Aalto University
School of Science
Degree programme in Engineering Physics and Mathematics

Discrete Choquet integral and multilinear forms

Bachelor's thesis
4.7.2014

Tiina Kangasniemi

AALTO UNIVERSITY
SCHOOL OF SCIENCE
PO Box 11000, FI-00076 AALTO
http://www.aalto.fi

ABSTRACT OF THE BACHELOR'S THESIS

Author: Tiina Kangasniemi

Title: Discrete Choquet integral and multilinear forms

Degree programme: Degree programme in Engineering Physics and Mathematics

Major subject: System sciences	Major subject code: F3010

Supervisor: Prof. Harri Ehtamo

Instructor: Ph.D. Matteo Brunelli

Abstract:

Multi-criteria decision making aims at comparing alternatives with respect to multiple criteria. To decide which criteria to use for evaluations is not always a straight-forward task. Single criteria values need to be aggregated into one value that represents the alternative, but the most commonly used aggregation functions, such as the weighted arithmetic mean, require that criteria must be independent.

The aim of this thesis is to provide a comprehensive introduction to two different functions that do not require criteria independence: the discrete Choquet Integral and Multilinear forms. With these functions, the alternatives can be evaluated even with non-independent criteria, and thus they provide capabilities to model the preferences of the decision maker in a more flexible way. In this study, the differences and similarities of the functions are compared on a qualitative level.

Date: 4.7.2014	Language: English	Number of pages: 36+3

Keywords: Aggregation function, Choquet integral, Multilinear extension, Multilinear form, Multi-criteria decision making

Sisältö

1 Introduction and motivation

Consider a situation of choosing between cars. You have certain qualities that are important to you, for example spaciousness, color, fuel consumption and acceleration. It is a fairly believable assumption that the value of color would be independent of all the other criteria, but are not the ecological cars usually smaller and have relatively lower acceleration than cars with larger fuel consumption? Using for example the weighted arithmetic mean to calculate comparable values for different cars might provide misleading preference order, since the effect of criteria interaction is neglected.

The objective of this thesis is to introduce two different functions that are able to take criteria interaction into account: the discrete Choquet Integral and Multilinear forms. Both of these functions provide capabilities to model the preferences of the decision maker in a more flexible way, when the evaluation criteria should not be assumed to be independent.

Let us next introduce the concept of aggregation functions. Let $X = X_1 \times ... \times X_n$ be a set of potential alternatives, where each alternative is described by a vector of n attributes (criteria), $x := (x_1, ..., x_n)$. The decision maker (later DM) is supposed to have a preference over X.

In decision theory there is a wide range of methods to find out what the preferences of the DM are and how the DM values alternatives compared to other alternatives. Quite often a lot of time and effort is put into finding good and reliable ways to express these preference relations and finding numerical values of attributes i (i.e. finding a value function $u_i(x_i)$). Mathematically different preference relations (preferable, less preferable and equally preferred) are presented with "$\succeq$", "$\preceq$" and "$\sim$" respectively. In order to actually utilize this information, we need an *aggregation function* F to aggregate attribute-specific values into one single numeric value to represent the DM's preference relations so that

$$x \succeq y \Leftrightarrow F(u_1(x_1), ..., u_n(x_n)) \geq F(u_1(y_1), ..., u_n(y_n)),$$

where $u_i : X_i \rightarrow S$, $i = 1, ..., n$ and S is a scale, often the closed interval [0,1]. To simplify the notation of aggregation functions, let us from now on denote the aggregated value for alternative x with $F(x) := F(u_1(x_1), ..., u_n(x_n))$.

Traditionally, effort put into this phase of decision making is smaller and usually only most basic aggregation functions are used. The most common aggregation function is the weighted arithmetic mean (WAM):

$$\mathrm{WAM}(x) = \sum_{i=1}^{n} w_i u_i(x_i). \tag{1}$$

Nevertheless, it can easily be shown by simple examples, that it fails to represent more complex, yet intuitive preference relations.

In Section 2 we represent the Choquet integral and discuss its qualities and possible modifications, and in Section 3 similar introduction to Multilinear forms is conducted. In the qualitative comparison in Section 4 differences and similarities of these methods are discussed, and conclusions are presented in Section 5.

2 The discrete Choquet integral

2.1 History

The evolution of the Choquet integral started in the 1950's with the *Theory of capacities* by Choquet [1954]. Its development towards a useful tool for multi-criteria decision making started in 1970's, when the concept of fuzzy integrals and fuzzy measures was proposed by Sugeno [1974]. The special feature of fuzzy integrals is that with them it is possible to model the interaction between criteria, namely redundancy and synergy (negative and positive interaction, respectively).

Later the Choquet integral was discussed with respect to fuzzy measures by Murofushi and Sugeno [1989], which has later on led to Choquet integral being called a fuzzy integral and capacity being called a fuzzy measure. Usage of the Choquet integral in the field of decision making did start only in the end of 1980's for decisions under uncertainty and in the beginning of the 1990's for multi-criteria decision aid. A snapshot of the evolution of the Choquet integral and its usage is provided by Grabisch and Labreuche [2010].

2.2 Introduction

In decision making we often encounter situations where we are no longer able to represent the preferences of the decision maker with the most basic and widely used aggregation functions such as the weighted arithmetic mean. In these situations we want to find more general aggregation functions to use. Let us illustrate this kind of a situation with a simple example presented by Grabisch and Labreuche [2010]:

Let a, b, c be three alternatives that are evaluated on two criteria as follows:

$$\begin{aligned} u_1(a) &= 0.4, & u_2(a) &= 0.4 \\ u_1(b) &= 0, & u_2(b) &= 1 \\ u_1(c) &= 1, & u_2(c) &= 0 \end{aligned}$$

The scores u for each alternative and criteria are given in [0,1]. Suppose the DM says that $a \succ b \sim c$. Let us find weights w_1 and w_2 so that this preference relation can be presented with values calculated with the WAM

$\text{WAM}(x) = w_1u_1(x) + w_2u_2(x)$. We get the following:

$$b \sim c \Leftrightarrow w_1 = w_2$$
$$a \succ b \Leftrightarrow 0.4w_1 + 0.4w_2 > w_2$$

this leads to $0.8w_2 > w_2$, which is impossible.

As we see with this example, the WAM fails to represent the preferences of a DM who prefers more balanced alternatives to those that are fully satisfactory on one criterion but fail totally on the other. It is important to understand, that this situation is in no way hypothetical but something that can be a part of a real-life decision problem.

We will later learn that the weighted arithmetic mean can be seen as a special case of the Choquet integral, when certain assumptions regarding the weighting coefficients are made. In the example we assumed that the criteria fulfilled alone were as valuable as criteria fulfilled simultaneously. In order to overrule that assumption and enable modeling interaction between criteria, let us define a new weight w_{12} to represent the importance of both criteria being satisfied simultaneously. As the situation where both criteria are evaluated as 1 must be the most desirable one in this case, we can choose $w_{12} = 1$ and scale other weights accordingly without loss of generality. This gives the alternative a value 0.4 in the example. On the other hand, criteria satisfied alone are not as attractive to the DM as criteria satisfied together (as defined earlier) and the weights w_1 and w_2 should be chosen accordingly. In the example situation all weights that satisfy $w_1 = w_2 < 0.4$ are adequate to represent the preferences of the DM.

The alternatives in the example above are chosen so, that they fit perfectly to the problem. To give a wider perspective of the situation, Grabisch and Labreuche [2010] introduce a situation with a fourth alternative d, which is evaluated on the same criteria:

$$u_1(d) = 0.2, \text{ and } u_2(d) = 0.8.$$

DMs preferences with the four alternatives are $a \succ d \succ b \sim c$. Now the alternative is partially satisfied on both criteria but not as balanced as alternative a.

To calculate the value for alternative d in the same manner as earlier for other alternatives, let us assume that the total value of d can be calculated as a sum of two fictitious alternatives, d' and d'', that are defined by:

$$u_1(d') = 0.2, \quad \text{and} \quad u_2(d') = 0.2$$
$$u_1(d'') = 0, \quad \text{and} \quad u_2(d'') = 0.6$$

Now the overall score for d can be calculated as a sum of the scores of alternatives d' and d''. If we for example chose that $w_1 = w_2 = 0.3$ and $w_{12} = 1$, the alternatives would get following scores:

$$\begin{aligned} F(a) &= 0.4 \\ F(b) &= 0.3 \\ F(c) &= 0.3 \\ F(d) &= 1 \cdot 0.2 + 0.3 \cdot 0.6 = 0.38 \end{aligned}$$

which satisfies the preferences of the DM.

Basically what we used here is the *Choquet integral*. To define the general form of the Choquet integral, we shall first define the *capacity* and its generalization *game* for a set N:

Definition 2.1.

1. A function $\nu : 2^N \to \mathbb{R}$ is a *game* if it satisfies $\nu(\emptyset) = 0$
2. A game μ which satisfies $\mu(A) \leq \mu(B)$ whenever $A \subseteq B$ (monotonicity) is called a *capacity*. The capacity is *normalized* if in addition $\mu(N) = 1$.

