Identification of a fuzzy measure with an L^1 entropy

Christophe Labreuche
Thales Research & Technology
RD128
91767 Palaiseau Cedex, France
christophe.labreuche@thalesgroup.com

Abstract

The aim of the paper is to propose a *linearized* entropy of a fuzzy measure. This entropy is used as an optimization functional for the identification of a fuzzy measure from learning data. The problem of identifying a fuzzy measure can thus be transformed into a linear program while keeping the good properties of the entropy.

Keywords: Choquet integral, fuzzy measure, entropy, parameter identification.

1 Introduction

The Choquet integral appears as a useful aggregation function in many fields such as multi-criteria decision aid and data mining. An important practical problem is the identification of the parameters of the Choquet integral [1, 9]. Classically, the identification of these parameters is performed from learning data. This preferential information can classically be of three types: a comparison of two profiles, an overall score associated to a profile, or less intuitive preferential information related to the importance or interaction of the attributes. The information of the second and the third types are not always easy to obtain from a decision maker, especially in multi-criteria decision aid. Moreover, as argued in [2], the main information obtained from empirical data of the second type is a requirement stating that the data set shall preserve the order deduced from the overall scores, which corresponds to the first type of information. We focus thus on the first type of information in this paper.

There exist several learning methods based on the first type of preferential information [14, 8]. Yet, in all cases, the preferential information is considered as constraints to be satisfied. Any method consists thus in determining the value of the parameters that fulfills previous constraints and that maximizes some functional. In [14], the functional to be maximized is the difference between the overall scores of the two profiles in each pair belonging to the learning data. In [8], the functional is the entropy or the variance of the parameters. The main asset of the first method is that it leads to linear programming, which can be easily solved. The second method yields non-linear optimization. However, the maximization of the entropy or the variance tends to find the value of the parameters that is the less *specific*, which is very attractive in the case of scarse empirical data. The goal of this paper is to define a functional that has the good properties of the entropy but still leads to linear programming.

The prerequisites on the Choquet integral are given in Section 2. Two models are described: the general Choquet integral and the two-additive Choquet integral. Section 3 defines the identification problem and shows the classical approaches to solve it. A *linearized* entropy is proposed in Section 4 for the general Choquet integral. The case of the two-

L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (eds): Proceedings of IPMU'08, pp. 1476–1483
Torremolinos (Málaga), June 22–27, 2008

additive Choquet integral is dealt with in Section 5.

2 Background on fuzzy measures and Choquet integral

We denote by $N = \{1, \dots, n\}$ the set of criteria or attributes. The alternatives are supposed to be described by a vector in $[0,1]^n$, where the i^{th} coordinate depicts the satisfaction degree w.r.t. attribute i. Considering two alternatives $x, y \in [0,1]^n$ and $A \subseteq N$, we use the notation $(x_A, y_{N\setminus A})$ to denote the compound alternative $z \in [0,1]^n$ such that $z_i = x_i$ if $i \in A$ and y_i otherwise.

2.1 Fuzzy measures and Choquet integral

The Choquet integral is a generalization of the commonly used weighted sum. It is based on the concept of fuzzy measures [16].

Definition 1 *A* fuzzy measure μ *on* N *is a function* $\mu : 2^N \longrightarrow [0,1]$, *satisfying the following axioms.*

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

(ii) $A \subseteq B \subseteq N$ *implies* $\mu(A) \leq \mu(B)$.

We denote by $\mathcal{M}$ the set of fuzzy measures on N. Additive measures are particular cases of fuzzy measures when $\mu(A \cup B) = \mu(A) + \mu(B)$ for any pair of disjoint sets A,B. An additive measure μ is described by a probability p: $\mu(A) = \sum_{i \in A} p_i$ for $A \subseteq N$. The Möbius transform m of a fuzzy measure μ is defined by

$$m(A) := \sum_{B \subseteq A} (-1)^{|A|-|B|} \mu(B).$$

Definition 2 *Let* μ *be a fuzzy measure on* N. *The* discrete Choquet integral *of a vector* $x \in [0,1]^N$ *with respect to* μ *is defined by*

$$C_\mu(x) = \sum_{i=1}^{n} x_{\tau(i)} \Big[\mu(\{\tau(i), \dots, \tau(n)\}) - \mu(\{\tau(i+1), \dots, \tau(n)\}) \Big]$$

where τ *is a permutation satisfying* $x_{\tau(1)} \leq \dots \leq x_{\tau(n)}$.

