Fuzzy Measures and Integrals

Linett Montaño Guzmán

Universidad Católica Boliviana
Departamento de Ciencias Exactas
Cochabamba, Bolivia
e-mail: lmontano@ucbcba.edu.bo

Abstract

In this paper, we will study the aggregation problem with interacting criteria, and we will introduce the concepts of fuzzy measures and integrals. The fuzzy integrals are powerful aggregation operators which are able to take into consideration the interaction among criteria and the fuzzy measures can be used for modeling the importance of all subsets of criteria.

We will concentrate in “the Choquet integral”, this fuzzy integral is a weighted aggregation operator, which takes into account not only the importance of the criteria, but also the importance of all subsets of them. The information about the importance and the interaction between criteria is used in the aggregation process of the partial evaluations. This paper analyse some characteristics and properties of this operator.

keywords: Multiple criteria analysis; Fuzzy measures; aggregation operator.

1 Introduction

Multicriteria Decision Aid (MCDA) is an important branch of Operational Research. MCDA discipline develops and implements decision support tools and methodologies to deal with complex decision problems that involve multiple criteria, goals or objectives of conflicting nature. The tools and methodologies provided by MCDA are not just mathematical models aggregating criteria, points of view or attributes but furthermore they are decision support oriented. Most Multicriteria Decision Aid (MCDA) methods need somewhere in their procedure a fundamental operation: “the aggregation”. The most common aggregation tool used today, is still the weighted arithmetic mean. This operator has a necessary condition for the representation of preferences: “the independence”. But, in many situations, however, the criteria considered are not independent, they interact.

In fact, in a decision problem there are links between criteria. For instance, there will be a very strong relation between the quality of a product and its cost. This type

Acta Nova; Vol. 2, N°1, diciembre 2002 · 216
of relation should be taken into account in aggregation methods and especially in the process of determining the weights. Thus, the fuzzy measures and integrals take into account this information.

In the following sections, we will introduce some important notions of the theory of fuzzy measures and integration. For more details and proofs about this theory, the reader is referred to the following papers and books [1, 2, 3, 6, 7, 8, 10].

2 Interaction among criteria

In a decision problem the evaluation of the alternatives under various points of view (criteria) often shows interactions. These links express the correlation or the dependence with respect to the preferences among criteria. This does not mean that the corresponding points of view are redundant and should be eliminated. On the opposite, this means, that the attributes that are used to reflect these points of view are linked by logical or factual interdependencies and this information should be used. We will analyse two types of interaction among criteria: the correlation and the preferential dependence.

2.1 Correlation

The word correlation is used in everyday life to denote some form of association.

The statistical definition of the notion of correlation is the following:

**Definition 2.1** The correlation measures the inter-relationship between two variables $X$ and $Y$. The output of this measurement is called the “correlation coefficient”, denoted by $r$. It is sometimes called Pearson’s correlation coefficient after its originator and is a measure of linear association.

\[
    r = \frac{n \sum_{i=1}^{n} X_i Y_i - \left( \sum_{i=1}^{n} X_i \right) \left( \sum_{i=1}^{n} Y_i \right)}{\sqrt{n \sum_{i=1}^{n} X_i^2 - \left( \sum_{i=1}^{n} X_i \right)^2} \sqrt{n \sum_{i=1}^{n} Y_i^2 - \left( \sum_{i=1}^{n} Y_i \right)^2}}
\]  

(1)

The correlation coefficient is measured on a scale that varies from $-1$ through 0 to $+1$. Complete correlation between two variables is expressed by either $+1$ or $-1$. When one variable increases as the other increases the correlation is positive; when one decreases as the other increases it is negative. Complete absence of correlation is represented by 0.

In statistical terms the notion of correlation is used to denote an association between two quantitative variables. It is also assumed that the association is linear, that is, that one variable increases or decreases a fixed amount for a unit increase or decrease of the other.
The correlation among criteria is probably the best known and most intuitive type of dependence. For instance, in the selection of personnel problem, the age of the candidate and the number of years of professional experience, criteria used in the evaluation, are likely to be positively correlated.

