Aggregation functions: Means

Michel Grabisch, Jean-Luc Marichal, Radko Mesiar, Endre Pap

Aggregation Functions. Part I: Means

Michel GRABISCH
University of Panthéon-Sorbonne
Paris, France
Jean-Luc MARICHAL
University of Luxembourg
Luxembourg
Radko MESIAR
Slovak University of Technology
Bratislava, Slovakia,
Endre PAP
University of Novi Sad
Novi Sad, Serbia

November 13, 2009

Abstract

The two-parts state-of-art overview of aggregation theory summarizes the essential information concerning aggregation issues. Overview of aggregation properties is given, including the basic classification of aggregation functions. In this first part, the stress is put on means, i.e., averaging aggregation functions, both with fixed arity (n-ary means) and with open arity (extended means).

1 Introduction

Aggregation functions play an important role in many of the technological tasks scientists are faced with nowadays. They are specifically important in many problems related to the fusion of information. More generally, aggregation functions are widely used in pure mathematics (e.g., functional equations, theory of means and averages, measure and integration theory), applied mathematics (e.g., probability, statistics, decision mathematics), computer and engineering sciences (e.g., artificial intelligence, operations research, information theory, engineering design, pattern recognition and image analysis, data fusion, automated reasoning), economics and finance (e.g., game theory, voting theory, decision making), social sciences (e.g., representational measurement, mathematical psychology) as well as many other applied fields of physics and natural sciences. Thus, a main characteristic of the aggregation functions is that they are used in a large number of areas and disciplines.

The essence of aggregation is that the output value computed by the aggregation function should represent or synthesize "in some sense" all individual inputs, where quotes are put to emphasize the fact that the precise meaning of this expression is highly dependent on the context. In any case, defining or choosing the right class of aggregation functions for a specific problem is a difficult task, considering the huge variety of potential aggregation functions. In this respect, one could ask the following question: Consider a set of n values, lying in some real interval $[a, b]$ to be aggregated. Is any function from $[a, b]^n$ to $\mathbb{R}$ a candidate to be an aggregation function? Obviously not. On the other hand, it is not that easy to define the minimal set of properties a function should fulfill to be an aggregation function. A first natural requirement comes from the fact that the output should be a synthetic value. Then, if inputs are supposed to lie in the interval $[a, b]$, the output should also lie in this interval. Moreover, if all input values are identical to the lower bound a, then the output should also be a, and similarly for the case of the upper bound b. This defines a boundary condition. A second natural requirement is nondecreasing monotonicity. It means that if some of the input values increase, the representative output value should reflect this increase, or at worst, stay constant. These two requirements are commonly accepted in the field, and we adopt them as the basic definition of an aggregation function.

Thus defined, the class of aggregation functions is huge, making the problem of choosing the right function (or family) for a given application a difficult one. Besides this practical consideration, the study of the main classes of aggregation functions, their properties and their relationships, is so complex and rich that it becomes a mathematical topic of its own.

A solid mathematical analysis of aggregation functions, able to answer both mathematical and practical concerns, was the main motivation for us to prepare a monograph [**?**]. From related recent monographs, recall the handbook [**?**] and [**?**]. The aim of these two-parts invited state-of-art overview is to summarize the essential information about aggregation functions for Information Sciences readers, to open them the door to the rich world of tools important for information fusion. With a kind permition of the publisher, some parts of [**?**] were used in this manuscript. Moreover, to increase the transparentness, proofs of several introduced results are not included (or sketched only), however, for interested readers always an indication where the full proofs can be found is given.

The paper is organized as follows. In the next section, basic properties of aggregation functions and several illustrative examples are given. Section 3 is devoted to means related to the arithmetic mean and means with some special properties. In Section 4, non-additive integral - based aggregation functions are discussed, stressing a prominent role of the Choquet and Sugeno integrals. Part I ends with some concluding remarks. In Part II, Section 2 deals with conjunctive aggregation functions, especially with triangular norms and copulas. In Section 3, disjunctive aggregation functions are summarized, exploring their duality to conjunctive aggregation functions. Moreover, several kinds of aggregation functions mixing conjunctive and disjunctive aggregation functions, are also included (uninorms, nullnorms, gamma operators, etc.). Several construction methods for aggregation functions are introduced in Section 4. These methods are not only a summarization showing how several kinds of aggregation functions were introduced from simpler ones, but they allow to readers to taylor their aggregation model when solving nonstandard problems resisting to standard aggregation functions being used and fit to real data constraints. Finally, several concluding remarks are included.

2 Basic definitions and examples

2.1 Aggregation functions

Aggregation functions are special real functions with inputs from a subdomain of the extended real line $\overline{\mathbb{R}} = [-\infty, \infty]$. We will deal with interval $\mathbb{I}$ domains, independently of their type (open, closed, ...). The framework of aggregation functions we deal with is constrained by the next definition, see also [**?**, **?**, **?**, **?**] for the case of closed domain $\mathbb{I}$.

Definition 1. An *aggregation function* in $\mathbb{I}^n$ is a function $\mathsf{A}^{(n)} : \mathbb{I}^n \to \mathbb{I}$ that

(i) is nondecreasing (in each variable)

(ii) fulfills the boundary conditions

$$\inf_{\mathbf{x} \in \mathbb{I}^n} \mathsf{A}^{(n)}(\mathbf{x}) = \inf \mathbb{I} \quad \text{and} \quad \sup_{\mathbf{x} \in \mathbb{I}^n} \mathsf{A}^{(n)}(\mathbf{x}) = \sup \mathbb{I}. \tag{1}$$

An *extended aggregation function* in $\cup_{n\in\mathbb{N}} \mathbb{I}^n$ is a mapping $\mathsf{A} : \bigcup_{n\in\mathbb{N}} \mathbb{I}^n \to \overline{\mathbb{R}}$ whose restriction $\mathsf{A}^{(n)} := \mathsf{A}|_{\mathbb{I}^n}$ to $\mathbb{I}^n$ is an aggregation function in $\mathbb{I}^n$ for any $n \in \mathbb{N}$.

Due to this definition, $\text{Ran}\mathsf{A} \subset \mathbb{I}$ and thus we will consider $\mathsf{A} : \bigcup_{n\in\mathbb{N}} \mathbb{I}^n \to \mathbb{I}$ ($\mathsf{A}^{(n)} : \mathbb{I}^n \to \mathbb{I}$) in the rest of the paper.

The integer n represents the arity of the aggregation function, that is, the number of its variables. When no confusion can arise, the aggregation functions will simply be written A instead of $\mathsf{A}^{(n)}$.

To give a first instance, the arithmetic mean, defined by

$$\mathsf{AM}^{(n)}(\mathbf{x}) := \frac{1}{n} \sum_{i=1}^{n} x_i \,, \tag{2}$$

is clearly an extended aggregation function in any domain $\mathbb{I}^n$ (if $\mathbb{I} = \overline{\mathbb{R}}$, convention $+\infty + (-\infty) = -\infty$ is often considered). Moreover, ternary function $\mathsf{AM}^{(3)}$ is an example of a ternary aggregation function. Among other basic (extended) aggregation functions recall:

(i) the product $\prod(\mathbf{x}) = \prod_{i=1}^n x_i$ ($\mathbb{I} \in \{|0,1|, |0,\infty|, |1,\infty|\}$, where $|a,b|$ means any of four kinds of intervals, with boundary points a and b, and with convention $0 \cdot \infty = 0$;

(ii) the geometric mean $\mathsf{GM}(\mathbf{x}) = \left(\prod_{i=1}^n x_i\right)^{1/n}$ ($\mathbb{I} \subset [0,\infty], 0 \cdot \infty = 0$);

(iii) the minimum function $\mathsf{Min}(\mathbf{x}) = \min\{x_1, \ldots, x_n\}$ (arbitrary $\mathbb{I}$);

(iv) the maximum function $\mathsf{Max}(\mathbf{x}) = \max\{x_1, \ldots, x_n\}$ (arbitrary $\mathbb{I}$);

(v) the sum function $\sum(\mathbf{x}) = \sum_{i=1}^n x_i$ ($\mathbb{I} \in \{|0,\infty|, |-\infty, 0|, |-\infty,\infty|\}$, $+\infty + (-\infty) = -\infty$).

Based on many valued logics connectives [**?**, **?**] we have the next classification of aggregation functions.

Definition 2. Consider an (extended) aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ ($\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$). Then

(i) A is called conjunctive whenever $\mathsf{A} \leqslant \mathsf{Min}$, i.e., $\mathsf{A}(\mathbf{x}) \leqslant \mathsf{Min}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{I}^n$ ($\mathbf{x} \in \cup_{n\in\mathbb{N}}\mathbb{I}^n$).

(ii) A is called disjunctive whenever $\mathsf{A} \geqslant \mathsf{Max}$.

(iii) A is called internal whenever $\mathsf{Min} \leqslant \mathsf{A} \leqslant \mathsf{Max}$.

(iv) A is called mixed if it is not conjunctive neither disjunctive nor internal.

In particular case $\mathbb{I} = [0,1]$ (or $\mathbb{I} = [a,b] \subset \mathbb{R}$), the standard duality of aggregation functions is introduced.