The main drawback of the fuzzy measure is that it contains much more parameters than a simple weighted sum - namely 2^n parameters for a fuzzy measure vs. n parameters for a weighted sum. The use of sub-classes of fuzzy measures containing much less than 2^n parameters is crutial in practice.

An interesting sub-class of fuzzy measures are the so-called 2-additive fuzzy measures. These are fuzzy measures for which the interactions between more than two criteria vanish, or equivalently for which all Möbius coefficients for coalitions of more than two criteria are zero. The importance of criteria and the interaction between pairs of criteria can be modeled with 2-additive fuzzy measures. The parameters of a 2-additive fuzzy measure are the importances (between 0 and 1) of the criteria and the interactions (between -1 and 1) of the pairs of criteria. We denote by v_i^μ the importance of criterion i and $I_{i,j}^\mu$ the interaction between the criteria i and j. The expression of the Choquet integral with respect to a 2-additive fuzzy measure has the following form :

$$C_\mu^2(x) = \sum_{i \in N} v_i^\mu \, x_i - \sum_{\{i,j\} \subseteq N} I_{i,j}^\mu \frac{|x_i - x_j|}{2} .$$

The importances of the criteria yield a weighted sum whereas the interactions provide a non-linear expression proportional to the distance between the values of the criteria. The parameters v^μ and I^μ must satisfy the following constraints

$$\sum_{i \in N} v_i^\mu = 1 ,$$

$$\forall i \in N \ \forall \varepsilon_1, \dots, \varepsilon_n \in \{-1, 1\}$$

$$v_i + \frac{1}{2} \sum_{j \in N \setminus \{i\}} \varepsilon_j \times I_{i,j} \geq 0$$

The set of two-additive fuzzy measures on N is denoted by $\mathcal{M}^2$.

2.2 Interpretation of the fuzzy measures

According to the following fundamental property

$$C_\mu(1_A, 0_{N\setminus A}) = \mu(A) \qquad (1)$$

the parameter $\mu(A)$ corresponds to an alternative that is perfectly satisfactory on attributes A and completely unsatisfactory on the remaining criteria.

Concerning the 2-additive Choquet integral, relation (1) is also true. Moreover, one has

$$C^2_\mu(1_A, 0_{N\setminus A}) = \sum_{i\in A} v^\mu_i - \frac{1}{2} \sum_{i\in A, j\in N\setminus A} I^\mu_{i,j}. \quad (2)$$

One has

$$C^2_\mu(1_i, 0.5_{N\setminus i}) = v^\mu_i + \sum_{k\in N\setminus i} \frac{v^\mu_k}{2} - \frac{1}{2} \sum_{k\in N\setminus i} I^\mu_{i,k}$$

$$C^2_\mu(0_i, 0.5_{N\setminus i}) = \sum_{k\in N\setminus i} \frac{v^\mu_k}{2} - \frac{1}{2} \sum_{k\in N\setminus i} I^\mu_{i,k}$$

Therefore, $C^2_\mu(1_i, 0.5_{N\setminus i}) - C^2_\mu(0_i, 0.5_{N\setminus i}) = v^\mu_i$. The importance index v^μ_i corresponds to the difference of global score between the two profiles $(1_i, 0.5_{N\setminus i})$ and $(0_i, 0.5_{N\setminus i})$. The importance of attribute i is thus the added-value on the overall score when improving the score on attribute i from 0 to 1, in the situation where the other attributes have the mean score 0.5.

¿From relation (2), one has

$$C^2_\mu(1_{\{i,j\}}, 0_{N\setminus\{i,j\}}) - C^2_\mu(1_i, 0_{N\setminus i}) - C^2_\mu(1_j, 0_{N\setminus j}) = I^\mu_{i,j}$$

Hence the interaction index $I^\mu_{i,j}$ between attribute i and j measures the difference between being good at both attributes i and j together, and being good at attributes i and j separately, in the situation where the other attributes are very ill-satisfied.