Definition 2.2. Consider $f : N \to \mathbb{R}_+$. The Choquet integral of f with respect to a capacity μ is given by

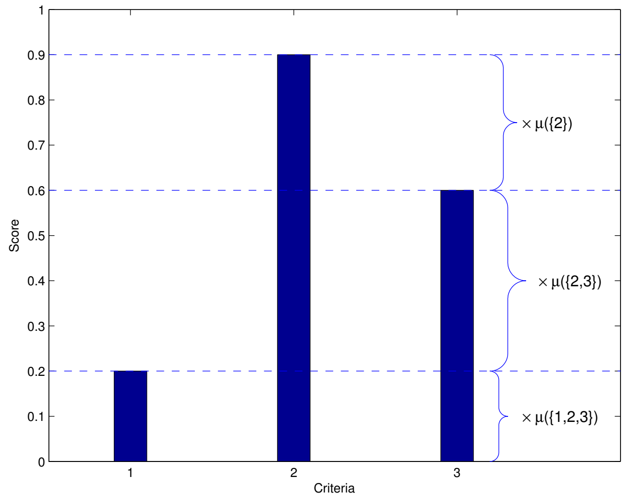
$$C_\mu(f) := \sum_{i=1}^{n} [f_{\sigma(i)} - f_{\sigma(i-1)}]\mu(\{\sigma(i), \ldots, \sigma(n)\}), \tag{2}$$

where f_i stands for $f(i)$, σ is a permutation on N such that $f_{\sigma(1)} \leq \cdots \leq f_{\sigma(n)}$ and $f_{\sigma(0)} := 0$ (Grabisch [1996]).

In order to simplify further notations, let us from now on refer to values of capacity $\mu(A)$ as *weights* for set A and the Choquet integral for alternative x as $C_\mu(x)$.

2.3 Graphical interpretation

To better understand the idea underlying the Choquet integral, let us demonstrate graphically a situation with an alternative being evaluated w.r.t three criteria. Consider an alternative a that has been given the following



Kuva 1: Graphical example of the Choquet integral with three criteria.

scores:

$$u_1(a) = 0.2$$
$$u_2(a) = 0.9$$
$$u_3(a) = 0.6$$

The scores are illustrated in Figure 1. The weights assigned for sets of criteria (i.e. capacity) are marked with $\mu(\{A\})$ where A is the set of criteria considered. The Choquet integral for the alternative a can be calculated with the Definition 2.2:

$$\begin{aligned} C_\mu(u) &= \sum_{i=1}^{3}[u_{\sigma(i)} - u_{\sigma(i-1)}]\mu(\{\sigma(i),\ldots,\sigma(n)\}) \\ &= (u_1 - u_0)\mu(\{1,2,3\}) + (u_2 - u_1)\mu(\{2,3\}) + (u_3 - u_2)\mu(\{2\}) \\ &= 0.2\mu(\{1,2,3\}) + 0.4\mu(\{2,3\}) + 0.3\mu(\{2\}). \end{aligned}$$

As an example, let us assign example weights for the three criteria and calculate the value for alternative a. The same weights will be used later to illustrate also other qualities of the Choquet integral.

Example 2.1. Assume that all criteria satisfied separately are of the same value, but that criteria 1 and 2 have synergy on each other and criteria 2 and 3 have redundancy on each other. Criteria 1 and 3 do not interact. These assumptions can be satisfied, for example, with the following weights:

$$\begin{aligned}
\mu(\{1\}) = \mu(\{2\}) = \mu(\{3\}) &= 0.3 \\
\mu(\{1,2\}) &= 0.75 \\
\mu(\{2,3\}) &= 0.55 \\
\mu(\{1,3\}) &= 0.6 \\
\mu(\{1,2,3\}) &= 1.
\end{aligned}$$

Note that $\mu(\{1,2\}) > \mu(\{1\}) + \mu(\{2\}) = 0.6$ which represents the synergy the criteria have on each other. Respectively $\mu(\{2,3\}) < 0.6$ to represent the redundancy.

With these weights the Choquet integral for alternative a is

$$C_\mu(a) = 0.2 \cdot 1 + 0.4 \cdot 0.55 + 0.3 \cdot 0.3 = 0.51.$$

The effects of redundancy and synergy are easier to notice when another alternative b is introduced. Choose b so that it is given following scores:

$$\begin{aligned}
u_1(b) &= 0.9 \\
u_2(b) &= 0.6 \\
u_3(b) &= 0.2
\end{aligned}$$

Note that the scores are exactly the same as with the alternative a except for different criteria. The Choquet integral for alternative b is

$$C_\mu(b) = 0.2 \cdot 1 + 0.4 \cdot 0.75 + 0.3 \cdot 0.3 = 0.59$$

which is higher than that of alternative a, even though the separate criteria were evaluated to be equally important ($\mu(\{1\}) = \mu(\{2\}) = \mu(\{3\}) = 0.3$) and the absolute values of the scores were same for both alternatives. The difference in values of alternatives is explained by the synergy and redundancy between criteria: for alternative b the positively interacting criteria 1 and 2 had higher scores than those of alternative a and simultaneously the criterion 3 was rated lower for alternative b than for alternative a. Even though criterion 3 alone was evaluated equally important as criteria 1 and 2, the negative effect it has on criterion 2 makes it less important than the other two criteria.

2.4 Special cases

We have now introduced an aggregation function, that is more general than the weighted arithmetic mean, but with the cost of exponentially growing amount of weights (2^n-2) to determine. As defining 2^n-2 coefficients gets laborious when the number of criteria n increases, reducing the complexity of the model gets desirable. By making some simplifying assumptions about the capacity to be defined, we can efficiently reduce the amount of weights that must be determined separately, and even achieve some well-known aggregation functions.

The most extreme variations of the Choquet integral are minimum and maximum and also the most common aggregation function, the weighted arithmetic mean, can be derived from the Choquet integral. Another interesting special case of the Choquet integral is the Ordered Weighted Average (OWA). A short discussion of these special cases and easing the determination of capacity is presented next.

Minimum and maximum

The most extreme functionality of the Choquet integral is achieved, when we define $\mu(A) = 0$ for all $A \subsetneq N$ and $\mu(N) = 1$. Choquet integral defined with these weights narrows down to the minimum function. On the other hand, if we define capacity such that $\mu(A) = 1$ for all $A \neq \emptyset$ and $\mu(\emptyset) = 0$, we get the maximum function. Practically this means, that for an alternative x and all possible capacities μ the following holds:

$$\min(x) \leq C_\mu(x) \leq \max(x).$$

This kind of aggregation functions are called *averaging* aggregation functions. In the Example 2.1 both alternatives would get value 0.2 when using the minimum function and value 0.9 when using the maximum function.

k-additive capacities and the weighted arithmetic mean

A capacity is *additive*, if for all disjoint sets $A, B \subseteq N$, we have $\mu(A \cup B) = \mu(A) + \mu(B)$. In the case of an additive capacity, the Choquet integral collapses into the weighted arithmetic mean: (2) can be written as

$$
\begin{aligned}
C_\mu(f) &= \sum_{i=1}^{n}[f_{\sigma(i)} - f_{\sigma(i-1)}]\mu(\{\sigma(i),\ldots,\sigma(n)\}) \\
&= \sum_{i=1}^{n}\left\{[f_{\sigma(i)} - f_{\sigma(i-1)}]\sum_{j=i}^{n}\mu(\{\sigma(j)\})\right\} \\
&= \sum_{i=1}^{n}\left\{f_{\sigma(i)}\sum_{j=i}^{n}\mu(\{\sigma(j)\}) - f_{\sigma(i-1)}\sum_{j=i}^{n}\mu(\{\sigma(j)\})\right\} \\
&= \sum_{i=1}^{n}\left\{f_{\sigma(i)}\sum_{j=i}^{n}\mu(\{\sigma(j)\}) - f_{\sigma(i-1)}\sum_{j=i-1}^{n}\mu(\{\sigma(j)\}) + f_{\sigma(i-1)}\mu(\{\sigma(i-1)\})\right\} \\
&= \underbrace{\sum_{i=1}^{n}\left\{f_{\sigma(i)}\sum_{j=i}^{n}\mu(\{\sigma(j)\}) - f_{\sigma(i-1)}\sum_{j=i-1}^{n}\mu(\{\sigma(j)\})\right\}}_{\text{Telescoping sum}} + \sum_{i=1}^{n} f_{\sigma(i-1)}\mu(\{\sigma(i-1)\}) \\
&= f_{\sigma(n)}\mu(\{\sigma(n)\}) - f_{\sigma(0)}\sum_{j=0}^{n}\mu(\{\sigma(j)\}) + f_{\sigma(0)}\mu(\{\sigma(0)\}) + \sum_{i=1}^{n-1} f_{\sigma(i)}\mu(\{\sigma(i)\})
\end{aligned}
$$

$f_{\sigma(0)} = 0$ by definition, and thus we get

$$C_\mu(f) = \sum_{i=1}^{n} f_{\sigma(i)}\mu(\{\sigma(i)\})$$

the order of terms is irrelevant in a sum, thus this brings us to

$$C_\mu(f) = \sum_{i=1}^{n} \mu(\{i\}) f_i$$

which is the weighted arithmetic mean.

From this we see that, with the assumption of additivity of the capacity, the interaction between criteria disappears. Nevertheless, the interaction of criteria is a desirable quality in some situations, as we discovered earlier. Consider a situation with a large number of criteria of which some have redundancy and some have synergy on each other. Because of the interaction, the WAM does not manage to represent the situation in a desirable way, but as the amount of weights to be defined is large, using the original form of Choquet integral is extremely laborious.

The key question is, whether we can somehow ease defining the capacity by making simplifying assumptions but still preserve some of the interaction in the system. Is it necessary to assume that there is significant interaction between sets of m criteria or can we consider that the capacity for larger sets is in fact additive?

