0.23	0.16	0.16	0.22
MV	0.23	0.22	0.18	0.18	0.20
MD	0.22	0.21	0.18	0.19	0.21

[LP]	S	P	E	M	En
S	NA	-0.30	0.05	0.05	0.12
P	-0.30	NA	0.07	0.14	0.05
E	0.05	0.07	NA	-0.24	0.05
M	0.05	0.14	-0.24	NA	0.05
En	0.12	0.05	0.05	0.05	NA

[MV]	S	P	E	M	En
S	NA	-0.13	0.05	0.05	0.05
P	-0.13	NA	0.05	0.05	0.05
E	0.05	0.05	NA	-0.05	0.05
M	0.05	0.05	-0.05	NA	0.05
En	0.05	0.05	0.05	0.05	NA

[MD]	S	P	E	M	En
S	NA	-0.21	0.05	0.05	0.05
P	-0.21	NA	0.05	0.05	0.05
E	0.05	0.05	NA	-0.12	0.05
M	0.05	0.05	-0.12	NA	0.05
En	0.05	0.05	0.05	0.05	NA

As expected, the constraints additionally imposed by the DM are satisfied.

The overall utilities computed using the Choquet integral w.r.t. the 3-additive solutions are given in the following table:

	S	P	E	M	En	Mean	LP	MV	MD
a	18	11	11	11	18	13.8	14.06	14.26	14.45
b	18	11	18	11	11	13.8	13.55	13.76	13.95
c	11	11	18	11	18	13.8	13.04	13.26	13.45
d	18	18	11	11	11	13.8	12.53	12.76	12.95
e	11	11	18	18	11	13.8	12.02	12.26	12.45
f	18	18	11	11	11	12.4	11.51	11.76	11.95
g	11	11	18	11	11	12.4	11.00	11.26	11.45

We hereby conclude the process of the modeling of the DM's preferences. In the following section we imagine a scenario where the DM considers that the 3-additive solutions above are too complex and where he/she prefers to have a simpler description of his preferences. In such a case, the DM has to take into account that some of his preferences will be violated.

5.6 A simpler solution

Assume that for the sake of simplicity the DM absolutely wants a 2-additive solution for the problem described in the previous subsections. In that case, it is possible to use the generalized least squares approach described in Subsection 4.4 to obtain an approximate solution.

First of all, the weak order over the students has to be encoded into a 6 row R matrix:

```
> rk.proto <- rbind(c(1,2), c(2,3), c(3,4), c(4,5), c(5,6), c(6,7))
```

The integers correspond to the line indices of the alternatives a, b, c, d, e, f and g in the matrix `C` defined in Subsection 5.2.

The generalized least squares approach can then be called by typing:

```
> gls <- ls.ranking.capa.ident(5, 2, C, rk.proto, 0.5,
                               A.Shapley.preorder = Asp,
                               A.interaction.interval = Aii)
```

The first argument sets the number of criteria, the second the desired order of k-additivity, the third the matrix containing the partial evaluations, the fourth the matrix containing the weak order and the fifth argument is the value of the threshold δ_y. The two last arguments contain the matrices encoding the additional constraints on the Shapley value and on the interaction indices respectively.

Although we know from the previous subsections that there are no 2-additive capacities compatible with the imposed constraints, this approach provides a solution with a non zero objective function as we could have expected. The following table gives the aggregated overall utilities $\{C_{m_\nu}(u(x))\}_{x\in\mathcal{O}}$ in the last column and the overall utilities $\{y(x)\}_{x\in\mathcal{O}}$ in the last but one column:

	S	P	E	M	En	Mean	y	GLS
a	18	11	11	11	18	13.8	13.94	13.67
b	18	11	18	11	11	13.8	13.44	13.44
c	11	11	18	11	18	13.8	12.94	12.81
d	18	18	11	11	11	13.8	12.44	12.57
e	11	11	18	18	11	13.8	11.94	11.81
f	18	18	11	11	11	12.4	11.44	11.57
g	11	11	18	11	11	12.4	10.94	11.21

As one can see, the ranking provided by the DM is not violated but the minimal threshold δ_y is not always respected (for example $C_{m_\nu}(f) - C_{m_\nu}(g) < \delta_y$). The Shapley value and the interaction indices of this 2-additive solution are:

S	P	E	M	En
0.23	0.22	0.17	0.16	0.22

	S	P	E	M	En
S	NA	-0.22	0.05	0.05	0.14
P	-0.22	NA	0.06	0.09	0.06
E	0.05	0.06	NA	-0.08	0.15
M	0.05	0.09	-0.08	NA	0.05
En	0.14	0.06	0.15	0.05	NA

which as expected satisfy the constraints imposed by the DM. It is now up to the DM to evaluate if the violation of δ_y does not deteriorate significantly the overall quality of the model.

References

[1] R. L. Keeney and H. Raiffa. *Decision with multiple objectives*. Wiley, New-York, 1976.

[2] D. Bouyssou and M. Pirlot. Conjoint measurement tools for MCDM: A brief introduction. In J. Figuera, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 73–132. Springer, 2005.

[3] D.H. Krantz, R.D. Luce, P. Suppes, and A. Tversky. *Foundations of measurement, volume 1: Additive and polynomial representations*. Academic Press, 1971.