3 Description of the identification problem

3.1 Statement of the problem

The determination of the aggregation function is carried out through a learning phase on the basis of some preferential information (learning data). These data are composed of a set

$$P := \{(x^k, y^k) \ : \ k \in \{1, \ldots, m\}\}$$

of pairs of profiles in $[0,1]^n$, where, for each pair (x^k, y^k), the first profile x^k is judged at least as good as the second one y^k by the decision maker. When the decision maker provides these comparisons, he has no doubt on them so that they must be considered as hard constraints [14]. We are thus lead to considering the set of fuzzy measures that fulfill these constraints:

$$\mathcal{M}(P) := \Big\{ \mu \in \mathcal{M} \ : \ \forall k \in \{1, \ldots, m\} \quad C_\mu(x^k) \geq C_\mu(y^k) \Big\}.$$

For the two-additive Choquet integral, one obtains

$$\mathcal{M}^2(P) := \Big\{ \mu \in \mathcal{M}^2 \ : \ \forall k \in \{1, \ldots, m\} \quad C^2_\mu(x^k) \geq C^2_\mu(y^k) \Big\}$$

The two sets $\mathcal{M}(P)$ and $\mathcal{M}^2(P)$ are clearly polytopes since all constraints on the fuzzy measure are linear.

When the decision maker says that profile x^k is at least as good as y^k, he probably means that x^k is significantly preferred to y^k. It seems then reasonable to try to maximize the difference between the scores obtained by x^k and y^k [14]

$$\max \ \varepsilon \quad \text{under} \quad \begin{cases} \varepsilon \geq 0 \\ \mu \in \mathcal{M}(P) \\ \forall i \in \{1, \ldots, m\} \, , \, C_\mu(x^k) \geq C_\mu(y^k) + \varepsilon \end{cases}$$

This problem is a linear problem. A similar problem can be obtained for two-additive fuzzy measures.

3.2 Entropy of a fuzzy measure

We consider here the case when the preferential information is scarse. We make the analogy with the management of uncertainty. Consider a random experiment having n possible outcomes with probabilities $p_1, \ldots, p_n$. In order to appraise the average uncertainty associated with the prediction of these outcomes, or equivalently, the amount of information received from the knowledge of which these outcomes occurred, several information measures have been introduced. Among them, the best known is probably the Shanon entropy [15]

$$H_S(p) := \sum_{i=1}^{n} h(p_i) \ ,$$

where $h(u) = -u \ \ln u$ with the convention $0 \ln 0 := 0$.

For assessing a probability distribution, the *maximum entropy* principle is often used. More precisely, in the case of partial knowledge on the probabilities (this might be the case when one knows the probability distribution on a coarser frame of discernment), the probability distribution that is chosen is the one that maximizes the entropy. A nice property of this maximization is that one gets always a unique solution thanks to the convexity of the entropy functional.

The problem we aim to solve here is quite similar, since we want to determine the fuzzy measure from preferential information that does not fully specify it. The idea would thus be to compute the fuzzy measure satisfying to the information provided by the DM that maximizes the entropy. One has then to define the notion of entropy for fuzzy measures.

There remains to define the notion of entropy for fuzzy measures. Let $\mathcal{Q}(N)$ denote the set of maximal chains of the Hasse diagram $\left(2^N, \subseteq\right)$. Recall that a maximal chain in $\left(2^N, \subseteq\right)$ is a sequence $\emptyset, \{\tau(1)\}, \{\tau(1), \tau(2)\}, \ldots, \{\tau(1), \ldots, \tau(n)\}$ where τ is a permutation of N. On each maximal chain, one can define the following probability distribution: $p^\tau_{\tau(i)} = \mu(\{\tau(i), \ldots, \tau(n)\}) - \mu(\{\tau(i+1), \ldots, \tau(n)\})$. The entropy of μ is then defined as the mean value of the entropy of each maximal chain [13] :

$$\begin{aligned} \hat{H}_S(\mu) &:= \frac{1}{n!} \sum_{\tau} H_S(p^\tau) \\ &= \sum_{i=1}^{n} \sum_{A \subseteq N \setminus \{i\}} \gamma^n_{|A|} h(\delta_i \mu(A)) \ , \end{aligned}$$

where $\gamma^n_p = \frac{(n-p-1)!p!}{n!}$ and $\delta_i \mu(A) := \mu(A \cup \{i\}) - \mu(A)$. This leads to solve the following problem

$$\max_{\mu \in \mathcal{M}(P)} \hat{H}_S(\mu) \ \text{ or } \ \max_{\mu \in \mathcal{M}^2(P)} \hat{H}_S(\mu). \qquad (3)$$