For this we introduce the concept of *k-additivity* proposed by Grabisch [1997]. What k-additivity basically means, is that interaction between criteria can occur only in sets of at most k criteria. As we have seen, the WAM is 1-additive, meaning that interaction can only occur in sets of 1 criteria (which is no interaction). In order to define k-additivity, the concept of Möbius transform for a game is introduced:

Definition 2.3. Let ν be a game on N. The Möbius transform of ν, m^ν, is

$$m^\nu(A) = \sum_{B\subseteq A} (-1)^{|A\setminus B|}\nu(B), \qquad \forall A \subseteq N.$$

The Möbius transform is an alternative way to represent the capacity and the capacity ν can be recovered from its Möbius transform:

$$\nu(A) = \sum_{B\subseteq A} m^\nu(B), \qquad \forall A \subseteq N \tag{3}$$

For the Möbius transformation monotonicity and boundary conditions of the capacity are formulated respectively as (Beliakov et al. [2008])

$$\sum_{B\subseteq A|i\in B} m^\mu(B) \geq 0, \qquad \forall A \subseteq N \quad \text{and} \quad \forall i \in A$$
$$m^\mu(\emptyset) = 0$$
$$\sum_{A\subseteq N} m^\mu(A) = 1$$

Definition 2.4. A capacity is *k-additive* if its *Möbius transform* m^μ satisfies $m^\mu = 0$ for all $A \subseteq N$ such that $|A| > k$, and there exists $A \subseteq N$, $|A| = k$ such $m^\mu \neq 0$.

In case of k-additive capacities for n criteria, $\sum_{i=1}^{k}\binom{n}{i}$ coefficients need to be determined separately and the rest can be calculated from their Möbius transformations by using the inverse transform (3).

Example 2.2. Let us calculate the Möbius transform for the capacity μ

defined in Example 2.1.

$$
\begin{aligned}
m^{\mu}(\{1\}) &= (-1)^{|\{1\}|}\mu(\emptyset) + (-1)^{|\emptyset|}\mu(\{i\}) \\
&= -1 \cdot 0 + 1 \cdot 0.3 = 0.3 \\
m^{\mu}(\{2\}) &= 0.3 \\
m^{\mu}(\{3\}) &= 0.3 \\
m^{\mu}(\{1,2\}) &= -0.3 - 0.3 + 0.75 = 0.15 \\
m^{\mu}(\{2,3\}) &= -0.3 - 0.3 + 0.0.55 = -0.05 \\
m^{\mu}(\{1,3\}) &= -0.3 - 0.3 + 0.6 = 0 \\
m^{\mu}(\{1,2,3\}) &= 0.3 + 0.3 + 0.3 - 0.55 - 0.6 - 0.75 + 1 = 0.
\end{aligned}
$$

From these values we notice, that the capacity in the example is in fact 2-additive. Another interesting quality of the Möbius transform is, that the transformations for sets of one criterion is equal to the value of capacity for that criterion. Capacity was defined in such way, that criteria 1 and 2 had synergy and criteria 2 and 3 redundancy whereas criteria 1 and 3 had no interaction. These interactions can also be interpreted from the Möbius transform; positive value for $m^{\mu}(\{1,2\})$, negative value for $m^{\mu}(\{2,3\})$ and zero for $m^{\mu}(\{1,3\})$.

Symmetric capacities and the Ordered Weighted Average

A capacity is *symmetric* if and only if for any subsets $A, B, |A| = |B|$ implies that $\mu(A) = \mu(B)$. Let us denote $\mu_i = \mu(A)$ when $|A| = i$. Thus $\mu(\{\sigma(i), \ldots, \sigma(n)\})$ can be marked simply as μ_{n-i+1}. With this in mind, the Choquet integral (2) can be expressed as follows

$$
\begin{aligned}
C_{\mu}(f) &= \sum_{i=1}^{n} [f_{\sigma(i)} - f_{\sigma(i-1)}]\mu_{n-i+1} \\
&= \sum_{i=1}^{n} f_{\sigma(i)}\mu_{n-i+1} - \sum_{i=1}^{n} f_{\sigma(i-1)}\mu_{n-i+1}
\end{aligned}
$$

since $f_{\sigma(0)}$ is defined as zero, the first term of the latter sum is zero. On the other hand, as the value of capacity of an empty set is also zero, i.e. $\mu_0 = 0$

we get

$$
\begin{aligned}
C_\mu(f) &= \sum_{i=1}^{n} f_{\sigma(i)}\mu_{n-i+1} - \sum_{i=2}^{n} f_{\sigma(i-1)}\mu_{n-i+1} + f_{\sigma(n+1-1)}\mu_{n-n-1+1} \\
&= \sum_{i=1}^{n} f_{\sigma(i)}\mu_{n-i+1} - \sum_{i=2}^{n+1} f_{\sigma(i-1)}\mu_{n-i+1} \\
&= \sum_{i=1}^{n} f_{\sigma(i)}\mu_{n-i+1} - \sum_{i=1}^{n} f_{\sigma(i)}\mu_{n-i} \\
C_\mu(f) &= \sum_{i=1}^{n} (\mu_{n-i+1} - \mu_{n-i}) f_{\sigma(i)}, \qquad (4)
\end{aligned}
$$

which is better known as the OWA operator or the ordered weighted average. OWA operators are not as widely known and as much used as the arithmetic mean.

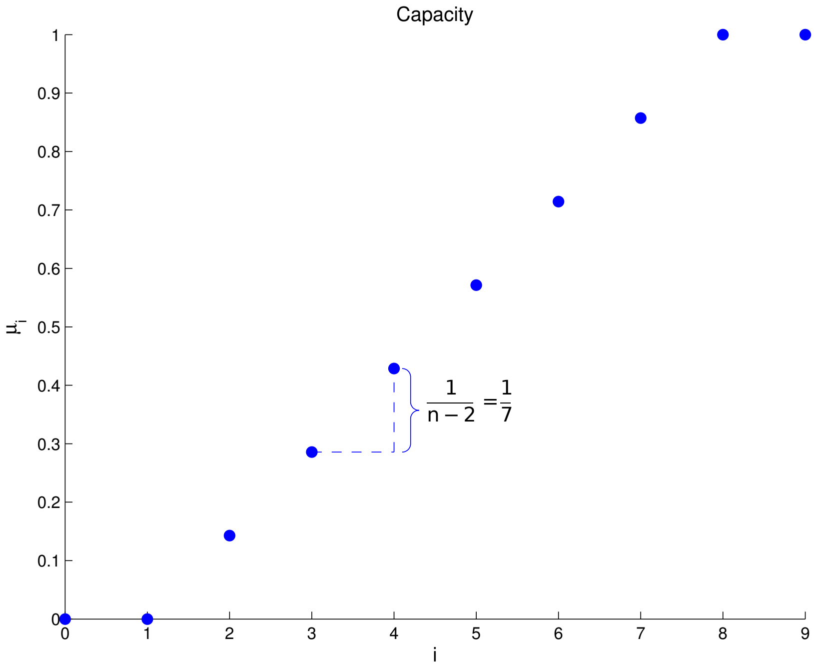
OWA operators are useful in situations where there are criteria used to evaluate the alternatives, but the DM does not value one criteria to another; only the relative order of magnitude matters. An example of usage of OWA operators is, when the performance of an athlete, for example a ski jumper, is evaluated w.r.t. evaluations from multiple referees. The opinion of a specific referee is not more important than that of another, but usually the extreme values are left out from evaluation to avoid distortion.

Example 2.3. Consider the evaluation situation of the style of a ski jumper with nine referees. The scores for the ski jumpers are achieved when the highest and the lowest scores given by referees are left out and then an average is taken from the rest. The performances are evaluated in [0,20]. In order to represent this with Choquet integral we need a symmetric capacity denoted as μ_i where i is the cardinality of the set of criteria. For simplicity we define $\mu_{n-i+1} - \mu_{n-i} = \omega_i$, and the equation (4) can be written as

$$C_\mu(f) = \sum_{i=1}^{n} \omega_i f_{\sigma(i)} =: \mathrm{OWA}(f),$$

where ω_i are the coefficients for the OWA operator and they satisfy $\sum_i \omega_i = 1$. The desired score aggregation is achieved with coefficients $\omega_1 = \omega_9 = 0$ and $\omega_i = \frac{1}{7}$ when $i \in \{2...8\}$.

The capacity with which this is achieved is represented in Figure 2.



Kuva 2: Values of capacity for sets of cardinality i

If a ski jumper was given for example evaluations $(15, 12, 16, 18, 18, 17, 16, 14, 19)$ the value of this Choquet integral would be

$$C_\mu = 0 \cdot 12 + \frac{1}{7} \cdot 14 + \frac{1}{7} \cdot 15 + \frac{1}{7} \cdot 16 + \frac{1}{7} \cdot 16 + \frac{1}{7} \cdot 17 + \frac{1}{7} \cdot 18 + \frac{1}{7} \cdot 18 + 0 \cdot 19$$
$$\approx 16.29.$$

2.5 Indices and interpretation of the capacity

In addition to the exponentially growing amount of weights, another problem occurs with the Choquet integral: interpretation of capacities. Compared for example to the WAM, the weights for criteria do no longer directly represent the importance of that criterion as the criterion might be redundant or have notable synergy to other criteria. In this section we learn two indices, the importance index and the interaction index, that help us better understand and interpret the capacities. In addition, the concept of *orness* is presented.

The importance index

With Choquet integral, a numerical value to represent the importance of a criterion comes in need, as the weight assigned for the criterion itself no longer perfectly describes the significance of the criterion (compare with the situation of the weighted arithmetic mean). In other words, we are interested in knowing, how large is the effect when criterion i is added to some coalition A of criteria. The *Shapley importance index* (Shapley [1952]) can be calculated to tell the relative importance of the criterion among criteria:

$$\phi_i(\mu) = \sum_{A \subseteq N \setminus \{i\}} \underbrace{\frac{|A|!(n-|A|-1)!}{n!}}_{\text{Normalization factor}} (\mu(A \cup \{i\}) - \mu(A)). \tag{5}$$

Basically this is the average weight of criterion i over all possible profiles in $[0,1]^n$.