2.2 Preferential dependence

Another way to have dependence or interaction among criteria is when the evaluation of the alternatives follow different points of view, do not respect the property of preferential independence.

A simple test, to prove the independence in the sense of preferences, consists in asking the decision maker about his preferences, on pairs of alternatives, that share the same profile, for a subset of attributes; varying the common profile should not reverse the preferences when the points of view are independent.

Let us consider a set of alternatives $A = \{a, b, \ldots, m\}$ and a set of criteria $F = \{g_1, g_2, \ldots, g_n\}$. Each alternative $a \in A$ is associated with a profile $g(a) = (g_1(a), g_2(a), \ldots, g_n(a)) \in \mathbb{R}^n$ where $g_j(a)$ represents the utility of $a$ related to the criterion $j$.

Suppose that the preferences over $A$ of the decision maker are known and expressed by a binary relation $\succeq$.

**Definition 2.2** (Vincke 1992 [14]) Let $F$ be the family of criteria, $S$ a subset of $F$ and $\overline{S}$ the complementary subset in $F$. $S$ is preferentially independent in $F$ if, given four actions $a, b, c, d$ such that:

\[
\begin{align*}
g_j(a) &= g_j(b) & \forall j \in \overline{S} \\
g_j(c) &= g_j(d) & \forall j \in \overline{S} \\
g_j(a) &= g_j(c) & \forall j \in S \\
g_j(b) &= g_j(d) & \forall j \in S
\end{align*}
\]

we get

$$aPb \Leftrightarrow cPd,$$

where $P$ is the global preference relation taking into account all the criteria.

In other words, $S$ is preferentially independent in $F$ if the preferences between actions, which differ only by their evaluations according to the criteria in $S$, do not depend upon the values yielded by the criteria of $\overline{S}$.

**Example 1** Consider the selection of personnel problem: A company is hiring and has a short list of 4 candidates: Anne, Bernard, Céline and David. The HR department has a list of three criteria: age of the candidate ($g_1$), number of years of professional experience in a relevant related job ($g_2$) and number of years of higher education ($g_3$). The evaluation involving 4 candidates is described in the Table 1.
Table 1: Selection of personnel problem.

<table>
<thead>
<tr>
<th>Candidate</th>
<th>( g_1 )</th>
<th>( g_2 )</th>
<th>( g_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne ((a))</td>
<td>28</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Bernard ((b))</td>
<td>28</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Céline ((c))</td>
<td>35</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>David ((d))</td>
<td>35</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

The HR department expresses its opinion about the candidates. Evidently, the preferences \( a \succeq c \) and \( b \succeq d \) are immediately suggested; but the other comparisons are not so obvious because the associated profiles interlace. After the analysis of the problem, the HR department expresses the following preferences:

\[
\text{candidate } a \succeq \text{ candidate } b \text{ and candidate } d \succeq \text{ candidate } c
\]

For the HR department, the reason is that as the age increases, the years of professional experience become more important than the years of higher education. However, when the candidate is younger, it is more important to have a high education than experience. In this case the professional experience and the number of years of higher education are not preferentially independent of the age of the candidate. Thus, the first two criteria are not preferentially independent of the third. Therefore, in this case, there is no additive measure that could reproduce the HR department ranking.

A well known example of dependence in the sense of preferences is the preference for white wine or red wine depending on whether you are eating fish or meat.

In order to take into account the behavior of criteria: the correlation and dependence in the sense of preferences and to have a flexible representation of these interactions between criteria, it is useful to introduce the concept of fuzzy measures and integrals. Thus, in the next sections we introduce these main concepts, and we show how fuzzy measures provide a means of representing the relationship between criteria.

3 Fuzzy measures

The additivity property of the “measure”, one of the most important concepts in mathematics, is often inflexible or too rigid to represent the many faces of human reasoning. Therefore, to be able to express human subjectivity, Sugeno [13] proposed to replace the additivity property by a weaker one, the monotonicity, and he called these non-additive monotonic measures “fuzzy measures”.