This problem is non-linear and even non quadratic, which is not so easy to solve. One can obtain a quadratic problem under linear constraints, considering a special case of Rényi entropy. This amounts to take function $h(u) = -u^2$ [8] :

$$\hat{H}_R(\mu) := -\sum_{i=1}^{n} \sum_{A \subseteq N \setminus \{i\}} \gamma^n_{|A|} (\delta_i \mu(A))^2 \, .$$

We obtain

$$\max_{\mu \in \mathcal{M}(P)} \hat{H}_R(\mu) \ \text{ or } \ \max_{\mu \in \mathcal{M}^2(P)} \hat{H}_R(\mu). \qquad (4)$$

A unique solution of (3) and (4) is attained with previous two entropies.

4 Towards a *linearized* entropy for a general fuzzy measure

In practice, it may be very interesting to rely only on linear programming. The question that arises is the following one: does there exist a version of the entropy that fits into linear programming?

4.1 Definition of an L^1 entropy

The first idea is to consider function $f(u) = -|u|$. This leads to the following definition

$$\hat{H}_l(\mu) := -\sum_{i=1}^{n} \sum_{A \subseteq N \setminus \{i\}} \gamma^n_{|A|} \left|\delta_i \mu(A)\right| .$$

Going back to the interpretation of the entropy of a fuzzy measure in terms of chains, we notice that

$$\hat{H}_l(\mu) = \frac{1}{n!} \sum_{\tau} H_l(p^\tau)$$

where

$$H_l(p^\tau) := -\sum_{i=1}^{n} \left| p^\tau_{\tau(i)} \right| .$$

By monotonicity of the fuzzy measure, we obtain

$$\begin{aligned} H_l(p^\tau) &= -\sum_{i=1}^{n} \Big(\mu(\{\tau(i), \ldots, \tau(n)\}) \\ &\qquad - \mu(\{\tau(i+1), \ldots, \tau(n)\}) \Big) \\ &= \mu(\emptyset) - \mu(N) = -1 \end{aligned}$$

and thus

$$\hat{H}_l(\mu) = -1. \tag{5}$$

This shows that previous definition is not satisfactory.

An entropy H of a probability distribution on N shall satisfy several fundamental properties:

- Symmetry: $H(\sigma \circ p) = H(p)$ for any permutation σ on N, where $(\sigma \circ p)_i = p_{\sigma(i)}$;
- Maximality: H takes its maximal value at the uniform probability measure $p = (1/n, \ldots, 1/n)$;
- Minimality: H takes its minimal value at the extreme probability measure $p = (0, \ldots, 0, 1, 0, \ldots, 0)$;

¿From (5), the maximality and minimality properties are clearly not satisfied with H_l. In order to make sure to satisfy the minimality property, an option is to define an entropy on probability distributions as the L^1 error between the probability and the uniform probability:

$$H_L(p) := -\sum_{i=1}^{n} \left| p_i - \frac{1}{n} \right| .$$

This leads to the following definition of the L^1 entropy on fuzzy measures:

$$\begin{aligned} \hat{H}_L(\mu) &= \frac{1}{n!} \sum_{\tau} H_L(p^\tau) \\ &= -\sum_{i=1}^{n} \sum_{A \subseteq N \setminus \{i\}} \gamma^n_{|A|} \left| \delta_i \mu(A) - \frac{1}{n} \right| . \end{aligned}$$

¿From this expression, one easily shows that the three properties Symmetry, Maximality and Minimality are fulfilled. The fuzzy measure can then be obtained by solving the following optimization problem.