In Section 2.3 we assigned example weights for a three criteria system. Let us now use the same capacity and calculate the importance index for each criteria. The weights were assigned in following manner:

$$\begin{aligned} \mu(\{1\}) = \mu(\{2\}) = \mu(\{3\}) &= 0.3 \\ \mu(\{1,2\}) &= 0.75 \\ \mu(\{2,3\}) &= 0.55 \\ \mu(\{1,3\}) &= 0.6 \\ \mu(\{1,2,3\}) &= 1 \\ \mu(\{\emptyset\}) &= 0 \end{aligned}$$

The important indices can be calculated for each criterion with the equation

(5):

$$
\begin{aligned}
\phi_1(\mu) = &\sum_{A \subseteq \{1,2\}} \frac{|A|!(3-|A|-1)!}{3!}(\mu(A \cup \{1\}) - \mu(A)) \\
=&\frac{|\emptyset|!(3-|\emptyset|-1)!}{3!}(\mu(\{1\}) - \mu(\emptyset)) \\
&+ \frac{|\{2\}|!(3-|\{2\}|-1)!}{3!}(\mu(\{1,2\}) - \mu(\{2\})) \\
&+ \frac{|\{3\}|!(3-|\{3\}|-1)!}{3!}(\mu(\{1,3\}) - \mu(\{3\})) \\
&+ \frac{|\{2,3\}|!(3-|\{2,3\}|-1)!}{3!}(\mu(\{1,2,3\}) - \mu(\{2,3\})) \\
=&\frac{0!(3-0-1)!}{6}(0.3-0) + \frac{1!(3-1-1)!}{6}(0.75-0.3)) \\
&+ \frac{1!(3-1-1)!}{6}(0.6-0.3) + \frac{2!(3-2-1)!}{6}(1-0.55) \\
=&\frac{2}{6} \cdot 0.3 + \frac{1}{6} \cdot 0.45 + \frac{1}{6} \cdot 0.3 + \frac{2}{6} \cdot 0.45 \\
=&0.375.
\end{aligned}
$$

When the indices for criteria 2 and 3 are calculated in the same manner, the following values are achieved:

$$
\begin{aligned}
\phi_2 &= 0.35 \\
\phi_3 &= 0.275
\end{aligned}
$$

From these values we can see that even though the separate criteria were evaluated equally important (all given the same weight 0.3), because of the positive interaction between criteria 1 and 2 and the negative interaction between criteria 2 and 3, the actual importance of criteria is different. The criterion 1 having no negative interactions has the highest relative importance whereas the criterion 3 that had negative interaction with criterion 2 has the lowest relative importance.

The interaction index

Since the Choquet integral is able to model interaction between criteria, we are also interested in knowing how two criteria interact. To represent the synergy or redundancy between criteria, we can calculate the *interaction index* discussed by Grabisch and Labreuche [2010]:

$$I_{\{i,j\}}(\mu) = \sum_{A \subseteq N \setminus \{i,j\}} \frac{|A|!(n-|A|-2)!}{(n-1)!} \delta^A_{\{i,j\}}(\mu) \tag{6}$$

where

$$\begin{aligned} \delta^A_{\{i,j\}}(\mu) :&= \delta^A_{ij}(\mu) - \delta^A_i(\mu) - \delta^A_j(\mu) \\ &= \mu(A \cup \{i,j\}) - \mu(A \cup \{i\}) - \mu(A \cup \{j\}) + \mu(A). \end{aligned}$$

In Example 2.1 the weights were assigned so, that there was synergy between criteria 1 and 2, redundancy between criteria 2 and 3, and no interaction between criteria 1 and 3. By calculating the interaction indices we can prove that these weights actually do satisfy the assumptions regarding interaction. With equation (6), the following interaction indices are achieved:

$$\begin{aligned} I_{\{1,2\}}(\mu) &= 0.15 \\ I_{\{2,3\}}(\mu) &= -0.05 \\ I_{\{1,3\}}(\mu) &= 0 \end{aligned}$$

The positive value of $I_{\{1,2\}}$ indicates synergy between criteria 1 and 2, whereas the negative values of interaction index indicate redundancy.

An interesting connection with the Möbius transform introduced in Section 2.4 is discovered when the interaction indices are compared to the Möbius transform of the capacity calculated in Example 2.2. We notice the Möbius transform for sets of two criteria is equal to the interaction indices calculated above. The reason to this is actually the 2-additivity of our example as for 2-additive capacities the following holds: $\forall A \subseteq N$, $|A| = 2$ we have $I_A = m^\mu(A)$ (Grabisch and Labreuche [2010]).

Orness

Orness is an index to describe an *averaging* aggregation function. As we have earlier learned, an averaging aggregation function is such, that it always valuates alternatives higher than or equal to minimum function but lower than or equal to maximum function. Practically orness gives values in [0,1] and the value represents how close the aggregation function is to maximum.

Orness of the minimum is 0 and the maximum 1. The orness of an aggregation function F is defined as

$$\text{orness}(F) := \frac{\overline{F} - \overline{\min}}{\overline{\max} - \overline{\min}},$$

where $\overline{F}$ is the expected value of F in $[0,1]^n$. The expected values of min and max are

$$\overline{\min} = \frac{1}{n+1}$$
$$\overline{\max} = \frac{n}{n+1}$$

and thus we can express orness as

$$\text{orness}(F) = -\frac{1}{n-1} + \frac{n+1}{n-1}\overline{F}.$$

As Choquet integral is an averaging aggregation function, its orness can be calculated (Marichal [2004]) as:

$$\text{orness}(C_\mu) := \frac{1}{n-1}\sum_{t=1}^{n-1}\frac{1}{\binom{n}{t}}\sum_{\substack{T\subseteq N\\|T|=t}}\mu(T).$$

Earlier in Example 2.1 we assigned example weights for a Choquet integral of three criteria. The orness of it calculated with this formula is $\text{orness}(C_\mu) \approx 0.47$ which indicates a slightly intolerant DM. Aggregation functions with orness < 0.5 fit for intolerant DMs who demand that most criteria are satisfied. Aggregation functions with orness > 0.5 on the contrary represent the preferences of a more tolerant DM.

Usage

With these indices we can better understand and describe our models and explain the behavior of the DM but in addition they can also be used in model construction. Defining the capacity to meet the DMs preferences is a complicated task, and it often is not possible to directly set weights for sets of criteria so, that they would actually model the DMs preferences. These indices help giving the mathematical form to the verbal statements given by the DM.

In situations where relatively large set of criteria is taken into account and a need to model the interactions between these criteria occurs, 2-additivity offers an alternative to extremely complicated models with large number of coefficients. As learned earlier, the interaction index has a connection to 2-additive capacities and therefore it can be used to ease the determination of weights for the Choquet integral.

2.6 In short

The Choquet integral provides an adaptable way to construct an aggregation function for different situations and different kinds of decision makers. As a downside for flexibility, the number of coefficients to be determined grows exponentially and the interpretation of these coefficients gets more complicated.

We have learned that with certain assumptions regarding the capacity, the Choquet integral can be reduced to more simple aggregation functions such as the weighted arithmetic mean and the ordered weighted average. Concept of k-additive capacities was introduced to offer a compromise between the difficulty of determining the coefficients and complexity of the model. Especially 2-additive capacities were found sensible, as information of desired criteria interaction could be utilized in capacity determination.

Even with simplified capacities the interpretation of coefficients is not as straight-forward as it is for example with the weighted arithmetic mean. Nevertheless, the importance and interaction of criteria can be presented with different indices to ease the interpretation of capacity and the aggregation function. Also the fundamental nature of aggregation functions can be described with different kinds of indices, of which we have discussed the concept of orness. Orness can be calculated for averaging aggregation functions and it describes how similar the function is to the maximum -function.

3 Multilinear forms

3.1 History

Multilinear forms were originally introduced by Owen [1972] as the *multilinear extension* (MLE), and they were used to help compute values of large, n-person games ν (see Definition 2.1). The connections between games and their multilinear extension enabled using the MLE to observe certain properties of games.

Zeleny [1982] discussed the usage of the multilinear forms among other functions in the field of utility measurement in multiple criteria decision making problems, pointing out the complexity and increasing number of scaling constants (later: weights) when the number of attributes increases. Concentrating on Multi-attribute utility theory (MAUT), focus was set more on verifying the independence conditions required and methods for weight determination than the features and selection of aggregation functions. However, Zeleny [1982] and Von Winterfeldt and Edwards [1986] pointed out that using multilinear forms required looser independence conditions than, for example, the most commonly used aggregation function, the arithmetic mean, as the interaction between attributes could be modeled.

The usage of multilinear forms in the field of decision making with multiple objectives was discussed by Keeney and Raiffa [1993], also showing how weighting coefficients could be evaluated from the evaluations of the decision maker.

3.2 Introduction

Multilinear functions are functions that are linear with respect to each separate variable, but might contain terms with products of variables, e.g. for example $f(x, y) = x + y + xy$ is a multilinear function whereas $f(x, y) = x + y + x^2$ is not.

Multilinear forms, also referred to as multilinear extension (MLE, Grabisch et al. [2009]) and multilinear model (Zeleny [1982]) can be used for value determination in decision making problems.