Before the introduction of fuzzy measures by Sugeno in 1974 (to generalize additive measures), this concept has been introduced in 1953 by Choquet [1] as “capacities”.

Thus, the fuzzy measures are also called non-additive measures, or capacities.

**Definition 3.1** A discrete “fuzzy measure” on \( N \) is a set function \( \mu: 2^N \rightarrow [0,1] \) satisfying the following conditions:
i) \( \mu(\emptyset) = 0 \), \( \mu(N) = 1 \)

ii) \( S \subset T \subset N \implies \mu(S) \leq \mu(T) \)

A fuzzy measure is said to be:

i) Additive if \( \mu(S \cup T) = \mu(S) + \mu(T) \) whenever \( S \cap T = \emptyset \).

ii) Super-additive if \( \mu(S \cup T) \geq \mu(S) + \mu(T) \) whenever \( S \cap T = \emptyset \).

iii) Sub-additive if \( \mu(S \cup T) \leq \mu(S) + \mu(T) \) whenever \( S \cap T = \emptyset \).

To understand the meaning of fuzzy measures in practical applications we give an example proposed by Murfushi and Sugeno [12]:

**Example 2** Let \( N \) be a set of workers in a workshop, and suppose they produce the same products. Suppose that a group of workers \( A \subset N \) works in the most efficient way. Let \( \mu(A) \) be the number of products made by \( A \) in one hour. \( \mu \) can be considered as a measure of productivity, and can be normalized dividing by \( \mu(N) \). Then, the normalized function of the set function \( \mu : 2^N \to \mathbb{R}_+ \) is monotone and vanishes at the empty set, and therefore it is a fuzzy measure.

Observe that in this situation, \( \mu \) may be non-additive, since two groups of workers \( A \) and \( B \) together can produce more (or less) than if they work separately. If \( A \) and \( B \) work separately, then \( \mu(A \cup B) = \mu(A) + \mu(B) \). Nevertheless, generally if they work jointly, they interact and the equality may not be kept. Thus, an effective cooperation of members of \( A \cup B \) gives the inequality \( \mu(A \cup B) > \mu(A) + \mu(B) \). On the other hand, the incompatibility between the groups of workers \( A \) and \( B \) could give the opposite inequality \( \mu(A \cup B) < \mu(A) + \mu(B) \).

In a decision aid framework, the fuzzy measure \( \mu(T) \) represents the weight of importance of the subset of criteria \( T \). Therefore, in addition to the usual weights on criteria taken into account separately, weights on all the combinations of criteria will be defined as well.

In the case where the fuzzy measure is additive, we can consider that there is no interaction between criteria and it suffices to define \( n \) weights \( \{\mu(1), \ldots, \mu(n)\} \) to define the measure entirely. When the fuzzy measure is super-additive, this means that there is a positive interaction between criteria (synergy). Otherwise, if the fuzzy measures is sub-additive it means that there is a negative interaction between criteria (redundancy).

There are several representations of a fuzzy measure. According to Grabisch [4] a representation of a fuzzy measure \( \mu \) could be defined by “any set function from which it is possible to recover \( \mu \) without loss of information”. The most utilized representations are the following:

- The usual (measure) representation; this is simply \( \mu \) itself.
- The Möbius representation or polynomial representation \( (w) \).
- Interaction representation or Shapley value \( (I) \).
4 Fuzzy integral

"Fuzzy integral" is a generic term for integrals with respect to fuzzy measures. There are many kinds of fuzzy integrals: The integrals of Choquet, Šipoš, Sugeno, t-conorm integral, etc. In this paper we will mainly concentrate on the Choquet integral.

Murofushi and Sugeno proposed the so-called Choquet fuzzy integral, referring to a function defined by Choquet in a different context. The Choquet integral is a fuzzy integral based on a fuzzy measure that provides alternative computational scheme for aggregating information.

In fact, this aggregation operator, in a decision aid context, takes into account, in the evaluation process, not only the importance of each criteria, but also the importance of all subsets. The weights are represented by the coefficients of the fuzzy measure.