$$\max_{\mu \in \mathcal{M}(P)} \hat{H}_L(\mu). \tag{6}$$

4.2 Properties of the L^1 entropy

We introduce the following problem $\mathcal{P}_1$:

Find $s := \left\{ \left\{ s^+_{A,i}, s^-_{A,i} \right\} \ : \ i \in N \ , \ A \subseteq N \setminus \{i\} \right\}$ such that

$$\min \ F_1(s) := \sum_{i=1}^{n} \sum_{A \subseteq N \setminus \{i\}} \gamma_{|A|}(n) \Big(s^+_{A,i} + s^-_{A,i} \Big)$$

under

$$\left\{ \begin{array}{l} \mu \in \mathcal{M}(P) \\ \forall i \in N \ , \ \forall A \subseteq N \setminus \{i\} \quad s^+_{A,i} \geq 0 \\ \qquad \text{and } \delta_i \mu(A) - \frac{1}{n} \leq s^+_{A,i} \\ \forall i \in N \ , \ \forall A \subseteq N \setminus \{i\} \quad s^-_{A,i} \geq 0 \\ \qquad \text{and } \delta_i \mu(A) - \frac{1}{n} \geq -s^-_{A,i} \end{array} \right. \tag{7}$$

Problem $\mathcal{P}_1$ is a linear program.

Let

$$\begin{aligned} \mathcal{M}_1(s) = \Big\{ & \mu \in \mathcal{M} \ : \ \forall i \in N \ , \ \forall A \subseteq N \setminus \{i\} \\ & - s^-_{A,i} \leq \delta_i \mu(A) - \frac{1}{n} \leq s^+_{A,i} \Big\} \end{aligned}$$

$\mathcal{M}_1(s)$ corresponds to the inequalities in (7). Set

$$\Sigma_1 = \left\{ s \in \mathbb{R}_+^{(n\, 2^n)} \ : \ \mathcal{M}_1(s) \neq \emptyset \right\} \ .$$

We define the preference relation $\succeq_1$ on $\mathbb{R}_+^{(n\, 2^n)}$ by

$$\begin{aligned} s \succeq_1 s' \iff & \forall i \in N \ , \ \forall A \subseteq N \setminus \{i\} \\ & s^+_{A,i} \geq s'^+_{A,i} \text{ and } s^-_{A,i} \geq s'^-_{A,i} \end{aligned}$$

Let $\succ_1$ by the asymmetric part of $\succeq_1$. We have $s \succ_1 s'$ iff for all $i \in N$ and all $A \subseteq N \setminus \{i\}$ $s_{A,i} \geq s'_{A,i}$ and if there exists $i \in N$ and $A \subseteq N \setminus \{i\}$ such that $s^+_{A,i} > s'^+_{A,i}$ or $s^-_{A,i} > s'^-_{A,i}$.

Lemma 1 *Let $s \in \Sigma_1$ non dominated in Σ_1 in the sense of $\prec_1$ (i.e. there does not exist $s' \in \Sigma_1$ such that $s' \prec_1 s$). Then for all $i \in N$*

and $A \subseteq N \setminus \{i\}$ one has either $s^+_{A,i} = 0$ or $s^-_{A,i} = 0$. Moreover

$$\mathcal{M}_1(s) = \Big\{\mu \in \mathcal{M} \ : \ \forall i \in N \,,\, \forall A \subseteq N \setminus \{i\}$$
$$\delta_i\mu(A) - \frac{1}{n} = s^+_{A,i} \quad \textit{if } s^-_{A,i} = 0$$
$$\delta_i\mu(A) - \frac{1}{n} = -s^-_{A,i} \quad \textit{if } s^+_{A,i} = 0\Big\}$$

The next lemma shows that $\mathcal{M}_1(s)$ reduces to a singleton for a non-dominated s.

Lemma 2 *Let $s \in \Sigma_1$ non dominated in Σ_1 in the sense of $\prec_1$. Then $\mathcal{M}_1(s)$ reduces to a singleton.*

We now study the relationship between $\mathcal{P}_1$ and (6).

Lemma 3 *The solution(s) to $\mathcal{P}_1$ is(are) non-dominated in the sense of $\prec_1$.*

We have the following result.