Definition 3.1. Consider alternative $x = (x_1, x_2, ..., x_n)$ evaluated with respect to n attributes, and monotone increasing single attribute utility functions $u_i : N \to \mathbb{R}_+$. The total value of alternative x given by the *Multilinear* *form* MLE(x) with respect to weighting coefficients λ_A is

$$\mathrm{MLE}(x) := \sum_{A \subseteq N} \lambda_A \prod_{i \in A} u_i(x_i), \tag{7}$$

which is equivalent to

$$\begin{aligned}\mathrm{MLE}(x) = &\sum_{i=1}^{n} \lambda_i u_i(x_i) + \sum_{i=1}^{n} \sum_{j>i} \lambda_{ij} u_i(x_i) u_j(x_j) \\ &+ \sum_{i=1}^{n} \sum_{j>i} \sum_{k>j} \lambda_{ijk} u_i(x_i) u_j(x_j) u_k(x_k) \\ &+ \cdots + \lambda_{123\cdots n} u_1(x_1) u_2(x_2) \cdots u_n(x_n).\end{aligned}$$

For simplification here we assume that single attribute utility functions $u_i : N \to [0,1]$ and MLE(x) : $[0,1]^n \to [0,1]$, and that all the functions are monotone increasing with respect to each variable. Zero denotes the lowest, and one the highest level of preference. These requirements set constraints the values that can be set to weights λ_A, when the function is utilized in decision making problems. More of how these weights can be evaluated from evaluations by the DM was presented by Keeney and Raiffa [1993].

When the MLE is applied to three criteria situation, the effect of weights λ is easier to interpret:

$$\begin{aligned}\mathrm{MLE}(x_1, x_2, x_3) = &\lambda_1 u_1(x_1) + \lambda_2 u_2(x_2) + \lambda_3 u_3(x_3) \\ &+ \lambda_{12} u_1(x_1) u_2(x_2) \\ &+ \lambda_{13} u_1(x_1) u_3(x_3) \\ &+ \lambda_{23} u_2(x_2) u_3(x_3) \\ &+ \lambda_{123} u_1(x_1) u_2(x_2) u_3(x_3).\end{aligned}$$

Since we demand that each criterion is evaluated so, that a bigger value is always more desirable than a smaller value, we notice that $\lambda_1, \lambda_2, \lambda_3 \geq 0$, but the sign for other weights can, at least in theory, also be negative. Next we present an example to examine closer how the weighting coefficients of product terms reflect the preferences of the DM.

Example 3.1. Consider a situation with a decision maker who prefers alternatives with balanced attribute values, for example with alternatives a, b

and c evaluated as

$$\begin{aligned} u_1(a) &= 0.4, & u_2(a) &= 0.4 \\ u_1(b) &= 0, & u_2(b) &= 1 \\ u_1(c) &= 1, & u_2(c) &= 0 \end{aligned}$$

the DM considers $a \succ b \sim c$. Earlier we have shown that the basic weighted arithmetic mean fails to represent the preference relations of the DM, but can we find such weights λ that the multilinear form represents the DM's preferences? Assume single criteria utility functions here are linear, namely $u_i(x) = x$.

The multilinear form for two criteria is

$$\text{MLE}(x_1, x_2) = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_{12} x_1 x_2.$$

Because $b \sim c$ we get $\lambda_1 = \lambda_2 = \text{MLE}(b) = \text{MLE}(c)$. With this in mind, for alternative $a \succ b$ we demand that

$$\begin{aligned} \text{MLE}(a) &= \lambda_1 0.4 + \lambda_2 0.4 + \lambda_{12} 0.16 \\ &= \lambda_1 0.8 + \lambda_{12} 0.16 > \text{MLE}(b) = \lambda_1 \\ \Rightarrow \lambda_1 &< \lambda_{12} 0.8. \end{aligned}$$

When we also demand that $\text{MLE}(1,1) = 1$, we get $\lambda_1 + \lambda_2 + \lambda_{12} = 1$ and the inequality results in $\lambda_1 < 0.30769$. For example with $\lambda_1 = \lambda_2 = 0.25$ and $\lambda_{12} = 0.5$ the alternatives get the following values:

$$\begin{aligned} \text{MLE}(a) &= 0.28 \\ \text{MLE}(b) &= 0.25 \\ \text{MLE}(c) &= 0.25 \end{aligned}$$

which match the preference order defined by the decision maker.

In the example we had a situation where the criteria had positive interaction and the value of simultaneously satisfied criteria was greater than the value of criteria satisfied separately. In case of synergy within A we notice that the weight λ_A was positive thus giving the product term a positive effect on total score. If we consider a situation of redundancy and negatively interacting criteria, we notice that the desired effect on total values is achieved when the weight for the product term is set negative.

3.3 Special cases

Using multilinear forms enable modeling more complex preference relations since with them it is possible to evaluate situations, where the criteria have synergy or redundancy on each other. Using MLE however requires the determination of exponentially growing number of weights ($2^n - 2$), but the complexity of the model can be reduced with simplifying assumptions regarding the weights.

In this section we consider two special cases of multilinear forms: the weighted arithmetic mean and the multiplicative model, and what kind of assumptions regarding weighting coefficients reduce the MLE to these aggregation functions.

Weighted arithmetic mean

WAM (1) is achieved from MLE (7), when weights for sets of criteria of cardinality 1 sum up to 1, and all other weights are set to zero. As in multilinear forms the effect of criteria interaction is created by the product terms of multiple criteria values, setting only weights λ_i non-zero, disables the possibility for criteria interaction.

Multiplicative model

The multiplicative model, also known as log additive model (Zeleny [1982]), reduces the amount of weights by redefining interaction terms weights λ_A with weights for individual criteria λ_i scaled with a constant $k \in \mathbb{R}$ to the power of the cardinality of interaction:

$$\lambda_A = k^{|A|-1} \prod_{i \in A} \lambda_i.$$

For three criteria the multiplicative model can be written as

$$\begin{aligned} u(x_1, x_2, x_3) = &\lambda_1 u_1(x_1) + \lambda_2 u_2(x_2) + \lambda_3 u_3(x_3) \\ &+ k\lambda_1\lambda_2 u_1(x_1)u_2(x_2) \\ &+ k\lambda_1\lambda_3 u_1(x_1)u_3(x_3) \\ &+ k\lambda_2\lambda_3 u_2(x_2)u_3(x_3) \\ &+ k^2\lambda_1\lambda_2\lambda_3 u_1(x_1)u_2(x_2)u_3(x_3). \end{aligned}$$

This reduces the amount of weights to $n+1$ but simultaneously constrains the interaction that can be modeled with it (Von Winterfeldt and Edwards [1986]). Observe that in a situation of only two criteria, the multiplicative model is exactly the same as the corresponding multilinear form with $\lambda_{12} = k\lambda_1\lambda_2$. Possible values for k are determined by the boundary and monotonicity conditions defined.

3.4 Interpretation of the weights

In decision making problems the weights are derived from evaluations of a DM, but can these weights be given any interpretation, for example to reflect the DM's attitudes towards criteria? Next we discuss the weights' connection to the Möbius transform of a game.

Connection to Möbius transform and capacity

It has been proved by Grabisch et al. [2009] that the constants λ in Definition 3.1 can be expressed as the Möbius transform of a game ν and thus the MLE can be rewritten as

$$\text{MLE}_\nu(x) = \sum_{A \subseteq N} m^\nu(A) \prod_{i \in A} u_i(x_i), \tag{8}$$

where $m^\nu(A)$ is the Möbius transform of $\nu(A)$.

Proposition 3.1. Let $u_i : N \to [0,1]$. If the game ν is a capacity (Definition 2.1), the Multilinear extension $\text{MLE}_\nu(x)$ is monotone increasing.

Todistus. MLE(x) is monotone increasing if and only if for each u_i, $\frac{\partial \text{MLE}}{\partial u_i} \geq 0$. The partial derivative of (8) with respect to u_i is

$$\frac{\partial \text{MLE}}{\partial u_i} = m^\mu(i) + \sum_{\substack{A \subseteq N \\ i \in A}} m^\mu(A) \prod_{j \in A \setminus \{i\}} u_j(x_j), \tag{9}$$

which is constant with respect to variable u_i.

The monotonicity condition for the Möbius transform of the capacity is, as presented in Section 2.4:

$$\sum_{A \subseteq B | i \in A} m^\mu(A) \geq 0, \qquad \forall B \subseteq N \quad \text{and} \quad \forall i \in B.$$

Let us first consider the monotonicity of the MLE on the vertices of the $[0,1]^n$ space. Assume $B \subseteq N$ so, that $i \in B$ and each $j \neq i, j \in B$ gets value $u_j(x_j) = 1$ and all $j \notin B$ get value $u_j(x_j) = 0$. The partial derivate (9) is reduced to

$$\frac{\partial \text{MLE}}{\partial u_i} = \sum_{\substack{A \subseteq B \\ i \in A}} m^{\mu}(A),$$

which is greater or equal to zero because of the monotonicity condition of the capacity $\rightarrow$ MLE is monotone increasing in the vertices of the $[0,1]^n$ space.

Next consider the partial derivative (9) in $(0,1)^n$. Weierstrass theorem states that a continuous function attains its maximum and minimum values in a closed bounded interval. Since the partial derivative (9) is continuous and attains only non-negative values at the boundaries of the $[0,1]^n$ space, the only possibility for the partial derivative to be negative within the interval is if its gradient is zero at some point within the open interval $(0,1)^n$. The gradient is zero if and only if at some point x the partial derivatives of (9) w.r.t. every variable u_j are zero. The partial derivative of (9) is

$$\frac{\partial}{\partial u_j} \frac{\partial \text{MLE}}{\partial u_i} = m^{\mu}(\{i,j\}) + \sum_{\substack{A \subseteq N \\ i,j \in A}} m^{\mu}(A) \prod_{k \in A \setminus \{i,j\}} u_k(x_k),$$

which is constant w.r.t u_j and thus (9) attains its minimum and maximum values at the boundaries of the interval $\Rightarrow$ MLE is monotone increasing in $[0,1]^n$. □

Since the weighting coefficients of the MLE can be expressed with a capacity, it is possible to evaluate the criteria with the Importance index and the Interaction index presented in Section 2.5.