[4] D. Bouyssou and M. Pirlot. Preferences for multi-attributed alternatives: Traces, dominance, and numerical representations. *J. of Mathematical Psychology*, 48:167–185, 2004.

[5] P. Vincke. *Multicriteria Decision-aid*. Wiley, 1992.

[6] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953.

[7] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, Tokyo, Japan, 1974.

[8] M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89:445–456, 1992.

[9] J.-L. Marichal. An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria. *IEEE Transactions on Fuzzy Systems*, 8(6):800–807, 2000.

[10] C. Labreuche and M. Grabisch. The Choquet integral for the aggregation of interval scales in multicriteria decision making. *Fuzzy Sets and Systems*, 137:11–16, 2003.

[11] M. Grabisch, Ch. Labreuche, and J.C. Vansnick. On the extension of pseudo-Boolean functions for the aggregation of interacting criteria. *European Journal of Operational Research*, 148:28–47, 2003.

[12] C.A. Bana e Costa and J.C. Vansnick. Preference relations in MCDM. In T. Gal and T. Hanne T. Steward, editors, *MultiCriteria Decision Making: Advances in MCDM models, algorithms, theory and applications*. Kluwer, 1999.

[13] M. Grabisch and C. Labreuche. Fuzzy measures and integrals in MCDA. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis*, pages 563–608. Springer, 2004.

[14] M. Grabisch, I. Kojadinovic, and P. Meyer. `kappalab`*: Non additive measure and integral manipulation functions*, 2006. R package version 0.3.

[15] R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2005. ISBN 3-900051-00-3.

[16] T. Murofushi and M. Sugeno. Fuzzy measures and fuzzy integrals. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals: Theory and Applications*, pages 3–41. Physica-Verlag, 2000.

[17] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

[18] G-C. Rota. On the foundations of combinatorial theory. I. Theory of Möbius functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 2:340–368 (1964), 1964.

[19] A. Chateauneuf and J-Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17(3):263–283, 1989.

[20] J.-L. Marichal. Behavioral analysis of aggregation in multicriteria decision aid. In J. Fodor, B. De Baets, and P. Perny, editors, *Preferences and Decisions under Incomplete Knowledge*, pages 153–178. Physica-Verlag, 2000.

[21] J-L. Marichal. Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. *European Journal of Operational Research*, 155(3):771–791, 2004.

[22] 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.

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

[24] T. Murofushi and S. Soneda. Techniques for reading fuzzy measures (iii): Interaction index. In *9th Fuzzy System Symposium*, pages 693–696, Saporo, Japan, 1993.

[25] M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92(2):167–189, 1997.

[26] K. Fujimoto, I. Kojadinovic, and J-L. Marichal. Axiomatic characterizations of probabilistic and cardinal-probabilistic interaction indices. *Games and Economic Behavior*, 55:72–99, 2006.

[27] P. Wakker. *Additive Representations of Preferences: A New Foundation of Decision Analysis*. Kluwer Academic Publishers, Dordrecht, 1989.

[28] T. Mori and T. Murofushi. An analysis of evaluation model using fuzzy measure and the Choquet integral. In *5th Fuzzy System Symposium*, pages 207–212, Kobe, Japan, 1989. In Japanese.

[29] M. Grabisch, H.T. Nguyen, and E.A. Walker. *Fundamentals of uncertainty calculi with applications to fuzzy inference*. Kluwer Academic, Dordrecht, 1995.

[30] P. Miranda and M. Grabisch. Optimization issues for fuzzy measures. *Int. J. of Uncertainty, Fuzziness, and Knowledge Based Systems*, 7:545–560, 1999.

[31] K. Ishii and M. Sugeno. A model of human evaluation process using fuzzy measure. *Int. J. Man-Machine Studies*, 67:242–257, 1996.

[32] M. Grabisch. A new algorithm for identyfing fuzzy measures and its application to pattern recognition. In *Int. 4th IEEE Conf. on Fuzzy Systems*, pages 145–150, Yokohama, Japan, March 1995.

[33] J-L. Marichal and M. Roubens. Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research*, 124:641–650, 2000.

[34] I. Kojadinovic. Minimum variance capacity identification. *European Journal of Operational Research*, 177(1):498–514, 2007.

[35] 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.

[36] I. Kojadinovic. Quadratic distances for capacity and bi-capacity approximation and identification. *4OR: A Quarterly Journal of Operations Research*, page in press, 2006.

[37] J-L. Marichal. *Aggregation operators for multicriteria decison aid*. PhD thesis, University of Liège, Liège, Belgium, 1998.

[38] P. Meyer and M. Roubens. Choice, ranking and sorting in fuzzy Multiple Criteria Decision Aid. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 471–506. Springer, New York, 2005.

[39] Michel Berkelaar et al. `lpSolve`*: Interface to LPSolve v.5 to solve linear/integer programs*, 2005. R package version 1.1.9.

[40] B.A. Turlach and A. Weingessel. `quadprog`*: Functions to solve quadratic programming problems.*, 2004. R package version 1.4-7.

[41] Alexandros Karatzoglou, Alex Smola, Kurt Hornik, and Achim Zeileis. `kernlab`: An S4 package for kernel methods in R. *Journal of Statistical Software*, 11(9):1–20, 2004.