**Definition 4.1** Let, $\mu$ be a fuzzy measure on $N$. The discrete Choquet integral of a function $g : N \rightarrow \mathbb{R}$ with respect to $\mu$ is defined by:

$$C_\mu(g_1, \ldots, g_n) = \sum_{i=1}^{n}(g_i - g_{i-1})\mu(A_i)$$

(3)

where $(\cdot)$ indicates that the indices have been permuted, such that $\{0 \leq g_1 \leq g_2 \ldots \leq g_n\}$; $A_i = \{(i), \ldots, (n)\}$ and $g_0 = 0$.

![Figure 1: The Choquet integral.](image)

For instance, if $g_0 \leq g_1 \leq g_2$, then $g_1 = g_3$, $g_2 = g_1$ and $g_3 = g_2$, the fuzzy measures $\mu(A_{11}) = \mu(312)$, $\mu(A_{12}) = \mu(12)$ and $\mu(A_{33}) = \mu(2)$ and the fuzzy integral will be:

$$C_\mu(g_1, g_2, g_3) = (g_3 - g_0)\mu(3,1,2) + (g_1 - g_2)\mu(1,2) + (g_2 - g_3)\mu(2)$$

In Figure 1 we can see the representation of the Choquet integral.
5 Properties of the Choquet integral

In order to show the suitability of fuzzy integrals in the multicriteria decision aid framework, we present in this section the main properties for aggregation of this operator.

The Choquet integral is a monotonic increasing function $C_μ$ defined from $[0,1]^n$ in $[0,1]$ and has the following properties:

i) **Boundary conditions:** This property expresses the behavior of the aggregation operator in the worst and in the best case.

$$C_μ(0, \ldots, 0) = 0 \quad C_μ(1, \ldots, 1) = 1$$

ii) **Idempotence:** For all $g_j$ identical ($= a$), this aggregation operator restitutes the common value ($a$).

$$C_μ(a, \ldots, a) = a$$

iii) **Continuity:** This property means that the Choquet integral is continuous with respect to each of its variables. It means that, this operator does not present any chaotic reaction to a small change in the arguments.

iv) **Monotonicity:** This operator is non-decreasing with respect to each variable.

$$g_i \geq g_j \implies C_μ(g_1, \ldots, g_i, \ldots, g_n) \geq C_μ(g_1, \ldots, g_j, \ldots, g_n)$$

v) **Decomposability:** This property means that any subset of elements $g \in \mathbb{R}^n$ can be replaced by their partial aggregation without changing the global aggregation.

$$C_μ^{(n)}(g_1, \ldots, g_k, g_{k+1}, \ldots, g_n) = C_μ^{(n)}(g_1, \ldots, g_k, g_{k+1}, \ldots, g_n)$$

where $g = C_μ^{(k)}(g_i, \ldots, g_k)$

vi) **Stability under the same positive linear transformations:** This property translates a stability of the operator for a change of measurement scale.

$$C_μ(rg_1 + t, \ldots, rg_n + t) = rC_μ(g_1, \ldots, g_n) + t \quad \forall r > 0, \forall t \in \mathbb{R}.$$  

From this property, we can say that changing the scale does not change the result. This is an extremely important property in utility theory, given that the evaluations of an action according to each criterion are defined with respect to a positive linear transformation. The global utility of the action have thus to keep this property.

From boundary conditions i), idempotency ii) and monotonicity iv) properties, the following lemma can be deduced.
Lemma 1  The Choquet integrals are comprised between the minimum and the maximum value.

\[
\min\{g_1, \ldots, g_n\} \leq C_\mu (g) \leq \max\{g_1, \ldots, g_n\}
\]

Proof: From boundary conditions i), idempotency ii) and monotonicity iv) properties we have:

For \( g = \{g_1, \ldots, g_n\} \)

Let \( g_f = \min\{g\} \) be the minimum value of \( g \), and \( g_k = \max\{g\} \) be the maximum value of \( g \)

\[
g_f \leq g_i \quad \forall i \in \{1, \ldots, n\}
\]

\[
g_k \geq g_i \quad \forall i \in \{1, \ldots, n\}
\]

\[
\{g_f, \ldots, g_f\} \leq \{g_1, \ldots, g_n\} \leq \{g_k, \ldots, g_k\}
\]