3.5 In short

We have here introduced the multilinear form, the multilinear aggregation function that can be utilized in decision making problems. With multilinear forms it is possible to evaluate alternatives with respect to nonindependent criteria, but when the dimension of the problem increases, using Multilinear forms gets more difficult due to the growing number of weights to be determined. We also introduced some simple ways to reduce the number of weights with the cost of the flexibility of the model and even the weighted arithmetic mean was achieved. The multiplicative model reduces the amount of weights to $n+1$ without completely removing the possibility for criteria interaction.

It was established that the MLE has a connection to games and certain assumptions regarding monotonicity and boundary conditions can be achieved when the weights coincide with the Möbius transform of a capacity.

4 Comparison

In Sections 2 and 3 we introduced two different aggregation functions that were both able to model criteria interaction, thus providing a possibility to model more demanding preference relations. Even though these two methods have evolved separately and from different premises, several connections between them can be found.

In this section we study similarities and differences between these methods and how are the decisions are affected by the selection between these aggregation functions.

4.1 Mathematical connection

It was proved by Grabisch et al. [2009] that an equivalent expression to the Choquet integral (2) is

$$C_\mu(x) = \sum_{A \subseteq N} m^\mu(A) \min_{i \in A}\{u_i(x_i)\}, \tag{10}$$

which is the same form as the MLE in (8), except that the product operator is replaced with a minimum operator. Basically, both aggregation functions can be expressed with respect to the same capacity, but when do these two functions actually coincide?

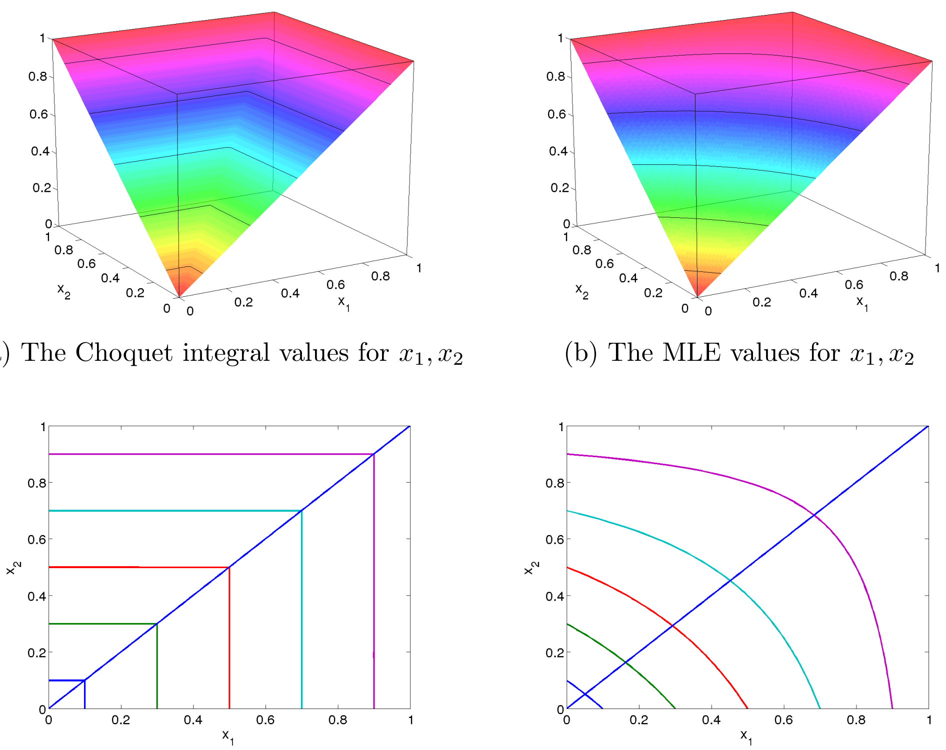
Earlier we have noticed that the MLE reduces to the weighted arithmetic mean when $\sum_i^n m^\mu(i) = 1$ and $m^\mu(A) = 0$ for all A such that $|A| > 1$. The minimum of a single value is the same as the product of a single value, thus the Choquet integral and MLE reduce to the same WAM.

We notice that the two functions also coincide when $u_i(x_i) \in \{0, 1\}$ since the product of zeros and ones is always equal to the minimum of the factors of the product. On the other hand $\min\{x_i\} = \prod_i x_i$ also when at most one factor has a value in (0,1) while others are in $\{0, 1\}$. Basically we notice that the two aggregation functions coincide on the edges of the $[0, 1]^n$ value space. Let us next demonstrate this graphically in a simple two-criteria situation.

4.2 Graphical connection

Since both the Choquet integral and MLE can be expressed in terms of the Möbius transform of the capacity, it is intuitive to compare these two when same weighting coefficients are used for both methods.

First consider a 2-criteria situation with weights $m^{\mu}(\{x_1\}) = m^{\mu}(\{x_2\}) = 1$ and $m^{\mu}(\{x_1, x_2\}) = -1$. The Choquet integral with these weights is reduced to the maximum function, as discussed earlier in Section 2.4. The value surface for the Choquet integral is drawn in Figure 3a and contour lines for values $C_{\mu} \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ are drawn in Figure 3c. Corresponding representations for the multilinear form are in Figures 3b and 3d. We notice that because of the multi-linearity of the MLE, the value surface is smooth unlike that of the Choquet.



(a) The Choquet integral values for x_1, x_2

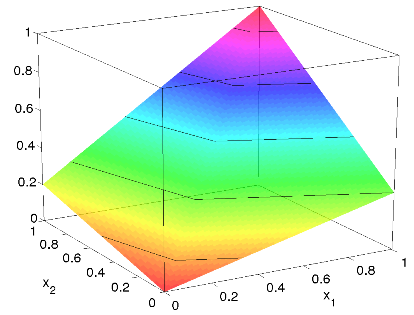
(b) The MLE values for x_1, x_2

(c) Contour plot for the Choquet integral

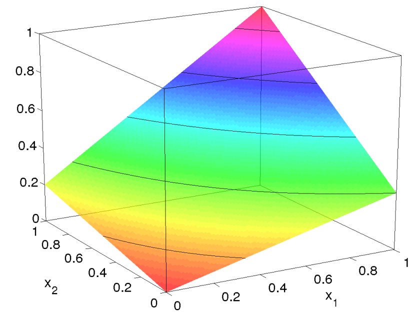
(d) Contour plot for the MLE

Kuva 3: Values given by the Choquet integral and the MLE in 2-criteria situation with weights $m^{\mu}(\{x_1\}) = m^{\mu}(\{x_2\}) = 1$ and $m^{\mu}(\{x_1, x_2\}) = -1$.

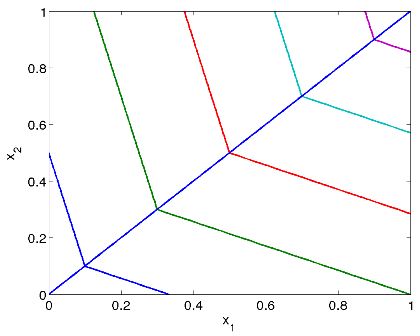
As earlier speculated from the mathematical equations, we can indeed notice that the two functions give the same values at the borders of the value surface. But how do the values of the functions change if we have criteria that have positive effect on each other? Let us, for example, choose the following weights: $m^{\mu}(\{x_1\}) = 0.2, m^{\mu}(\{x_2\}) = 0.3$ and $m^{\mu}(\{x_1, x_2\}) = 0.5$. The value surfaces and contour plots for the functions are shown in Figure 4.



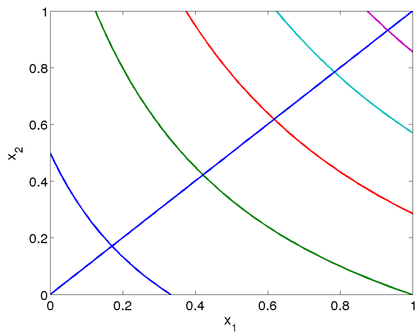
(a) The Choquet integral values for x_1, x_2



(b) The MLE values for x_1, x_2



(c) Contour plot for the Choquet integral



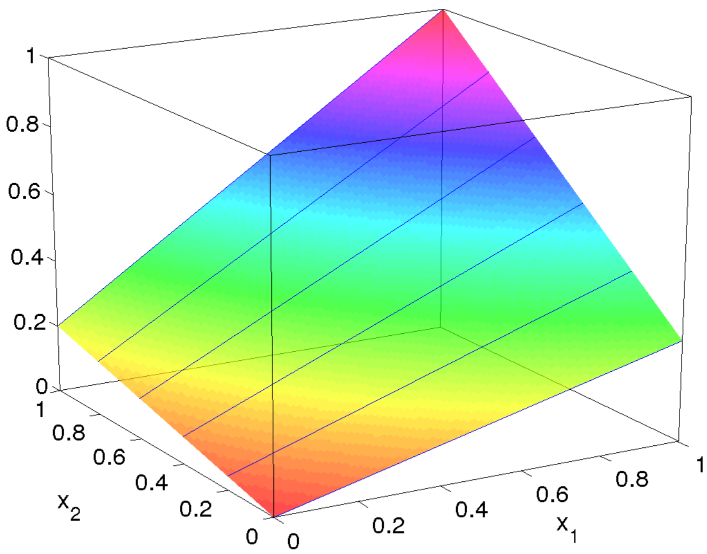
(d) Contour plot for the MLE

Kuva 4: Values given by the Choquet integral and the MLE in 2-criteria situation with weights $m^\mu(\{x_1\}) = 0.2, m^\mu(\{x_2\}) = 0.3$ and $m^\mu(\{x_1, x_2\}) = 0.5$.