Definition 3. Let $\mathsf{A} : [0,1]^n \to [0,1]$ be an aggregation function. Then the function $\mathsf{A}^d : [0,1]^n \to [0,1]$ given by

$$\mathsf{A}(\mathbf{x}) = 1 - \mathsf{A}(1 - x_1, \ldots, 1 - x_n) \tag{3}$$

is called a dual aggregation function (to A).

Evidently, A^d given by (3) is an aggregation function on $[0,1]$. Similarly, dual extended aggregation function A^d to A acting on $[0,1]$ can be introduced. If $\mathbb{I} = [a,b] \subset \mathbb{R}$, then (3) should be modified into

$$\mathsf{A}^d(\mathbf{x}) = a + b - \mathsf{A}(a + b - x_1, \ldots, a + b - x_n).$$

It is evident that dual to a conjunctive (respectively disjunctive, mixed) aggregation function is disjunctive (respectively conjunctive, mixed) aggregation function.

Note that many properties defined for n-ary functions can be naturally adapted to extended functions. For instance, with some abuse of language, the extended function $\mathsf{F} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \overline{\mathbb{R}}$ is said to be continuous if, for any $n \in \mathbb{N}$, the corresponding n-ary function $\mathsf{F}^{(n)} = \mathsf{F}|_{\mathbb{I}^n}$ is continuous. These adaptations are implicitly assumed throughout, for example in sections 2.2, 2.3. Properties defined for extended aggregation functions only will be stressed explicitly (note that these properties make an important link between aggregation functions with fixed but different arities). In some cases, properties of general real functions will be introduced.

2.2 Monotonicity properties

Definition 4. The function $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ is *strictly increasing* (in each argument) if, for any $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$,

$$\mathbf{x} < \mathbf{x}' \quad \Rightarrow \quad \mathsf{F}(\mathbf{x}) < \mathsf{F}(\mathbf{x}').$$

Thus, a function is strictly increasing if it is nondecreasing and if it presents a positive reaction to any increase of at least one input value.

An intermediate kind of monotonicity (between nondecreasingness and strict increasigness) is the unanimeonus increasigness, also called jointly strict increasingness.

Definition 5. The function $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ is *unanimously increasing* if it is nondecreasing and if, for any $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$,

$$x_i < x_i', \ \forall i \in \{1, \ldots, n\} \quad \Rightarrow \quad \mathsf{F}(\mathbf{x}) < \mathsf{F}(\mathbf{x}').$$

Clearly, strict increasing monotonicity ensures unanimous increasing monotonicity. For example, the arithmetic mean AM is strictly increasing, hence unanimously increasing. Functions Min and Max are unanimously increasing but not strictly increasing. The product Π is unanimously increasing on $[0,1]^n$. However, if 0 occurs among inputs, the strict monotonicity of Π is violated. The *bounded sum* $\mathsf{S}_{\mathbf{L}}(\mathbf{x}) = \mathsf{Min}(\sum_{i=1}^n x_i, 1)$ on $[0,1]^n$ is nondecreasing but not unanimously increasing.

2.3 Continuity properties

As already mentioned, the continuity of an extended aggregation function A means the classical continuity of all n-ary functions $\mathsf{A}^{(n)}$. The same holds for the other kinds of continuity which are therefore introduced for n-ary functions only. We recall only few of them, more details can be found in [**?**], section 2.2.2.

The continuity property can be strengthened into the well-known Lipschitz condition [**?**]; see Zygmund [**?**].

Definition 6. Let $\| \cdot \| : \mathbb{R}^n \to [0, \infty[$ be a norm. If a function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ satisfies the inequality

$$|\mathsf{F}(\mathbf{x}) - \mathsf{F}(\mathbf{y})| \leqslant c\, \|\mathbf{x} - \mathbf{y}\| \qquad (\mathbf{x}, \mathbf{y} \in \mathbb{I}^n), \tag{4}$$

for some constant $c \in \,]0, \infty[$, then we say that F satisfies the *Lipschitz condition* or is *Lipschitzian* (with respect to $\| \cdot \|$). More precisely, any function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ satisfying (4) is said to be *c-Lipschitzian*. The greatest lower bound d of constants $c > 0$ in (4) is called the *best Lipschitz constant* (which means that F is d-Lipschitzian but, for any $u \in \,]0, d[$, F is not u-Lipschitzian).

Important examples of norms are given by the Minkowski norm of order $p \in [1, \infty[$, namely

$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p},$$

also called the L_p-norm, and its limiting case $\|\mathbf{x}\|_\infty := \max_i |x_i|$, which is the Chebyshev norm.

The c-Lipschitz condition has an interesting interpretation when applied in aggregation. It allows us to estimate the relative output error in comparison with input errors

$$|\mathsf{F}(\mathbf{x}) - \mathsf{F}(\mathbf{y})| \leqslant c\,\varepsilon$$

whenever $\|\mathbf{x} - \mathbf{y}\| \leqslant \varepsilon$ for some $\varepsilon > 0$.

We also have the following result.

Proposition 1. For arbitrary reals $p, q \in [1, \infty]$, a function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ is Lipschitzian with respect to the norm $\| \cdot \|_p$ if and only if it is Lipschitzian with respect to the norm $\| \cdot \|_q$. Moreover, if F is c-Lipschitzian with respect to the norm $\| \cdot \|_p$ then it is also c-Lipschitzian with respect to the norm $\| \cdot \|_q$ for $q \leqslant p$, and for $q > p$ it is nc-Lipschitzian.

In the theory of aggregation functions, when speaking about Lipschitz property without mentioning the norm explicitly, always L_1-norm is considered. Observe that if $\mathbb{I} \subset \mathbb{R}$ then the Lipschitz property (with respect to any norm) implies the continuity but not vice-versa (consider the geometric mean GM on $\mathbb{I} = [0, 1]$, for example).

Proposition 2. Let $[a, b]$ be a real interval. The smallest and the greatest aggregation functions defined in $[a, b]^n$ that are 1-Lipschitzian with respect to the norm $\| \cdot \|$ are respectively given by $\mathsf{A}_*^{(n)} : [a, b]^n \to [a, b]$, with

$$\mathsf{A}_*^{(n)}(\mathbf{x}) := \mathsf{Max}(b - \|n \cdot b - \mathbf{x}\|, a),$$

and $\mathsf{A}^{*(n)} : [a, b]^n \to [a, b]$, with

$$\mathsf{A}^{*(n)}(\mathbf{x}) := \mathsf{Min}(a + \|\mathbf{x} - n \cdot a\|, b).$$

Note that if $[a, b] = [0, 1]$ and L_1-norm, $\mathsf{A}_* = \mathsf{T}_L$ (Lukasiewicz t-norm), where $\mathsf{T}_L(\mathbf{x}) = \max(0, \sum_{i=1}^n x_i - (n - 1)$, and $\mathsf{A}^* = \mathsf{S}_L$ (bounded sum).

Example 1. The arithmetic mean $\mathsf{AM} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{R}$ is 1-Lipschitzian (with respect to the L_1-norm) independently of the interval $\mathbb{I}$. For each $n \in \mathbb{N}$, the best Lipschitz constant for $\mathsf{AM}^{(n)}$ is $1/n$ and $\mathsf{AM}^{(n)}$ is the only n-ary aggregation function having this property. With respect to the L_∞-norm, $\mathsf{AM}^{(n)}$ is 1-Lipschitzian for all n, 1 being the best Lipschitz constant.

The extended aggregation function $\mathsf{Q} : \cup_{n\in\mathbb{N}}[0, 1]^n \to [0, 1]$ given by $\mathsf{Q}(\mathbf{x}) := \prod_i x_i^i$ is not Lipschitzian (with respect to the L_1-norm), though each $\mathsf{Q}^{(n)}$ is Lipschitzian (the best Lipschitz constant for $\mathsf{Q}^{(n)}$ is n).

We have also two next weaker forms of continuity (using the lattice notation $\vee$ for supremum and $\wedge$ for infimum.

Definition 7. A nondecreasing function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ is called *lower semi-continuous* or *left-continuous* if, for all $(\mathbf{x}^{(k)})_{k\in\mathbb{N}} \subset (\mathbb{I}^n)^\mathbb{N}$ such that $\vee_k \mathbf{x}^{(k)} \in \mathbb{I}^n$,

$$\bigvee_{k\in\mathbb{N}} \mathsf{F}(\mathbf{x}^{(k)}) = \mathsf{F}\Big(\bigvee_{k\in\mathbb{N}} \mathbf{x}^{(k)}\Big).$$

Definition 8. A nondecreasing function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ is called *upper semi-continuous* or *right-continuous* if, for all $(\mathbf{x}^{(k)})_{k\in\mathbb{N}} \subset (\mathbb{I}^n)^\mathbb{N}$ such that $\wedge_k \mathbf{x}^{(k)} \in \mathbb{I}^n$,

$$\bigwedge_{k\in\mathbb{N}} \mathsf{F}(\mathbf{x}^{(k)}) = \mathsf{F}\Big(\bigwedge_{k\in\mathbb{N}} \mathbf{x}^{(k)}\Big).$$

We have the following important result; see Klement et al. [**?**, Proposition 1.22].