Lemma 4 *The two problems (6) and $\mathcal{P}_1$ are equivalent.*

Of course, we cannot conclude that there is a unique solution to $\mathcal{P}_1$. However, there is a unique solution for each Pareto vector s.

Example 1 *Let μ_1 and μ_2 be defined by for all $A \subseteq N$:*

$$\mu_1(A) = \begin{cases} \frac{|A|}{n} & \textit{if } |A| \neq 1 \\ \frac{1}{n} - \varepsilon & \textit{if } |A| = 1 \end{cases}$$
$$\mu_2(A) = \begin{cases} \frac{|A|}{n} & \textit{if } |A| \neq 1 \\ \frac{1}{n} + \varepsilon & \textit{if } |A| = 1 \end{cases}$$

Let $i \in N$ and $A \subseteq N \setminus \{i\}$. If $|A| = 0$, then

$$\mu_1(A \cup \{i\}) - \mu_1(A) - \frac{1}{n} = -\varepsilon$$
$$\mu_2(A \cup \{i\}) - \mu_2(A) - \frac{1}{n} = \varepsilon$$

If $|A| = 1$, then

$$\mu_1(A \cup \{i\}) - \mu_1(A) - \frac{1}{n} = \varepsilon$$
$$\mu_2(A \cup \{i\}) - \mu_2(A) - \frac{1}{n} = -\varepsilon$$

If $|A| > 1$, then

$$\mu_1(A \cup \{i\}) - \mu_1(A) - \frac{1}{n} = 0$$
$$\mu_2(A \cup \{i\}) - \mu_2(A) - \frac{1}{n} = 0$$

The two fuzzy measures μ_1 and μ_2 lead to the following two vectors s^1 and s^2 respectively :

$$s^{1,+}_{A,i} = \begin{cases} \varepsilon & \textit{if } |A| = 1 \\ 0 & \textit{otherwise} \end{cases}$$
$$s^{1,-}_{A,i} = \begin{cases} \varepsilon & \textit{if } |A| = 0 \\ 0 & \textit{otherwise} \end{cases}$$

and

$$s^{2,+}_{A,i} = \begin{cases} \varepsilon & \textit{if } |A| = 0 \\ 0 & \textit{otherwise} \end{cases}$$
$$s^{2,-}_{A,i} = \begin{cases} \varepsilon & \textit{if } |A| = 1 \\ 0 & \textit{otherwise} \end{cases}$$

One has $F_1(s^1) = F_1(s^2)$. Problem $\mathcal{P}_1$ may lead to several solutions, namely s^1 and s^2. However, there is a unique fuzzy measure associated to each solution (thanks to the combining of Lemmas 3 and 2).

We conclude from the results of this section that solving $\mathcal{P}_1$ is equivalent to solving (6). Moreover, $\mathcal{P}_1$ is a linear problem. So, problem $\mathcal{P}_1$ is an interesting problem to solve (6).

5 The two-additive case

We wish to define a *linearized* entropy for a two-additive fuzzy measure. ¿From (1) and (2), we have

$$\delta_i\mu(A) = v_i - \frac{1}{2}\sum_{k\in A} I^\mu_{i,k} + \frac{1}{2}\sum_{k\in N\setminus(A\cup\{i\})} I^\mu_{i,k}$$

Plugging this in the expression of $\hat{H}_L(\mu)$, we obtain a complex expression of the entropy. We want to define a simplier expression of a *linearized* entropy. To this end, our starting point if the Maximality property, stating that the entropy shall take its maximal value for the arithmetic mean. Since the arithmetic mean is described by the following fuzzy measure

$$v^\mu_1 = \cdots = v^\mu_n = \frac{1}{n} \ , \ I^\mu_{1,2} = \cdots = I^\mu_{n-1,n} = 0$$

a possible definition of an approximate piece-wise linear entropy in the case of a two-additive measure can be :

$$\hat{H}^2_L(\mu) = -\sum_{i=1}^n \left| v_i^\mu - \frac{1}{n} \right| - \frac{1}{2} \sum_{\{i,j\}\subseteq N} \left| I^\mu_{i,j} \right| \ .$$

The symmetry and maximality properties are clearly satisfied with this definition.