Since the Choquet integral is an averaging aggregation function, its value is always between the maximum and the minimum of the scores of the criteria resulting that if $x_1 = ... = x_n$ the Choquet integral always gets value $C(x) = x_1$. The value surface in 2-criteria situation consists of two planes that connect at $x_1 = x_2$: the line where the value surface is continuous, but differentiable only if the Choquet integral is also a weighted arithmetic mean.

We perceive that unlike the value surface of the Choquet integral, that of the MLE is smooth and differentiable in the defined space, which in some applications might prove to be a useful quality. The contour lines are not linear, but the curvature is restricted by the multi-linearity of the MLE. In Figure 5 we have the value surface from Figure 4b with value lines on the surface to emphasize how the value increases with respect to the criterion x_1

when the score for criteria x_2 is held constant ($x_2 \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$). Because of the multi-linearity of the function, the value in these situations increases linearly, but as the criteria have synergy on each other, the value increases with greater rate with larger values of x_2.



Kuva 5: The MLE value surface.

4.3 Effects of aggregation function selection on decision guidelines

Since the contours for the Choquet integrals are piecewise linear but those of the MLE are curved (excluding the WAM), it is possible to find a preference relation that we can represent with the MLE but not with the Choquet integral and vice versa. In this section we examine an example of one of these situations, and consider whether the differences actually matter in real life applications.

Consider a situation with three alternatives, a, b and c evaluated on two criteria:

$$\begin{aligned} u_1(a) &= 0, & u_2(a) &= 0.9025 \\ u_1(b) &= 0.5, & u_2(b) &= 0.795 \\ u_1(c) &= 0.67, & u_2(c) &= 0.7045 \end{aligned}$$

None of the alternatives is dominated by other alternatives based on their scores. Consider DM with preference order $a \sim c \succ b$. With MLE this preference relation is achieved, yet with only a slight numerical difference, with weights $m^\mu(\{x_1\}) = m^\mu(\{x_2\}) = 1$ and $m^\mu(\{x_1, x_2\}) = -1$ resulting in following values for alternatives:

$$\begin{aligned} \text{MLE}_\mu(a) &= 0.9025 \\ \text{MLE}_\mu(b) &= 0.8975 \\ \text{MLE}_\mu(c) &= 0.9025 \end{aligned}$$

The Choquet integral with the same weights on the contrary evaluates alternatives as follows:

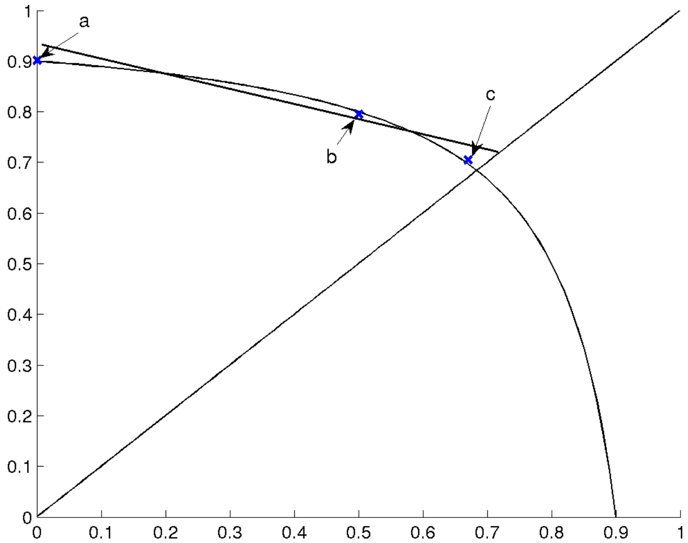
$$\begin{aligned} C_\mu(a) &= 0.9025 \\ C_\mu(b) &= 0.7950 \\ C_\mu(c) &= 0.7045 \end{aligned}$$

which does not at all reflect the preference order of the DM.

The three alternatives are plotted in (x_1, x_2) space in Figure 6 with the contour line of the earlier specified MLE. It is noteworthy that the three alternatives are all above the line $x_1 = x_2$, and that it is impossible to draw a straight line in such way, that it would go through both alternatives a and c, and the alternative b would be beneath the line.

Earlier we noticed that only a slight difference between alternatives was provided by the MLE with the chosen weights. Let us next find such weights for the Choquet integral that $a \sim c$, and re-calculate the values of the alternatives. For clarity we use notation $m^\mu(\{x_i\}) = w_i$, and $w_{1,2} = 1 - w_1 - w_2$ results from the boundary conditions of the Choquet integral.

$$\begin{aligned} C_\mu(a) &= C_\mu(c) \\ 0.9025w_2 &= 0.67w_1 + 0.7045w_2 + (1 - w_1 - w_2)0.67 \\ 0.9025w_2 &= 0.67w_1 - 0.67w_1 + 0.7045w_2 + -0.67w_2 + 0.67 \\ w_2 &\approx 0.7719. \end{aligned}$$



Kuva 6: Preference relation $a \sim c \succ b$ can be modeled with the MLE but not with the Choquet integral.

Since for all alternatives $u_1 < u_2$, the weight w_1 has no effect on the final score of the alternatives:

$$C_\mu(a) = 0.6966$$
$$C_\mu(b) = 0.7277$$
$$C_\mu(c) = 0.6966$$

and the Choquet integral gives the alternative b just a slightly better ranking than the other two alternatives.

Based on this example we know, that it is indeed possible to find a situation where using multilinear forms or the Choquet integral actually produce different preference relations even though the differences achieved were small. On the other hand, using the same weights for both aggregation functions in the example situation caused relatively large changes in score differences.

4.4 Qualitative comparison

The two aggregation functions that first did not seem too similar have proven to, not only to be able to allow for interaction between criteria, but also to produce rather similar decisions.

In decision making problems and related aggregation functions, the actual numeric values for alternatives are not that significant but the relative differences between the values of the alternatives. Saying that the value for alternative a is 0.4 does not describe the alternative at all, but relative preference is strongly influenced by values of other alternatives. Weather the value for alternative b would be 0.2 or 0.9 has significant influence on the degree of preference of alternative a.

It has been established, that both the Choquet integral and the MLE can be expressed with the same weights that are the Möbius transform of a capacity. Earlier in Section 2.5 we introduced the importance and interaction indices for capacities and since the Möbius transforms can be transformed back to capacity with the inverse transform (3), these indices are applicable to both functions. However, it is noteworthy that even though the two aggregation functions can be expressed with the same weighting coefficients, the same preference order is not always achieved with such weights. An example of such situation was presented in Section 4.3.

Both aggregation functions require determining an exponentially growing number of weights, unless some special case is used. Using the multiplicative forms is an effective way to reduce the number of weights, which might be desirable especially in situations with large number of criteria. For the Choquet integral we represented the concept of the k-additivity, which reduced the number of weights by assuming the Möbius transforms for sets larger than k would be zero. One could speculate if the same assumption for the multilinear forms would as well result in a useful sub-model.

5 Conclusion

The objective of this thesis was to comprehensively introduce two aggregation functions, the Choquet integral and the Multilinear forms, and present how these functions can be utilized in multi-criteria decision making problems. Also a qualitative comparison between these two was conducted.

In Section 2 we introduced the Choquet integral, which is a more flexible aggregation function, that is able to model the interaction between evaluation criteria. As a downside for flexibility, the number of coefficients to be determined grows exponentially when the amount of criteria grows, and thus the interpretation of these coefficients gets more complicated.

The selection of weighting coefficients affects the flexibility of the model; with certain assumptions regarding the capacity, and thus the coefficients, the Choquet integral can be reduced to more simple aggregation functions such as the weighted arithmetic mean and the ordered weighted average. The sub-models represent a trade-off between the challenges of determining the coefficients and complexity of the model. To offer a compromise between the difficulty of determining the coefficients and complexity of the model, concept of k-additive capacities was introduced, of which particularly 2-additive capacities have been found useful.

In Section 3 we introduced the multilinear form, an multilinear aggregation function that can be utilized in decision making problems. Similarly as the Choquet integral, also multilinear forms can be used to evaluate alternatives with respect to nonindependent criteria. When the dimension of the problem increases, the number of weighting coefficients grows, thus complicating the model construction. Some simple ways to reduce the number of weights with the cost of the flexibility of the model were introduced: with certain coefficients the multilinear form collapses into the weighted arithmetic mean. The multiplicative model reduces the amount of weights to $n + 1$ without completely removing the possibility for criteria interaction.

Both these functions have different premises but are applicable in similar situations; in Section 4 the differences and similarities of these functions were discussed. The functions have a mathematical connection and it is possible to use the same weighting coefficients with both functions. It is noteworthy, that even with same weights, same preference relations are not always achieved and thus aggregation function selection might affect the decisions.

Utilizing the Choquet integral and multilinear forms in multi-criteria decision making problems eliminates the need for non-interacting evaluation criteria, thus providing a more flexible way to model the decision makers preferences. As it is possible to represent slightly different preference relations with the Choquet integral than the multilinear forms, and vice versa, it would be interesting to, for example, conduct sensitivity analysis to research whether the differences between these two functions have any significance in practice. Probably the largest problem with these functions was the exponentially growing number of weighting coefficients, thus making different sub-models and special cases useful topics for future research.

Viitteet

G. Beliakov, A. Pradera, and T. Calvo. *Aggregation Functions: A Guide for Practitioners.* Springer Publishing Company, Incorporated, 1st edition, 2008.

G. Choquet. Theory of capacities. In *Annales de l'institut Fourier*, volume 5, pages 131–295, 1954.

M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89:445–456, 1996.

M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92:167–189, 1997. Fuzzy Measures and Integrals.