Proposition 3. An aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{R}$ is lower semi-continuous (respectively, upper semi-continuous) if and only if A is lower semi-continuous (respectively, upper semi-continuous) in each variable.

Proposition 4. An aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{R}$ is continuous if and only if it is both lower and upper semi-continuous.

Example 2. (i) An important example of a left-continuous (lower semi-continuous) but noncontinuous aggregation function is the nilpotent minimum $\mathsf{T}^{\mathbf{nM}} : [0,1]^2 \to [0,1]$,

$$\mathsf{T}^{\mathbf{nM}}(x_1, x_2) := \begin{cases} \mathsf{Min}(x_1, x_2) & \text{if } x_1 + x_2 > 1 \\ 0 & \text{otherwise.} \end{cases}$$

(ii) The drastic product $\mathsf{T}_{\mathbf{D}} : [0,1]^n \to [0,1]$, given by

$$\mathsf{T}_{\mathbf{D}}(\mathbf{x}) := \begin{cases} \mathsf{Min}(\mathbf{x}) & \text{if } |\{i \in \{1, \ldots, n\} \mid x_i < 1\}| < 2 \\ 0 & \text{otherwise,} \end{cases}$$

is a noncontinuous but upper semi-continuous aggregation function.

2.4 Symmetry

The next property we consider is *symmetry*, also called *commutativity*, *neutrality*, or *anonymity*. The standard commutativity of binary operations $x*y = y*x$, well known in algebra, can be easily generalized to n-ary functions, with $n \geqslant 2$, as follows.

Definition 9. $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ is a *symmetric* function if

$$\mathsf{F}(\mathbf{x}) = \mathsf{F}(\sigma(\mathbf{x}))$$

for any $\mathbf{x} \in \mathbb{I}^n$ and for any permutation σ of $(1, \ldots, n)$, where $\sigma(\mathbf{x}) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$.

The symmetry property essentially means that the aggregated value does not depend on the order of the arguments. This is required when combining criteria of equal importance or anonymous experts' opinions.[1]

Many aggregation functions introduced thus far are symmetric. For example, AM, GM, S_L, T_L, $\prod$, Min, Max are symmetric functions. A prominent example of non-symmetric aggregation functions is the weighted arithmetic mean $\mathsf{WAM}_{\mathbf{w}}$, $\mathsf{WAM}_{\mathbf{w}}(x_1, \ldots, x_n) = \sum_{i=1}^n w_i x_i$, where the nonnegative weights w_i are constrained by $\sum_{i=1}^n w_i = 1$ (and at least one weight $w_i \neq \frac{1}{n}$, else $\mathsf{WAM}_{\mathbf{w}} = \mathsf{AM}$ is symmetric).

The following result, well-known in group theory, shows that the symmetry property can be checked with only two equalities; see for instance Rotman [**?**, Exercise 2.9 p. 24].

Proposition 5. $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ is a symmetric function if and only if, for all $\mathbf{x} \in \mathbb{I}^n$, we have

(i) $\mathsf{F}(x_2, x_1, x_3, \ldots, x_n) = \mathsf{F}(x_1, x_2, x_3, \ldots, x_n)$,

(ii) $\mathsf{F}(x_2, x_3, \ldots, x_n, x_1) = \mathsf{F}(x_1, x_2, x_3, \ldots, x_n)$.

This simple test is very efficient, especially when symmetry does not appear immediately, like in the 4-variable expression

$$(x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_4) \vee (x_1 \wedge x_3 \wedge x_4) \vee (x_2 \wedge x_3 \wedge x_4),$$

which is nothing other than the 4-ary order statistic $x_{(2)}$.

[1] Of course, symmetry is more natural in voting procedures than in multicriteria decision making, where criteria usually have different importances.

2.5 Idempotency

Definition 10. $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ is an *idempotent* function if $\delta_{\mathsf{F}} = \mathrm{id}$, that is,

$$\mathsf{F}(n \cdot x) = x \qquad (x \in \mathbb{I}).$$

Idempotency is in some areas supposed to be a natural property of aggregation functions, e.g., in multicriteria decision making (see for instance Fodor and Roubens [**?**]), where it is commonly accepted that if all criteria are satisfied at the same degree x, implicitly assuming the commensurateness of criteria, then also the overall score should be x.

It is evident that AM, $\mathsf{WAM}_{\mathbf{w}}$, Min, Max, and Med are idempotent functions, while Σ and Π are not. Recall also that any nondecreasing and idempotent function $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ is an aggregation function.

2.5.1 Idempotent elements

Definition 11. An element $x \in \mathbb{I}$ is *idempotent* for $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ if $\delta_{\mathsf{F}}(x) = x$.

In $[0,1]^n$ the product Π has no idempotent elements other than the extreme elements 0 and 1. As an example of a function in $[0,1]^n$ which is not idempotent but has a non-extreme idempotent element, take an arbitrarily chosen element $c \in]0,1[$ and define the aggregation function $\mathsf{A}_{\{c\}} : [0,1]^n \to [0,1]$ as follows:

$$\mathsf{A}_{\{c\}}(\mathbf{x}) := \mathsf{Med}\Big(0, c + \sum_{i=1}^{n} (x_i - c), 1\Big),$$

where Med is the standard median function (i.e., in ternary case the "midle" input, between the smallest one and the greatest one). It is easy to see that the only idempotent elements for $\mathsf{A}_{\{c\}}$ are 0, 1, and c.

2.5.2 Strong idempotency

The idempotency property has been generalized to extended functions as follows; see Calvo et al. [**?**].

Definition 12. $\mathsf{F} : \cup_{n\in\mathbb{N}} \mathbb{I}^n \to \overline{\mathbb{R}}$ is *strongly idempotent* if, for any $n \in \mathbb{N}$,

$$\mathsf{F}(n \cdot \mathbf{x}) = \mathsf{F}(\mathbf{x}) \qquad (\mathbf{x} \in \cup_{m\in\mathbb{N}} \mathbb{I}^m).$$

For instance, if $\mathsf{F} : \cup_{n\in\mathbb{N}} \mathbb{I}^n \to \overline{\mathbb{R}}$ is strongly idempotent then we have

$$\mathsf{F}(x_1, x_2, x_1, x_2) = \mathsf{F}(x_1, x_2).$$

Proposition 6. Suppose $\mathsf{F} : \cup_{n\in\mathbb{N}} \mathbb{I}^n \to \overline{\mathbb{R}}$ is strongly idempotent. Then F is idempotent if and only if $\mathsf{F}(x) = x$ for all $x \in \mathbb{I}$.

According to our convention on unary aggregation functions, namely $\mathsf{A}(x) = x$ for all $x \in \mathbb{I}$, it follows immediately from the previous proposition that any strongly idempotent extended aggregation function is idempotent.

Example 3. (i) Let

$$\Delta = (w_{i,n} \mid n \in \mathbb{N}, i \in \{1, \dots, n\})$$

be the Pascal weight triangle, where

$$w_{i,n} = \frac{\binom{n-1}{i-1}}{2^{n-1}},$$

and let $\mathsf{WAM}_\Delta : \cup_{n\in\mathbb{N}}[0,1]^n \to [0,1]$ be the corresponding extended weighted arithmetic mean,

$$\mathsf{WAM}_\Delta(\mathbf{x}) = \sum_{i=1}^{n} w_{i,n} x_i.$$

Then WAM_Δ is idempotent but not strongly idempotent.

(ii) Projection to the first coordinate P_F is a nonsymmetric strongly idempotent extended aggregation function. It is a special extended weighted arithmetic mean, where

$$w_{i,n} = \begin{cases} 1 & \text{if } i = 1, \\ 0 & \text{else.} \end{cases}$$

(iii) Define an extended aggregation function $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ by

$$\mathsf{A}(\mathbf{x}) = \begin{cases} \mathsf{Min}(\mathbf{x}) & \text{if } n \text{ is odd,} \\ \mathsf{Max}(\mathbf{x}) & \text{if } n \text{ is even.} \end{cases}$$

Then A is symmetric and idempotent but not strongly idempotent.

2.6 Associativity

We consider first the associativity functional equation. Associativity of a binary operation $*$ means that $(x*y)*z = x*(y*z)$, so we can write $x*y*z$ unambiguously. If we write this binary operation as a two-place function $f(a,b) = a*b$, then associativity says that $f(f(a,b),c) = f(a,f(b,c))$. For general f, this is the *associativity functional equation*.

Definition 13. $\mathsf{F} : \mathbb{I}^2 \to \mathbb{I}$ is *associative* if, for all $\mathbf{x} \in \mathbb{I}^3$, we have

$$\mathsf{F}\big(\mathsf{F}(x_1,x_2),x_3\big) = \mathsf{F}\big(x_1,\mathsf{F}(x_2,x_3)\big). \tag{5}$$

A large number of papers deal with the associativity functional equation (5) even in the field of real numbers. In complete generality, its investigation naturally constitutes a principal subject of algebra. For a list of references see Aczél [**?**, Section 6.2] and Alsina et al. [**?**].

Basically, associativity concerns aggregation of only two arguments. However, as stated in the next definition, associativity can be extended to any finite number of arguments.