For $\{i,j\} \subseteq N$, $\mu^{\min}_{i,j}$ (resp. $\mu^{\max}_{i,j}$) is the capacity such that $v_i = v_j = \frac{1}{2}$, $I_{i,j} = 1$ (resp. -1), $v_k = 0$ for all $k \in N \setminus \{i,j\}$ and $I_{k,l} = 0$ for all $\{k,l\} \subseteq N$ with $\{k,l\} \neq \{i,j\}$. The Choquet integral w.r.t. this capacity is the minimum (resp. maximum) between the scores of criteria i and j. For $i \in N$, μ^d_i is such that $v_i = 1$, $v_k = 0$ for all $k \in N \setminus \{i\}$, and $I_{k,l} = 0$ for all $\{k,l\} \subseteq N$. The Choquet integral w.r.t. this capacity is the dictator for criterion i.

Lemma 5 *We have the following result:*

- *If $n \in \{2,3\}$, then the smallest value of $\hat{H}^2_L$ is attained at μ^d_i for all $i \in N$;*
- *If $n = 4$, then the smallest value of $\hat{H}^2_L$ is attained at μ^d_i for all $i \in N$, and also at $\mu^{\min}_{i,j}$ and $\mu^{\max}_{i,j}$ for all $\{i,j\} \subseteq N$;*
- *If $n \geq 5$, the smallest value of $\hat{H}^2_L$ is attained at $\mu^{\min}_{i,j}$ and $\mu^{\max}_{i,j}$ for all $\{i,j\} \subseteq N$.*

The optimization problem becomes:

$$\max_{\mu \in \mathcal{M}^2(P)} \hat{H}_L(\mu) \tag{8}$$

We introduce the following problem $\mathcal{P}_2$

Find $s := \left\{ \left\{ s^+_A, s^-_A \right\} : \ A \subseteq N \text{ with } |A| \in \{1,2\} \right\}$ such that

$$\min \ F_2(s) := \sum_{i=1}^n \left(s^+_{\{i\}} + s^-_{\{i\}} \right) + \\ + \frac{1}{2} \sum_{\{i,j\}\subseteq N} \left(s^+_{\{i,j\}} + s^-_{\{i,j\}} \right) \quad \text{under}$$

$$\begin{cases} \mu \in \mathcal{M}^2(P) \\ \forall i \in N \ , \quad s^+_{\{i\}}, s^-_{\{i\}} \geq 0 \\ \qquad \text{and } - s^-_{\{i\}} \leq v^\mu_i - \frac{1}{n} \leq s^+_{\{i\}} \\ \forall \{i,j\} \subseteq N \ , \quad s^+_{\{i,j\}}, s^-_{\{i,j\}} \geq 0 \\ \qquad \text{and } - s^-_{\{i,j\}} \leq I^\mu_{i,j} \leq s^+_{\{i,j\}} \end{cases} \tag{9}$$

Let

$$\mathcal{M}_2(s) = \\ \Big\{ \mu \in \mathcal{M}^2 \ : \ \forall i \in N \ - s^-_{\{i\}} \leq v^\mu_i - \frac{1}{n} \leq s^+_{\{i\}} \\ \text{and } \forall \{i,j\} \subseteq N \ - s^-_{\{i,j\}} \leq I^\mu_{i,j} \leq s^+_{\{i,j\}} \Big\}$$

and

$$\Sigma_2 = \left\{ s \in \mathbb{R}_+^{n(n+1)/2} \ : \ \mathcal{M}_2(s) \neq \emptyset \right\} \ .$$

We define the preference relation $\succeq_2$ on $\mathbb{R}_+^{n(n+1)/2}$ by

$$s \succeq_2 s' \iff \forall A \subseteq N \text{ with } |A| \in \{1,2\} \\ s^+_A \geq s'^+_A \text{ and } s^-_A \geq s'^-_A$$

Let $\succ_2$ by the asymmetric part of $\succeq_2$. We have $s \succ_2 s'$ iff $s \succeq s'$ and if there exists $A \subseteq N$ with $|A| \in \{1,2\}$ such that $s^+_A > s'^+_A$ or $s^-_A > s'^-_A$.