From the monotonicity property we have:

\[
C_\mu (g_f, \ldots, g_f) \leq C_\mu (g_1, \ldots, g_n) \leq C_\mu (g_k, \ldots, g_k)
\]

From the idempotency property

\[
g_f \leq C_\mu (g) \leq g_k
\]

as \( g_f = \min(g) \) and \( g_k = \max(g) \) then

\[
\min(g) \leq C_\mu (g) \leq \max(g)
\]

From this lemma, we can say that this aggregation operator is a "compromise operator".

In this section we mentioned the properties of the fuzzy integrals related to the representation of interactions between criteria. Now, we can illustrate, what is understood by interactions and how they can be modelled by fuzzy integral, with a simple example.

Example 2  Consider the same example of the selection of personnel problem shown in section 2.2, with other candidates, A, B and C. The HR department has a list of three criteria: age of the candidate (\( g_1 \)), number of years of professional experience in a relevant related job (\( g_2 \)) and number of years of higher education (\( g_3 \)). The evaluation involving these candidates is described in the Table 2.

The HR department wants well equilibrated candidates without weak evaluations. This means, that the ideal candidate should be young, with various years of professional experience and with a high level of education.

The normalized matrix\(^1\) and the weight sum results are shown in Table 3. The ranking of this table is not fully satisfactory for the HR department, since the candidate

\(^1\) The criterion \( g_1 \) should be minimized, and \( g_2, g_3 \) maximized. We choose here to maximize the inverse values \( \frac{1}{g_1} \) and for the normalization, we divide \( g_i \) by the total sum, i.e \( g'_i = \frac{g_i}{\sum g_i} \).
<table>
<thead>
<tr>
<th>Candidate</th>
<th>g₁</th>
<th>g₂</th>
<th>g₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>28</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>36</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>29</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 2: Selection of personnel problem.

<table>
<thead>
<tr>
<th>Candidate</th>
<th>g₁</th>
<th>g₂</th>
<th>g₃</th>
<th>Global evaluation (weighted sum)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.36</td>
<td>0.50</td>
<td>0.20</td>
<td>0.36</td>
</tr>
<tr>
<td>B</td>
<td>0.28</td>
<td>0.14</td>
<td>0.47</td>
<td>0.29</td>
</tr>
<tr>
<td>C</td>
<td>0.35</td>
<td>0.36</td>
<td>0.33</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Table 3: Normalized matrix and weighted sum results.

A has a weakness in the number of years of higher education (see Table 2), but has been considered better than candidate C, who has no weak point. The reason is that too much importance is given to age and the number of years of professional experience of the candidates, which are redundant. Usually, the age and the professional experience have a strong link. Solving this problem with the Choquet integral and an appropriate fuzzy measure we have:

i) For the HR department, the age and the professional experience are more important than the number of years of higher education.

\[ \mu(\{g₁\}) = \mu(\{g₂\}) = 0.35 \]
\[ \mu(\{g₃\}) = 0.3 \]

ii) Because the criteria g₁ and g₂ are redundant, the weight attributed to the set \{g₁, g₂\} should be less than the sum of the weights of the pair of criteria.

\[ \mu(\{g₁, g₂\}) = 0.45 < 0.35 + 0.35 \]

iii) Since, the criteria g₁, g₃ and g₂, g₃ have a positive interaction, the weight attributed to the set \{g₁, g₂\} and \{g₂, g₃\} should be greater than the sum of individual weights.

\[ \mu(\{g₁, g₃\}) = 0.8 > 0.35 + 0.3 \]
\[ \mu(\{g₂, g₃\}) = 0.8 > 0.35 + 0.3 \]
And \[ \mu(\{g₁, g₂, g₃\}) = 1. \]

Table 4 shows the results applying this fuzzy measure to the candidates.

The HR can see that the Choquet integral gives the expected results.

This example exposes how easy it is to translate the requirements of the decision maker into coefficients of fuzzy measures. Nevertheless, when the decision maker con-
siders more than 3 criteria, the evaluation of the fuzzy measure becomes difficult. Is for this reason that the next section introduce the $k$-order additive fuzzy measure and integral.