Recall that, for two vectors $\mathbf{x} = (x_1,\ldots,x_n)$ and $\mathbf{x}' = (x'_1,\ldots,x'_m)$, we use the convenient notation $\mathsf{F}(\mathbf{x},\mathbf{x}')$ to represent $\mathsf{F}(x_1,\ldots,x_n,x'_1,\ldots,x'_m)$, and similarly for more than two vectors. Also, if $\mathbf{x} \in \mathbb{I}^0$ is an empty vector then it is simply dropped from the function. For instance, $\mathsf{F}(\mathbf{x},\mathbf{x}') = \mathsf{F}(\mathbf{x}')$ and $\mathsf{F}(\mathsf{F}(\mathbf{x}),\mathsf{F}(\mathbf{x}')) = \mathsf{F}(\mathsf{F}(\mathbf{x}'))$.

Definition 14. $\mathsf{F} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is *associative* if $\mathsf{F}(x) = x$ for all $x \in \mathbb{I}$ and if

$$\mathsf{F}(\mathbf{x}, \mathbf{x}') = \mathsf{F}(\mathsf{F}(\mathbf{x}), \mathsf{F}(\mathbf{x}'))$$

for all $\mathbf{x}, \mathbf{x}' \in \cup_{n\in\mathbb{N}_0}\mathbb{I}^n$.

As the next proposition shows, associativity means that each subset of consecutive arguments can be replaced with their partial aggregation without changing the overall aggregation.

Proposition 7. $\mathsf{F} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is associative if and only if $\mathsf{F}(x) = x$ for all $x \in \mathbb{I}$ and

$$\mathsf{F}(\mathbf{x}, \mathsf{F}(\mathbf{x}'), \mathbf{x}'') = \mathsf{F}(\mathbf{x}, \mathbf{x}', \mathbf{x}'')$$

for all $\mathbf{x}, \mathbf{x}', \mathbf{x}'' \in \cup_{n\in\mathbb{N}_0}\mathbb{I}^n$.

Associativity is also a well-known algebraic property which allows one to omit "parentheses" in an aggregation of at least three elements. Implicit in the assumption of associativity is a consistent way of going unambiguously from the aggregation of n elements to $n+1$ elements, which implies that any associative extended function F is completely determined by its binary function $\mathsf{F}^{(2)}$. Indeed, by associativity, we clearly have

$$\mathsf{F}(x_1, \ldots, x_{n+1}) = \mathsf{F}\big(\mathsf{F}(x_1, \ldots, x_n), x_{n+1}\big).$$

For practical purpose we can start with the aggregation procedure before knowing all inputs to be aggregated. Additional input data are then simply aggregated with the current aggregated output.

Each associative idempotent extended function is necessary strongly idempotent. For a fixed arity $n > 2$, we can introduce the associativity as follows.

Definition 15. Let $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ be an n-ary function. Then it is associative if, for all $x_1, \ldots, x_{2n-1} \in \mathbb{I}$, we have

$$\begin{aligned} \mathsf{F}(\mathsf{F}(x_1, \ldots, x_n), x_{n+1}, \ldots, x_{2n-1}) &= \mathsf{F}(x_1, \mathsf{F}(x_2, \ldots, x_{n+1}), x_{n+2}, \ldots, x_{2n-1}) \\ &= \mathsf{F}(x_1, \ldots, x_{n-1}, \mathsf{F}(x_n, \ldots, x_{2n-1})). \end{aligned}$$

2.7 Decomposability

Introduced first in Bemporad [**?**, p. 87] in a characterization of the arithmetic mean, associativity of means has been used by Kolmogoroff [**?**] and Nagumo [**?**] to characterize the so-called mean values. More recently, Marichal and Roubens [**?**] proposed to call this property "decomposability" in order not to confuse it with classical associativity. Alternative names, such as *associativity with repetitions* or *weighted associativity*, could be naturally considered as well.

When symmetry is not assumed, it is necessary to rewrite this property in such a way that the first variables are not privileged. To abbreviate notations, for nonnegative integers m, n, we write $\mathsf{F}(m{\cdot}x, n{\cdot}y)$ what means the repetition of arguments, i.e., $\mathsf{F}(\underbrace{x, \ldots, x}_{m}, \underbrace{y, \ldots, y}_{n})$. We then consider the following definition.

Definition 16. $\mathsf{F} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is *decomposable* if $\mathsf{F}(x) = x$ for all $x \in \mathbb{I}$ and if

$$\mathsf{F}(\mathbf{x}, \mathbf{x}') = \mathsf{F}\big(k \cdot \mathsf{F}(\mathbf{x}), k' \cdot \mathsf{F}(\mathbf{x}')\big) \tag{6}$$

for all $k, k' \in \mathbb{N}_0$, all $\mathbf{x} \in \mathbb{I}^k$, and all $\mathbf{x}' \in \mathbb{I}^{k'}$.

By considering $k = 0$ or $k' = 0$ in (6), we see that any decomposable function is range-idempotent. Moreover, as the following proposition shows, decomposability means that each element of any subset of consecutive arguments can be replaced with their partial aggregation without changing the overall aggregation.

Proposition 8. $\mathsf{F} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is decomposable if and only if $\mathsf{F}(x) = x$ for all $x \in \mathbb{I}$ and

$$\mathsf{F}\big(\mathbf{x}, k' \cdot \mathsf{F}(\mathbf{x}'), \mathbf{x}''\big) = \mathsf{F}(\mathbf{x}, \mathbf{x}', \mathbf{x}'')$$

for all $k' \in \mathbb{N}_0$, all $\mathbf{x}' \in \mathbb{I}^{k'}$, and all $\mathbf{x}, \mathbf{x}'' \in \cup_{n\in\mathbb{N}_0}\mathbb{I}^n$.

2.8 Bisymmetry

Another grouping property is the bisymmetry.

Definition 17. $\mathsf{F} : \mathbb{I}^2 \to \mathbb{I}$ is *bisymmetric* if for all $\mathbf{x} \in \mathbb{I}^4$, we have

$$\mathsf{F}\big(\mathsf{F}(x_1, x_2), \mathsf{F}(x_3, x_4)\big) = \mathsf{F}\big(\mathsf{F}(x_1, x_3), \mathsf{F}(x_2, x_4)\big).$$

The bisymmetry property is very easy to handle and has been investigated from the algebraic point of view by using it mostly in structures without the property of associativity. For a list of references see Aczél [**?**, Section 6.4] (see also Aczél and Dhombres [**?**, Chapter 17], and Soublin [**?**]).

For n arguments, bisymmetry takes the following form (see Aczél [**?**]).

Definition 18. $\mathsf{F} : \mathbb{I}^n \to \mathbb{I}$ is *bisymmetric* if

$$\begin{aligned} &\mathsf{F}\big(\mathsf{F}(x_{11}, \ldots, x_{1n}), \ldots, \mathsf{F}(x_{n1}, \ldots, x_{nn})\big) \\ = \; &\mathsf{F}\big(\mathsf{F}(x_{11}, \ldots, x_{n1}), \ldots, \mathsf{F}(x_{1n}, \ldots, x_{nn})\big) \end{aligned}$$

for all square matrices

$$\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix} \in \mathbb{I}^{n\times n}.$$

Bisymmetry expresses the condition that aggregation of all the elements of any square matrix can be performed first on the rows, then on the columns, or conversely. However, since only square matrices are involved, this property seems not to have a good interpretation in terms of aggregation. Its usefulness remains theoretical. We then consider it for extended functions as follows; see Marichal et al. [**?**].

Definition 19. $\mathsf{F} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is *strongly bisymmetric* if $\mathsf{F}(x) = x$ for all $x \in \mathbb{I}$, and if, for any $n, p \in \mathbb{N}$, we have

$$\begin{aligned} &\mathsf{F}\big(\mathsf{F}(x_{11}, \ldots, x_{1n}), \ldots, \mathsf{F}(x_{p1}, \ldots, x_{pn})\big) \\ = \; &\mathsf{F}\big(\mathsf{F}(x_{11}, \ldots, x_{p1}), \ldots, \mathsf{F}(x_{1n}, \ldots, x_{pn})\big) \end{aligned}$$

for all matrices

$$\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pn} \end{pmatrix} \in \mathbb{I}^{p\times n}.$$

Remark 1. Contrary to bisymmetry, the strong bisymmetry property can be justified rather easily. Consider n judges (or criteria, attributes, etc.) giving a numerical score to each of p candidates. These scores, assumed to be defined on the same scale, can be put in a $p \times n$ matrix like

$$\begin{array}{c c} & \begin{array}{ccc} J_1 & \cdots & J_n \end{array} \\ \begin{array}{c} C_1 \\ \vdots \\ C_p \end{array} & \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pn} \end{pmatrix} \end{array}$$

Suppose now that we want to aggregate all the entries (scores) of the matrix in order to obtain an overall score of the p candidates. A reasonable way to proceed could be the following. First aggregate the scores of each candidate (aggregation over the rows of the matrix), and then aggregate these overall values. An alternative way to proceed would be to first aggregate the scores given by each judge (aggregation over the columns of the matrix), and then aggregate these values. The strong bisymmetry property means that these two ways to aggregate must lead to the same overall score, which is a natural property. Of course, we could as well consider only one candidate, n judges, and p criteria (assuming commensurateness of the scores along the criteria). In this latter setting, strong bisymmetry seems very natural as well.