M. Grabisch and C. Labreuche. A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *Annals of Operations Research*, 175:247–286, 2010.

M. Grabisch, J. Marichal, R. Mesiar, and E. Pap. *Aggregation functions*, volume 127. Cambridge University Press, 2009.

R. Keeney and H. Raiffa. *Decisions with multiple objectives: preferences and value trade-offs.* Cambridge University Press, 1993.

J. Marichal. Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. *European Journal of Operational Research*, 155:771–791, 2004.

T. Murofushi and M. Sugeno. An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure. *Fuzzy Sets and Systems*, 29:201–227, 1989.

G. Owen. Multilinear extensions of games. *Management Science*, 18(5):64–79, 1972.

L.S. Shapley. A value for n-person games. Technical report, DTIC Document, 1952.

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

D. Von Winterfeldt and W. Edwards. *Decision analysis and behavioral research*, volume 604. Cambridge University Press, 1986.

M. Zeleny. *Multiple criteria decision making*, volume 25. McGraw-Hill New York, 1982.

A Summary in Finnish

Monikriteeripäätöksenteossa eri päätösvaihtoehtoja vertaillaan useiden eri kriteerien avulla, ja näiden perusteella pyritään määrittämään vaihtoehtojen välinen paremmuusjärjestys. Tällöin kysymykseksi nousevatkin eri tavat kriteerien yhdistelyyn: millä menetelmällä saadaan laskettua vaihtoehdoille vertailukelpoiset arvot, jotka kuvastavat päätöksentekijän preferenssejä?

Kriteerien saamien arvojen yhdistelyyn käytettävät funktiot, aggregaatiofunktiot, ovat usein yksinkertaisia keskiarvofunktioita. Tyypillisesti päätöksenteon ongelmissa käytetään paljon aikaa kriteeriarvojen määrittämiseen, ja aggregaatiofunktiona päädytään käyttämään painotettua aritmeettista keskiarvoa, joka onkin yleisimmin käytetty aggregaatiofunktio. Tällaisten funktioiden käyttö kuitenkin edellyttää, että käytetyt arviointikriteerit ovat toisistaan riippumattomia. Riippumattomien arviointikriteerien löytäminen on usein vaikeaa, ellei jopa mahdotonta, ja virheellinen oletus riippumattomista arviointikriteereistä saattaa johtaa tilanteeseen, jossa mallin antamat päätössuositukset eivät vastaa päätöksentekijän preferenssejä.

Kriteeriarvojen yhdistämiseen käytettävän aggregaatiofunktion valinnalla on mahdollista vaikuttaa mallin joustavuuteen ja ominaisuuksiin. Tämän kandidaatintyön tavoitteena oli esitellä kaksi erilaista aggregaatiofunktiota, Choquet-integraali (Choquet integral) ja monilineaarinen malli (multilinear forms), joiden käyttö ei vaadi riippumattomien arviointikriteerien käyttöä. Tällaiset funktiot antavat joustavuutta päätöksenteon ongelmiin, sillä niillä on esimerkiksi mahdollista mallintaa tilanteita, joissa osa käytettävistä kriteereistä ovat redundantteja.

Sekä Choquet-integraali että monilineaarinen malli voidaan ajatella eräänlaisiksi aritmeettisen keskiarvon laajennuksiksi. Tämä tarkoittaa, että sopivasti valituilla painokertoimilla kumpikin funktioista palautuu aritmeettiseksi painotetuksi keskiarvoksi. Toisaalta voidaan siis ajatella, että mikäli painokerrointen määritys tehdään onnistuneesti, funktiot antavat vähintään yhtä hyviä tuloksia kuin painotettu aritmeettinen keskiarvo. Kirjallisuuskatsauksen lisäksi työssä vertailtiin ja esiteltiin funktioiden välisiä yhtäläisyyksiä ja eroja laadullisella tasolla.

Vaikka Choquet-integraali ja monilineaarinen malli ovat lähteneet liikkeelle varsin erilaisista lähtökohdista, on niiden perusajatus sama: sen lisäksi, että määritellään painokertoimet kuvastamaan yksittäisten kriteerien tärkeyttä, määritellään painokertoimet myös kaikille mahdollisille kriteerijoukoille. Yhtäläistä funktioille on myös, että samoja painokertoimia on mahdollista käyttää mallin rakentamisessa valitusta funktiosta riippumatta. Yleisesti päätöksenteon ongelmissa vaaditaan, että käytetty aggregaatiofunktio on monotonisesti kasvava, mikä asettaa rajoitteita mallin parametrien määrittämiseen. Choquet-integraali ja monilineaarinen malli ovat monotonisesti kasvavia, jos painokertoimet määritetään *kapasiteetin* avulla (Capacity). Kapasiteetti on monotoninen normalisoitu mitta.

Vaikka mallit voidaankin rakentaa samoja painokertoimia käyttämällä, funktiovalinnalla on vaikutus saatuihin tuloksiin. Funktiot mallintavat kriteerien välisen vuorovaikutuksen eri tavoin: Choquet-integraalissa vuorovaikutustermit pohjautuvat minimioperaattoriin, kun taas monilineaarisissa malleissa vuorovaikutus syntyy tulotermeistä. Tämän vuoksi niiden avulla on teoriassa mahdollista mallintaa hieman erilaisia preferenssirelaatioita, mutta saatavat erot ovat käytännön sovelluksissa varsin pieniä. Yksi tärkeä ero funktioiden välillä on, että monilineaariset mallit ovat jatkuvia funktioita, toisin kuin paloittain lineaarien Choquet-integraali.

Choquet-integraali ja monilineaarinen malli antavat joustavuutta ja helpottavat päätösmallien rakentamista. Vastapainona funktioiden tarjoamiin hyötyihin on, että määritettävien painokertoimien lukumäärä kasvaa eksponentiaalisesti, kun käytettävien arviointikriteerien määrää kasvatetaan; kymmenen päätöskriteerin tehtävässä määritettävänä on jo 1022 painokerrointa. Kertoimien suuri määrä hankaloittaa mallien rakentamista, eikä kumpikaan funktioista sellaisenaan ole aina paras mahdollinen valinta. Suuri parametrimäärä tekee funktioista mahdottomia käyttää suurten päätösongelmien ratkaisemisessa, mutta tekemällä erilaisia oletuksia määritettävistä painokertoimista on mahdollista vähentää määritettävien kertoimien määrää. Esimerkiksi minimi, maksimi ja, kuten todettu, painotettu aritmeettinen keskiarvo on mahdollista saavuttaa Choquet-integraalin erikoistapauksina. Mainitut funktiot eivät kuitenkaan ole kovinkaan joustavia. Esimerkiksi painotettu aritmeettinen keskiarvo saadaan olettamalla, että kriteerien välistä vuorovaikutusta ei ole lainkaan. Parametrien määrää on mahdollista vähentää myös joustavammin, esimerkiksi olettamalla, että kriteerien välinen vuorovaikutus suurissa kriteerijoukoissa ei ole merkittävää, tai määrittelemällä suurien kriteerijoukkojen painokertoimet yksittäisten kriteerien kertoimien avulla.

Mallit ja niiden käyttäytyminen pohjautuvat painokertoimien määrittämiseen; erilaisilla painokertoimilla mallit käyttäytyvät eri tavoin ja muodostavat erilaisia päätösmalleja. Painotetun keskiarvon tapauksessa eri kriteerien vaikutus ja tärkeys heijastuu suoraan kriteereille annetuista painokertoimista, mutta Choquet-integraalin ja monilineaarisen mallin kanssa eri kriteerien painoarvot eivät enää suoraan kuvaa kriteerin painoarvoa mallissa. Kriteerien tärkeyden kuvaamiseen on olemassa erilaisia tärkeysmittoja, joista tässä työssä esitellään Shapleyn indeksi (Shapley value, importance index). Shapleyn indeksi on eräänlainen painokeskiarvo, jonka avulla on mahdollista arvioida kriteerin kokonaisvaikutusta mallin käyttäytymiseen. Esimerkiksi kahden osittain redundantin kriteerin Shapleyn indeksit ovat niiden painokertoimia pienemmät, sillä negatiivinen vuorovaikutus heikentää niiden tärkeyttä mallissa.

Choquet-integraalin ja monilineaarisen mallin tärkein ominaisuus on mahdollisuus mallintaa kriteerien välistä vuorovaikutusta, joten parametrien ja siten mallin arvioinnin kannalta myös vuorovaikutuksien tutkiminen on mielenkiintoista. Kahden kriteerin välisen vuorovaikutuksen voimakkuutta voidaan kuvata vuorovaikutusindeksillä (interaction index). Positiivinen vuorovaikutusindeksi kertoo, että kriteerien välillä on synergiaa, kun taas negatiivinen indeksiarvo saadaan, kun kriteerit ovat osittain redundantteja.

Choquet-integraali ja monilineaariset mallit tarjoavat joustavuutta monikriteeripäätöksenteon ongelmiin, ja ovat siksi mielenkiintoinen tutkimuskohde. Ne helpottavat käytettävien päätöskriteereiden valintaa, sillä niiden avulla pystytään mallintamaan myös toisistaan riippuvien kriteerien vaikutuksia vaihtoehdon mielekkyyteen. Funktioiden käytön suurimpana haasteena on kuitenkin kasvava painokertoimien määrä. Päätösongelmissa, joissa vaihtoehtoja halutaan arvioida useiden eri kriteerien perusteella, painokertoimien suuri määrä vaikeuttaa funktioiden hyödyntämistä. Tämän vuoksi erityisesti tehokkaat, mutta kuitenkin joustavat alimallit saattavat tulevaisuudessa haastaa painotetun keskiarvon aseman aggregaatiofunktioiden joukossa.