Lemma 6 *Let $s \in \Sigma_2$ non dominated in Σ_2 in the sense of $\prec_2$. Then for all $A \subseteq N$ with $|A| \in \{1,2\}$, one has either $s^+_A = 0$ or $s^-_A = 0$. Moreover*

$$\mathcal{M}_2(s) = \Big\{ \mu \in \mathcal{M}^2 \ : \\ \forall i \in N \ , \quad v^\mu_i - \frac{1}{n} = s^+_{\{i\}} \quad \text{if } s^-_{\{i\}} = 0 \\ \forall i \in N \ , \quad v^\mu_i - \frac{1}{n} = -s^-_{\{i\}} \quad \text{if } s^+_{\{i\}} = 0 \\ \forall \{i,j\} \subseteq N \ , \quad I^\mu_{i,j} = s^+_{\{i,j\}} \quad \text{if } s^-_{\{i,j\}} = 0 \\ \forall \{i,j\} \subseteq N \ , \quad I^\mu_{i,j} = -s^-_{\{i,j\}} \quad \text{if } s^+_{\{i,j\}} = 0 \Big\}$$

Lemma 7 *Let $s \in \Sigma_2$ non dominated in Σ_2 in the sense of $\prec_2$. Then $\mathcal{M}_2(s)$ reduces to a singleton.*

Lemma 8 *The solution(s) to $\mathcal{P}_2$ is(are) non-dominated in the sense of $\prec_2$.*

Lemma 9 *The two problems (8) and $\mathcal{P}_2$ are equivalent.*

Hence, solving $\mathcal{P}_2$ is equivalent to solving (8). Moreover, $\mathcal{P}_2$ is a linear problem.

References

[1] G. Beliakov. How to build aggregation operators from data? Int. J. Intelligent Systems 18 (2003) 903–923.

[2] G. Beliakov, T. Calvo. Constructions of aggregation operatorsthat preserve ordering of the data. EUSFLAT conference, pp. 61–66, Sept 11-14 2007, Ostrava.

[3] G. Choquet. Theory of capacities. Annales de l'Institut Fourier 5 (1953) 131–295.

[4] M. Grabisch, The application of fuzzy integrals in multicriteria decision making, European Journal of Operational Research 89 (1996) 445-456.

[5] M. Grabisch, T. Murofushi and M. Sugeno, Fuzzy measures and integrals, Physica-Verlag, Heidelberg, New York, 2000.

[6] R.L. Keeney and H. Raiffa. Decision with Multiple Objectives. Wiley, New York, 1976.

[7] I. Kojadinovic. Quadratic distances for capacity and bi-capacity approximation and identification. Quaterly J. of Operations Research (4OR) 2 (2007) 117–142.

[8] I. Kojadinovic. Minimum variance capacity identification. Eur. J. of Operational Research 12 (2006) 23–36.

[9] M. Grabisch, I. Kojadinovic, P. Meyer. A review of capacity identification methods for Choquet integral based multi-attribute utility theory — Applications of the Kappalab R package. Eur. J. of Operational Research, to appear.

[10] M. Grabisch, Ch. Labreuche. Fuzzy Measures and Integrals in MCDA. In: Multiple Criteria Decision Analysis - State of the Art Surveys (J. Figueira, S. Greco, M. Ehrgott Eds). Springers International Series, pp 563-608, 2005.

[11] M. Grabisch and Ch. Labreuche. A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. Quarterly Journal of Operations Research (4OR), accepted.

[12] D.H. Krantz, R.D. Luce, P. Suppes and A. Tversky. Foundations of measurement, vol 1: Additive and Polynomial Representations. Academic Press, 1971.

[13] J.L. Marichal. Entropy of discrete Choquet capacities. European J. of Oper. Rech. 137 (2002) 612–624.

[14] J.L. Marichal, M. Roubens. Dependence between criteria and multiple criteria decision aid. 2nd Int. Workshop on Preferences and Decisions. Trento, Italy, pp. 69–75, 1998.

[15] C.E. Shannon, A mathematical theory of communication. Bell System Tech. J. 27 (1948) 379–423.

[16] M. Sugeno. Theory of fuzzy integrals and its applications. Tokyo Institute of Technology, 1974.