2.9 Neutral and annihilator elements

The neutral element is again a well-known notion coming from the area of binary operations. Recall that for a binary operation $*$ defined on a domain X, an element $e \in X$ is called a *neutral element* (of the operation $*$) if

$$x * e = e * x = x \qquad (x \in X).$$

Clearly, any binary operation $*$ can have at most one neutral element. From the previous equalities we can see that the action of the neutral element of a binary operation has the same effect as its omission. This idea is the background of the general definition.

Definition 20. Let $\mathsf{F} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \overline{\mathbb{R}}$ be an extended function. An element $e \in \mathbb{I}$ is called an *extended neutral element* of F if, for any $i \in \{1, \ldots, n\}$ and any $\mathbf{x} \in \mathbb{I}^n$ such that $x_i = e$, then

$$\mathsf{F}(x_1, \ldots, x_n) = \mathsf{F}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

So the extended neutral element can be omitted from the input values without influencing the aggregated value. In multicriteria decision making, assigning a score equal to the extended neutral element (if it exists) to some criterion means that only the other criteria fulfillments are decisive for the overall evaluation.

For n-ary functions, there is an alternative approach, given in the following definition:

Definition 21. An element $e \in \mathbb{I}$ is called a *neutral element* of a function $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ if, for any $i \in \{1, \ldots, n\}$ and any $x \in \mathbb{I}$, we have $\mathsf{F}(x_{\{i\}}e) = x$.

Clearly, if $e \in \mathbb{I}$ is an extended neutral element of an extended function $\mathsf{F} : \cup_{n \in \mathbb{N}} \mathbb{I}^n \to \mathbb{I}$, with $\mathsf{F}^{(1)}(x) = x$, then e is a neutral element of all $\mathsf{F}^{(n)}$, $n \in \mathbb{N}$. For instance, $e = 0$ is an extended neutral element for the extended sum function $\mathsf{\Sigma}$.

Definition 22. An element $a \in \mathbb{I}$ is called an *annihilator element* of a function $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ if, for any $\mathbf{x} \in \mathbb{I}^n$ such that $a \in \{x_1, \ldots, x_n\}$, we have $\mathsf{F}(\mathbf{x}) = a$.

Proposition 9. Consider an aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$. If A is conjunctive and $a := \inf \mathbb{I} \in \mathbb{I}$ then a is an annihilator element. Dually, if A is disjunctive and $b := \sup \mathbb{I} \in \mathbb{I}$ then b is an annihilator element.

The converse of Proposition 9 is false. For instance, in $[0, 1]^n$, 0 is an annihilator of the geometric mean GM, which is not conjunctive.

For more specific properties of aggregation functions we recommend to consider [**?**], chapter 2.

3 Means and averages

It would be very unnatural to propose a monograph on aggregation functions without dealing somehow with *means* and *averaging functions*. Already discovered and studied by the ancient Greeks,[2] the concept of mean has given rise today to a very wide field of investigation with a huge variety of applications. Actually, a tremendous amount of literature on the properties of several means (such as the arithmetic mean, the geometric mean, etc.) has already been produced, especially since the 19th century, and is still developing today. For a good overview, see the expository paper by Frosini [**?**] and the remarkable monograph by Bullen [**?**].

The first modern definition of mean was probably due to Cauchy [**?**] who considered in 1821 a *mean* as an internal function. We adopt this approach and assume further that a mean should be a nondecreasing function.

As usual, $\mathbb{I}$ represents a nonempty real interval, bounded or not. The more general cases where $\mathbb{I}$ includes $-\infty$ and/or ∞ will always be mentioned explicitly.

Definition 23. An n-ary *mean* in $\mathbb{I}^n$ is an internal aggregation function $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$. An *extended mean* in $\cup_{n \in \mathbb{N}} \mathbb{I}^n$ is an extended function $\mathsf{M} : \cup_{n \in \mathbb{N}} \mathbb{I}^n \to \mathbb{I}$ whose restriction to each $\mathbb{I}^n$ is a mean.

It follows that a mean is nothing other than an idempotent aggregation function. Moreover, if $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$ is a mean in $\mathbb{I}^n$, then it is also a mean in $\mathbb{J}^n$, for any subinterval $\mathbb{J} \subseteq \mathbb{I}$.

The concept of mean as an *average* or *numerical equalizer* is usually ascribed to Chisini [**?**, p. 108], who gave in 1929 the following definition:

> Let $y = \mathsf{F}(x_1, \ldots, x_n)$ be a function of n independent variables $x_1, \ldots, x_n$. A mean of $x_1, \ldots, x_n$ with respect to the function F

[2]See Antoine [**?**, Chapter 3] for a historical discussion of the various Greek notions of "mean".

is a number M such that, if each of $x_1, \ldots, x_n$ is replaced by M, the function value is unchanged, that is,

$$\mathsf{F}(M, \ldots, M) = \mathsf{F}(x_1, \ldots, x_n).$$

When F is considered as the sum, the product, the sum of squares, the sum of inverses, or the sum of exponentials, the solution of Chisini's equation corresponds respectively to the arithmetic mean, the geometric mean, the quadratic mean, the harmonic mean, and the exponential mean.

Unfortunately, as noted by de Finetti [**?**, p. 378] in 1931, Chisini's definition is so general that it does not even imply that the "mean" (provided there exists a real and unique solution to Chisini's equation) fulfills Cauchy's internality property.

To ensure existence, uniqueness, and internality of the solution of Chisini's equation, we assume that F is nondecreasing and idempotizable. Therefore we propose the following definition:

Definition 24. A function $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$ is an *average* in $\mathbb{I}^n$ if there exists a nondecreasing and idempotizable function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ such that $\mathsf{F} = \delta_{\mathsf{F}} \circ \mathsf{M}$. In this case, we say that M is an *average associated with* F *in* $\mathbb{I}^n$.

Averages are also known as *Chisini means* or *level surface means*. The average associated with F is also called the F-*level mean* (see Bullen [**?**, VI.4.1]). The following result shows that, thus defined, the concepts of mean and average coincide.

Proposition 10. The following assertions hold:

(i) Any average is a mean.

(ii) Any mean is the average associated with itself.

(iii) Let M be the average associated with a function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$. Then M is the average associated with a function $\mathsf{G} : \mathbb{I}^n \to \mathbb{R}$ if and only if there exists an increasing bijection $\varphi : \mathrm{ran}(\mathsf{F}) \to \mathrm{ran}(\mathsf{G})$ such that $\mathsf{G} = \varphi \circ \mathsf{F}$.

Proposition 10 shows that, thus defined, the concepts of mean and average are identical and, in a sense, rather general. Note that some authors (see for instance Bullen [**?**, p. xxvi], Sahoo and Riedel [**?**, Section 7.2], and Bhatia [**?**, Chapter 4]) define the concept of mean by adding conditions such as continuity, symmetry, and homogeneity, which is $\mathsf{M}(r\,\mathbf{x}) = r\,\mathsf{M}(\mathbf{x})$ for all admissible $r \in \mathbb{R}$.

3.1 Quasi-arithmetic means

A well-studied class of means is the class of *quasi-arithmetic means* (see for instance Bullen [**?**, Chapter IV]), introduced as extended aggregation functions as early as 1930 by Kolmogoroff [**?**], Nagumo [**?**], and then as n-ary functions in 1948 by Aczél [**?**]. In this section we introduce the quasi-arithmetic means and describe some of their properties and axiomatizations.

Definition 25. Let $f : \mathbb{I} \to \mathbb{R}$ be a continuous and strictly monotonic function. The n-ary *quasi-arithmetic mean generated by* f is the function $\mathsf{M}_f : \mathbb{I}^n \to \mathbb{I}$ defined as

$$\mathsf{M}_f(\mathbf{x}) := f^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big). \tag{7}$$

The *extended quasi-arithmetic mean generated by* f is the function $\mathsf{M}_f : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ whose restriction to $\mathbb{I}^n$ is the n-ary quasi-arithmetic mean generated by f.

Remark 2. (i) Each quasi-arithmetic mean M_f is a mean in the sense of Definition 23. It is also the average associated with $n(f\circ\mathsf{M}_f)$. For instance, the arithmetic mean $\mathsf{M}_f = \mathsf{AM}$ (with $f = \mathrm{id}$) is the average associated with the sum $\mathbf{\Sigma}$.

(ii) In certain applications it may be convenient to extend the range of f to the extended real line $\overline{\mathbb{R}} = [-\infty, \infty]$. Evidently, in this case it is necessary to define the expression $\infty - \infty$, which will often be considered as $-\infty$.

The class of quasi-arithmetic means comprises most of the algebraic means of common use such as the arithmetic mean and the geometric mean. Table 1 provides some well-known instances of quasi-arithmetic means.