6 $k$-order additive fuzzy measure and integral

In a decision aid problem involving $n$ criteria, if we take into account the interactions between criteria, we need to define the $2^n$ coefficients of the fuzzy measure $\mu$ corresponding to the $2^n$ subsets of $N$. The decision maker is not able to give such amount of information. In most cases, he will be able to guess the importance of each criterion and of each pair of criterion.

To overcome this problem, Grabisch [5] proposed to use the concept of $k$-order fuzzy measure.

**Definition 6.1** (Grabisch (97) [5]) A fuzzy measure $\mu$ defined on $N$ is said to be $k$-order additive if its corresponding pseudo-Boolean function is a multilinear polynomial of degree $k$, i.e. its Möbius transform $w(T) = 0$ for all $(T)$ such that $(T) > k$, and there exist at least one subset $T$ of $N$ of exactly $k$ elements such that $w(T) \neq 0$.

Thus, for the 2-additive fuzzy measures, coefficients of the Möbius representation are given by

$$\mu(i) = w(i), \quad i \in N,$$

$$\mu(ij) = w(i) + w(j) + w(ij), \quad \{i, j\} \in N$$

(4)

(5)

The monotonicity constraints and the normalisation of the coefficients of the 2-additive fuzzy measure are formulated as follows:

$$\begin{cases} 
\mu(\emptyset) = 0, \\
\sum_{\{i, j\} \subseteq N} \mu(ij) + (n - 2) \sum_{i \in N} \mu(i) = 1, \\
\mu(i) \geq 0, \quad \forall i \in N \\
\sum_{j \subseteq S} \mu(ij) - \sum_{j \subseteq S} \mu(j) - (n - 2) \mu(i) \geq 0, \quad \forall i \in N, \forall S \subseteq N \setminus i.
\end{cases}$$

(6)
In terms of the Möbius transform these constraints are:

\[
\begin{align*}
  w(\emptyset) &= 0, \\
  \sum_{i \in N} w(i) + \sum_{(i,j) \in N} w(i,j) &= 1, \\
  w(i) &\geq 0, \quad \forall i \in N \\
  w(i) + \sum_{j \in S} w(i,j) &\geq 0, \quad \forall i \in N, \forall S \subseteq N \setminus i.
\end{align*}
\] (7)

If we consider the 2-order model, the Choquet integral becomes:

\[
C_2(g) = \sum_{i \in N} w(i) g_i + \sum_{(i,j) \subseteq N} w(i,j) (g_i \wedge g_j) \quad \forall g \in \mathbb{R}^n
\] (8)

The 2-order case of the Choquet integral seems to be interesting in practical applications. It permits to model interaction between criteria while remaining very simple. Indeed, only \( n + (\binom{n}{2}) = (n(n+1))/2 \) coefficients are required to define the fuzzy measure. There are many applications based on the 2-order model \([1, 2, 3, 6, 7, 8, 9, 10]\).

A review of the literature \([3, 6, 7, 11]\), shows that there are mainly three approaches to find the fuzzy measures:

- Identification based on semantics
- Identification based on learning data
- Identification based on the combination of the semantics and learning data.

The first approach is based on the decision maker’s interpretation of the fuzzy measure. The second one is established on the assignment by the decision maker of the numerical score for each action and each criterion, and also a numerical global score for each action, from this information it is possible to build a learning system to find the fuzzy measure. Finally the third approach is based on the combination of the semantics and learning data, indeed this procedure combines semantical considerations and learning data which introduces very useful information, reduces the complexity and provides more efficient algorithms to find the fuzzy measure.

7 MCDA methods and Fuzzy measures and integrals

As we have shown in this paper, in many situations, the criteria considered in an MCDA evaluation process are not independent, they interact. In these situations the fuzzy integral becomes the appropriate aggregation operator to deal with this kind of problems. For this reason is very interesting the introduction of this powerful aggregation operator, “the Choquet integral”, in MCDA methods.
References