$f(x)$	$\mathsf{M}_f(\mathbf{x})$	name	notation
x	$\frac{1}{n}\sum_{i=1}^{n} x_i$	arithmetic mean	AM
x^2	$\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)^{1/2}$	quadratic mean	QM
$\log x$	$\left(\prod_{i=1}^{n} x_i\right)^{1/n}$	geometric mean	GM
x^{-1}	$\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}}$	harmonic mean	HM
x^α $(\alpha \in \mathbb{R}\setminus\{0\})$	$\left(\frac{1}{n}\sum_{i=1}^{n} x_i^\alpha\right)^{1/\alpha}$	root-mean-power	$\mathsf{M}_{\mathrm{id}^\alpha}$
$e^{\alpha x}$ $(\alpha \in \mathbb{R}\setminus\{0\})$	$\frac{1}{\alpha}\ln\left(\frac{1}{n}\sum_{i=1}^{n} e^{\alpha x_i}\right)$	exponential mean	EM_α

Table 1: Examples of quasi-arithmetic means

The function f occurring in (7) is called a *generator* of M_f. Aczél [**?**] showed that f is determined up to a linear transformation. More generally, we have the following result (see Bullen et al. [**?**, p. 226]):

Proposition 11. Let $f, g : \mathbb{I} \to \mathbb{R}$ be continuous and strictly monotonic functions. Assume also that g is increasing (respectively, decreasing). Then

(i) $\mathsf{M}_f \leqslant \mathsf{M}_g$ if and only if $g \circ f^{-1}$ is convex (respectively, concave);

(ii) $\mathsf{M}_f = \mathsf{M}_g$ if and only if $g \circ f^{-1}$ is linear, that is,

$$g(x) = rf(x) + s \qquad (r, s \in \mathbb{R},\ r \neq 0).$$

We now present an axiomatization of the class of quasi-arithmetic means as extended aggregation functions, originally called *mean values*. The next theorem brings an axiomatization of quasi-arithmetic means as extended aggregation functions. This axiomatization was obtained independently by Kolmogoroff [**?**] and Nagumo [**?**] in 1930.

Theorem 1. $\mathsf{F} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{R}$ is symmetric, continuous, strictly increasing, idempotent, and decomposable if and only if there is a continuous and strictly monotonic function $f : \mathbb{I} \to \mathbb{R}$ such that $\mathsf{F} = \mathsf{M}_f$ is the extended quasi-arithmetic mean generated by f.

Another axiomatization of n-ary quasi-arithmetic means is due to Aczél [**?**].

Theorem 2. The function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ is symmetric, continuous, strictly increasing, idempotent, and bisymmetric if and only if there is a continuous and strictly monotonic function $f : \mathbb{I} \to \mathbb{R}$ such that $\mathsf{F} = \mathsf{M}_f$ is the quasi-arithmetic mean generated by f.

Remark 3. Note that the results given in Theorem 1, and Theorem 2 can be extended to subintervals $\mathbb{I}$ of the extended real line containing ∞ or $-\infty$ with a slight modification of the requirements. Namely, the codomain of F should be $[-\infty, \infty]$, and the strict monotonicity and continuity are required for bounded input vectors only. Observe also that if $\mathbb{I} = [-\infty, \infty]$ then the corresponding quasi-arithmetic means are no more continuous due to the non-continuity of the standard summation on $[-\infty, \infty]$.

Adding some particular property, a special subfamily of quasi-arithmetic means is obtained. Due to Nagumo [**?**] we have the next result.

Theorem 3. (i) The quasi-arithmetic mean $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$ is difference scale invariant, i.e., $\mathsf{M}(\mathbf{x} + \mathbf{c}) = \mathsf{M}(\mathbf{x}) + \mathbf{c}$ for all $\mathbf{x} \in \mathbb{I}^n$ and $\mathbf{c} = (c, \ldots, c) \in \mathbb{R}^n$ such that $\mathbf{x} + \mathbf{c} \in \mathbb{I}^n$, if and only if either M is the arithmetic mean AM or M is the exponential mean, i.e., there exists $\alpha \in \mathbb{R} \setminus \{0\}$ such that

$$\mathsf{M}(\mathbf{x}) = \frac{1}{\alpha} \ln\Big(\frac{1}{n}\sum_{i=1}^{n} e^{\alpha x_i}\Big).$$

(ii) Assume $\mathbb{I} \subseteq \,]0, \infty[$. The quasi-arithmetic mean $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$ is ratio scale invariant, i.e., $\mathsf{M}(c\mathbf{x}) = c\mathsf{M}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{I}^n$ and $c > 0$ such that $c\mathbf{x} \in \mathbb{I}^n$, if and only if either M is the geometric mean GM or M is the root-mean-power, i.e., there exists $\alpha \in \mathbb{R} \setminus \{0\}$ such that

$$\mathsf{M}(\mathbf{x}) = \Big(\frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha}\Big)^{\frac{1}{\alpha}}.$$

A modification of Theorem 2, where the symmetry requirements is omitted, yields an axiomatic characterization of weighted quasi-arithmetic means, see Aczél [**?**] (these are called also quasi-linear means in some sources).

Theorem 4. The function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ is continuous, strictly increasing, idempotent, and bisymmetric if and only if there is a continuous and strictly monotonic

$f(x)$	$\mathsf{M}(\mathbf{x})$	name	notation
x	$\sum_{i=1}^{n} w_i\, x_i$	weighted arithmetic mean	$\mathsf{WAM}_{\mathbf{w}}$
x^2	$\Big(\sum_{i=1}^{n} w_i\, x_i^2\Big)^{1/2}$	weighted quadratic mean	$\mathsf{WQM}_{\mathbf{w}}$
$\log x$	$\prod_{i=1}^{n} x_i^{w_i}$	weighted geometric mean	$\mathsf{WGM}_{\mathbf{w}}$
x^α $(\alpha \in \mathbb{R}\setminus\{0\})$	$\Big(\sum_{i=1}^{n} w_i\, x_i^\alpha\Big)^{1/\alpha}$	weighted root-mean-power	$\mathsf{WM}_{\mathrm{id}^\alpha,\mathbf{w}}$

Table 2: Examples of quasi-linear means

function $f : \mathbb{I} \to \mathbb{R}$ and real numbers $w_1, \ldots, w_n > 0$ satisfying $\sum_i w_i = 1$ such that

$$\mathsf{F}(\mathbf{x}) = f^{-1}\Big(\sum_{i=1}^{n} w_i\, f(x_i)\Big) \qquad (\mathbf{x} \in \mathbb{I}^n). \tag{8}$$

Weighted quasi-arithmetic means can be seen as transformed weighted arithmetic means. The later means are trivially characterized by the additivity property, i.e., $\mathsf{F}(\mathbf{x}+\mathbf{y}) = \mathsf{F}(\mathbf{x}) + \mathsf{F}(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y}, \mathbf{x}+\mathbf{y} \in \mathrm{Dom}(\mathsf{F})$.

Proposition 12. $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}$ is additive, nondecreasing, and idempotent if and only if there exists a weight vector $\mathbf{w} \in [0,1]^n$ satisfying $\sum_i w_i = 1$ such that $\mathsf{F} = \mathsf{WAM}_{\mathbf{w}}$.

Corollary 1. $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}$ is additive, nondecreasing, symmetric, and idempotent if and only if $\mathsf{F} = \mathsf{AM}$ is the arithmetic mean.

Remark 4. As will be discussed in Section 4, the weighted arithmetic means $\mathsf{WAM}_{\mathbf{w}}$ are exactly the Choquet integrals with respect to additive normalized capacities; see Proposition 14 (v). If we further assume the symmetry property, we obtain the arithmetic mean AM.

A natural way to generalize the quasi-arithmetic mean consists in incorporating weights as in the quasi-linear mean (8). To generalize a step further, we could assume that the weights are not constant. On this issue, Losonczi [**?**, **?**] considered and investigated in 1971 nonsymmetric functions of the form

$$\mathsf{M}(\mathbf{x}) = f^{-1}\left(\frac{\sum_{i=1}^{n} p_i(x_i) f(x_i)}{\sum_{i=1}^{n} p_i(x_i)}\right),$$

where $f : \mathbb{I} \to \mathbb{R}$ is a continuous and strictly monotonic function and $p_1, \ldots, p_n : \mathbb{I} \to \,]0, \infty[$ are positive valued functions. The special case where $p_1 = \cdots = p_n$

was previously introduced in 1958 by Bajraktarević [?] who defined the concept of *quasi-arithmetic mean with weight function* as follows (see also Páles [?]).[3]

Definition 26. Let $f : \mathbb{I} \to \mathbb{R}$ be a continuous and strictly monotonic function and let $p : \mathbb{I} \to]0, \infty[$ be a positive valued function. The n-ary *quasi-arithmetic mean generated by f with weight function p* is the function $\mathsf{M}_{f,p} : \mathbb{I}^n \to \mathbb{I}$ defined as

$$\mathsf{M}_{f,p}(\mathbf{x}) = f^{-1}\left(\frac{\sum_{i=1}^n p(x_i)f(x_i)}{\sum_{i=1}^n p(x_i)}\right).$$

The *extended quasi-arithmetic mean generated by f with weight function p* is the function $\mathsf{M}_{f,p} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ whose restriction to $\mathbb{I}^n$ is the n-ary quasi-arithmetic mean generated by f with weight function p.

It is very important to note that, even though quasi-arithmetic means with weight function are (clearly) idempotent, they need not be nondecreasing, which implies that they need not be means or even aggregation functions. To give an example, consider the case where $n = 2$, $f(x) = x$, and $p(x) = 2x + 1$, that is,

$$\mathsf{M}_{f,p}(x_1, x_2) = \frac{2x_1^2 + 2x_2^2 + x_1 + x_2}{2x_1 + 2x_2 + 2}.$$

We can readily see that the section $x \mapsto \mathsf{M}_{f,p}(x, 1)$ of this binary function is not nondecreasing.

Assuming that the weight function p is nondecreasing and differentiable, Marques Pereira and Ribeiro [?] and Mesiar and Špirková [?] found sufficient conditions on p to ensure nondecreasing monotonicity of $\mathsf{M}_{\mathrm{id},p}$. Here we assume $\mathbb{I} = [0, 1]$ as in [?] but the conditions easily extend to arbitrary intervals. For extended arithmetic means $\mathsf{M}_{\mathrm{id},p}$ with weight function p, the simplest sufficient condition to ensure nondecreasing monotonicity is $p(x) \geqslant p'(x) \geqslant 0$. A more general one is

$$p(x) \geqslant (1 - x)p'(x) \geqslant 0 \qquad (x \in [0, 1]).$$

Remark 5. For extended quasi-arithmetic means $\mathsf{M}_{f,p}$ with weight function p, assuming that both f and p are increasing and differentiable and that $\mathrm{ran}(f) = [0, 1]$, the above sufficient condition generalizes into

$$f'(x)\, p(x) \geqslant (1 - f(x))\, p'(x) \geqslant 0 \qquad (x \in [0, 1]).$$

For n-ary arithmetic means $\mathsf{M}_{\mathrm{id},p}$ with weight function p, we also have the sufficient condition

$$\frac{p^2(x)}{(n-1)\, p(1)} + p(x) \geqslant (1 - x)p'(x) \qquad (x \in [0, 1]).$$

Remark 6. It is worth mentioning that $\mathsf{M}_{f,p}$ can also be obtained by the minimization problem

$$\mathsf{M}_{f,p}(\mathbf{x}) = \arg\min_{r\in\mathbb{R}} \sum_{i=1}^n p(x_i)\, (f(x_i) - f(r))^2.$$

Evidently, in the same way, classical quasi-arithmetic means (p is constant) and weighted quasi-arithmetic means (replace $p(x_i)$ with w_i) are obtained. For more details, see Calvo et al. [?] and Mesiar and Špirková [?].

[3] The subcase where $f = \mathrm{id}$, called *Beckenbach-Gini means* or *mixture operators*, has been investigated by Marques Pereira and Ribeiro [?], Matkowski [?], and Yager [?].

3.2 Constructions of means

There are several methods how to construct means (binary, n-ary, extended). Integral-based methods are discussed in Section 4.

Means can also be constructed by minimization of functions. This construction method will be thoroughly discussed in Part II, Section 4.

To give here a simple example, consider weights $w_1, w_2 \in]0, \infty[$ and minimize (in r) the expression

$$f(r) = w_1|x_1 - r| + w_2(x_2 - r)^2.$$

This minimization problem leads to the unique solution

$$r = \mathsf{M}(x_1, x_2) = \mathsf{Med}\big(x_1, x_2 - \frac{w_1}{2w_2}, x_2 + \frac{w_1}{2w_2}\big),$$

which defines a mean $\mathsf{M} : \mathbb{R}^2 \to \mathbb{R}$.

Any nonsymmetric function F can be symmetrized by replacing it variables $x_1, \ldots, x_n$ with corresponding order statistics functions $x_{(1)}$ (minimal input), $x_{(2)}, \ldots, x_{(n)}$ (maximal input).

One of the simplest examples is given by the *ordered weighted averaging function*

$$\mathsf{OWA}_{\mathbf{w}}(\mathbf{x}) = \sum_{i=1}^{n} w_i\, x_{(i)}, \tag{9}$$

which merely results from the symmetrization of the corresponding weighted arithmetic mean $\mathsf{WAM}_{\mathbf{w}}$.

Remark 7. The concept of ordered weighted averaging function was introduced by Yager[4] in 1988; see Yager [**?**], and also the book [**?**] edited by Yager and Kacprzyk. Since then, the family of these functions has been axiomatized in various ways; see for instance Fodor et al. [**?**] and Marichal and Mathonet [**?**]. Also, these functions are exactly the Choquet integrals with respect to symmetric normalized capacities; see Proposition 14 (vi).

The symmetrization process can naturally be applied to the quasi-linear mean (i.e., to the weighted quasi-arithmetic mean) (8) to produce the *quasi-ordered weighted averaging function* $\mathsf{OWA}_{\mathbf{w},f} : \mathbb{I}^n \to \mathbb{R}$, which is defined as

$$\mathsf{OWA}_{\mathbf{w},f}(\mathbf{x}) := f^{-1}\Big(\sum_{i=1}^{n} w_i f(x_{(i)})\Big),$$

where the generator $f : \mathbb{I} \to \mathbb{R}$ is a continuous and strictly monotonic function; see Fodor et al. [**?**].

The classical mean value formulas (Lagrange, Cauchy) lead to the concept of lagrange (Cauchy) means, see [**?**] and [**?**, **?**].

Definition 27. Let $f : \mathbb{I} \to \mathbb{R}$ be a continuous and strictly monotonic function. The *Lagrangian mean* $\mathsf{M}_{[f]} : \mathbb{I}^2 \to \mathbb{I}$ associated with f is a mean defined as

$$\mathsf{M}_{[f]}(x, y) := \begin{cases} f^{-1}\left(\dfrac{1}{y - x}\displaystyle\int_x^y f(t)\,\mathrm{d}t\right) & \text{if } x \neq y \\ x & \text{if } x = y. \end{cases} \tag{10}$$

[4]Note however that linear (not necessarily convex) combinations of ordered statistics were already studied previously in statistics; see for instance Weisberg [**?**] (and David and Nagaraja [**?**, Section 6.5] for a more recent overview).

The uniqueness of the generator is the same as for quasi-arithmetic means: Let f and g be two generators of the same Lagrangian mean. Then, there exist $r, s \in \mathbb{R}$, $r \neq 0$ such that $g(x) = rf(x) + s$; see [?, Corollary 7], [?, p. 344], and [?, Theorem 1].

Many classical means are Lagrangian. The arithmetic mean, the geometric mean, and the so-called *Stolarsky means* [?], defined by

$$\mathsf{M}_S(x,y) := \begin{cases} \left(\dfrac{x^r - y^r}{r(x-y)}\right)^{\frac{1}{r-1}} & \text{if } x \neq y \\ x & \text{if } x = y, \end{cases}$$

correspond to taking $f(x) = x$, $f(x) = 1/x^2$, and $f(x) = x^{r-1}$, respectively, in (10). The harmonic mean, however, is not Lagrangian.

In general, some of the most common means are both quasi-arithmetic and Lagrangian, but there are quasi-arithmetic means, like the harmonic one, which are not Lagrangian. Conversely, the *logarithmic mean*

$$\mathsf{M}(x,y) := \begin{cases} \dfrac{x-y}{\log x - \log y} & \text{for } x, y > 0,\ x \neq y \\ x & \text{for } x = y > 0, \end{cases}$$

is an example of a Lagrangian mean (actually a Stolarsky mean, $f(x) = 1/x$), that is not quasi-arithmetic. A characterization of the class of Lagrangian means and a study of its connections with the class of quasi-arithmetic means can be found in Berrone and Moro [?]. Further properties of Lagrangian means and other extensions are investigated for instance in Aczél and Kuczma [?], Berrone [?], Głazowska [?], Horwitz [?, ?], Kuczma [?], Sándor [?], and Wimp [?].

Definition 28. Let $f, g : \mathbb{I} \to \mathbb{R}$ be continuous and strictly monotonic functions. The *Cauchy mean* $\mathsf{M}_{[f,g]} : \mathbb{I}^2 \to \mathbb{I}$ associated with the pair (f, g) is a mean defined as

$$\mathsf{M}_{[f,g]}(x,y) := \begin{cases} f^{-1}\left(\dfrac{1}{g(y)-g(x)}\displaystyle\int_x^y f(t)\,\mathrm{d}g(t)\right) & \text{if } x \neq y \\ x & \text{if } x = y. \end{cases}$$

We note that any Cauchy mean is continuous, idempotent, symmetric, and strictly increasing.

When $g = f$ (respectively, g is the identity function), we retrieve the quasi-arithmetic (respectively, the Lagrangian) mean generated by f. The *anti-Lagrangian mean* [?] is obtained when f is the identity function. For example, the harmonic mean is an anti-Lagrangian mean generated by the function $g = 1/x^2$. We also note that the generator of an anti-Lagrangian mean is defined up to a non-zero affine transformation.

Further studies on Cauchy means can be found for instance in Berrone [?], Lorenzen [?], and Losonczi [?, ?]. Extensions of Lagrangian and Cauchy means, called *generalized weighted mean values*, including discussions on their monotonicity properties, can be found in Chen and Qi [?, ?], Qi et al. [?, ?, ?], and Witkowski [?].

3.3 Associative means

The class of continuous, nondecreasing, idempotent, and associative binary functions is described in the next theorem. The result is due to Fodor [?] who obtained this description in a more general framework, where the domain of variables is any connected order topological space. Alternative proofs were obtained independently in Marichal [?, Section 3.4] and [?, Section 5].

Theorem 5. $\mathsf{M} : \mathbb{I}^2 \to \mathbb{I}$ is continuous, nondecreasing, idempotent, and associative if and only if there exist $\alpha, \beta \in \mathbb{I}$ such that

$$\mathsf{M}(x,y) = (\alpha \wedge x) \vee (\beta \wedge y) \vee (x \wedge y). \tag{11}$$

Notice that, by distributivity of $\wedge$ and $\vee$, M can be written also in the equivalent form:

$$\mathsf{M}(x,y) = (\beta \vee x) \wedge (\alpha \vee y) \wedge (x \vee y).$$

The graphical representation of M is given in Figure 1.

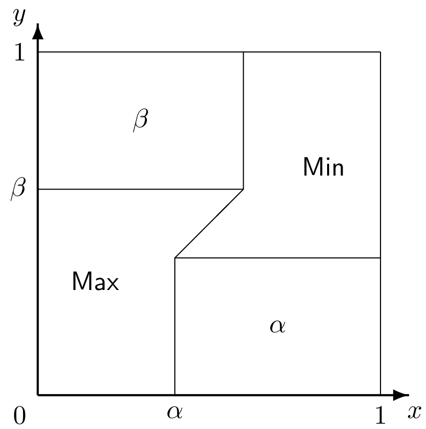

Figure 1: Representation on $[0,1]^2$ of function (11) when $\alpha \leqslant \beta$

Theorem 5 can be generalized straightforwardly to extended means as follows.

Theorem 6. $\mathsf{M} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is continuous, nondecreasing, idempotent, and associative if and only if there exist $\alpha, \beta \in \mathbb{I}$ such that, for any $n \in \mathbb{N}$,

$$\mathsf{M}^{(n)}(\mathbf{x}) = (\alpha \wedge x_1) \vee \Big(\bigvee_{i=2}^{n-1} (\alpha \wedge \beta \wedge x_i)\Big) \vee (\beta \wedge x_n) \vee \Big(\bigwedge_{i=1}^{n} x_i\Big). \tag{12}$$

Remark 8. Means described in Theorem 5 are nothing other than idempotent binary lattice polynomial functions, that is, binary Sugeno integrals; see Section 4. We also observe that the n-ary lattice polynomial function given in (12) is an n-ary Sugeno integral defined from a particular normalized capacity.

The special case of symmetric associative means was already discussed by Fung and Fu [?] and revisited in Dubois and Prade [?]. It turns out that these functions are the α-medians, i.e., for $\alpha \in \mathbb{I}$ α-median , $\mathsf{Med}_\alpha : \mathbb{I}^n \to \mathbb{I}$ is given by

$$\mathsf{Med}_\alpha(\mathbf{x}) = \mathsf{Med}(x_1, \dots, x_n, \underbrace{\alpha, \dots, \alpha}_{n-1}) = \mathsf{Med}(\mathsf{Min}(\mathbf{x}), \alpha, \mathsf{Max}(\mathbf{x})).$$

The description is the following.

Theorem 7. $\mathsf{M} : \mathbb{I}^2 \to \mathbb{I}$ is symmetric, continuous, nondecreasing, idempotent, and associative if and only if there exist $\alpha \in \mathbb{I}$ such that $\mathsf{M} = \mathsf{Med}_\alpha$. Similarly, $\mathsf{M} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is symmetric, continuous, nondecreasing, idempotent, and associative if and only if there exist $\alpha \in \mathbb{I}$ such that, for any $n \in \mathbb{N}$, $\mathsf{M}^{(n)} = \mathsf{Med}_\alpha^{(n)}$.

Remark 9. Since the conjunction of symmetry and associativity implies bisymmetry, we immediately see that the α-medians Med_α are particular nonstrict arithmetic means.

Czogała and Drewniak [**?**] have examined the case when M has a neutral element $e \in \mathbb{I}$. They obtained the following result.

Theorem 8. If $\mathsf{M} : \mathbb{I}^2 \to \mathbb{I}$ is nondecreasing, idempotent, associative, and has a neutral element $e \in \mathbb{I}$, then there is a nonincreasing function $g : \mathbb{I} \to \mathbb{I}$ with $g(e) = e$ such that, for all $x, y \in \mathbb{I}$,

$$\mathsf{M}(x,y) = \begin{cases} x \wedge y & \text{if } y < g(x) \\ x \vee y & \text{if } y > g(x) \\ x \wedge y \text{ or } x \vee y & \text{if } y = g(x). \end{cases}$$

Furthermore, if M is continuous, then $\mathsf{M} = \mathsf{Min}$ or $\mathsf{M} = \mathsf{Max}$.

Remark 10. (i) Fodor [**?**] showed that Theorem 8 still holds in the more general framework of connected order topological spaces.

(ii) The restriction of Theorem 8 to symmetric functions corresponds to idempotent uninorms.

4 Aggregation functions based on nonadditive integrals

The preceding section has developed the notion of means, which can be viewed as a variation of the idea of finite sum. Another generalization is the notion of integral, where the sum becomes infinite. Beside the classical Riemann integral, many types of integral exist, but there is one which is of particular interest to us, namely the Lebesgue integral, which is defined with respect to a measure. Indeed, the classical notion of measure extends the notion of weight to infinite universes, and the Lebesgue integral on a finite universe coincides with the weighted arithmetic mean. Therefore, the existence of more general notions of measure than the classical additive one, together with the appropriate integrals, offer a new realm of aggregation functions when these integrals are limited to a finite universe. These general measures may be called *nonadditive measures*, and the corresponding integrals *nonadditive integrals*, however a more precise term would be *monotonic measures* since additivity is replaced by monotonicity, although the most common name —which we will use—, is *capacity*, as coined by Choquet [?]. The term *fuzzy measure* introduced by Sugeno [?] is often used in the fuzzy set community, although this term is mathematically misleading since no fuzziness is involved there.

There are basically two types of integrals defined with respect to a capacity, namely the Choquet integral and the Sugeno integral, leading to two interesting classes of aggregation functions, developed in this section.

We start by introducing the notion of capacity.

Definition 29. Let $X = \{1, \dots, n\}$ be a given universe. A set function $\mu : 2^X \to [0, \infty[$ is called a *capacity* if $\mu(\varnothing) = 0$ and $\mu(A) \leqslant \mu(B)$ whenever $A \subset B \subset X$ (monotonicity). If in addition $\mu(X) = 1$, the capacity is *normalized*.

If μ does not satisfy monotonicity, it is called a *game*. If μ takes only values 0 and 1, then μ is called a *0-1 capacity*. For any capacity μ, the *dual* capacity is defined by $\mu^d(A) := \mu(X) - \mu(X \setminus A)$, for any $A \subseteq X$.

Let μ be a capacity on X and $A, B \subseteq X$. We say that μ is *additive* if $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever A and B are disjoint; it is *symmetric* if $\mu(A) = \mu(B)$ whenever $|A| = |B|$; it is *maxitive* if $\mu(A \cup B) = \mu(A) \vee \mu(B)$, and it is *minitive* if $\mu(A \cap B) = \mu(A) \wedge \mu(B)$.

We give several fundamental examples of capacities.

(i) The smallest normalized capacity is $\mu_{\min}(A) := 0$, $\forall A \subsetneq X$, while the greatest one is $\mu_{\max}(A) := 1$, $\forall A \subseteq X, A \neq \varnothing$;

(ii) For any $i \in X$, the *Dirac measure centered on i* is defined by, for any $A \subseteq X$

$$\delta_i(A) := \begin{cases} 1 & \text{if } i \in A \\ 0 & \text{otherwise.} \end{cases}$$

(iii) For any integer k, $1 \leqslant k \leqslant n$, the *threshold measure* τ_k is defined by

$$\tau_k(A) := \begin{cases} 1 & \text{if } |A| \geqslant k \\ 0 & \text{otherwise.} \end{cases}$$

The class of all capacities on X can be partitionned into the so-called k-additive capacities, $k = 1, \dots, n$. We need some additional definitions to introduce this new concept (details can be found in [?, ?, ?]). For a given capacity μ on X, their *Möbius transform* $m^\mu : X \to \mathbb{R}$ and *interaction transform* $I^\mu : X \to \mathbb{R}$ are defined by

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

$$I^\mu(A) := \sum_{B \supseteq A} \frac{1}{b - a + 1} m^\mu(B)$$

for all $A \subseteq X$, with $a = |A|, b = |B|$. A remarkable particular case is $A = \{i\}$, $i \in X$, for the interaction transform, since we recover the well-known Shapley value, whose more familiar expression is

$$I^\mu(\{i\}) =: \phi_i(\mu) = \sum_{A \subseteq X \setminus i} \frac{(n - a - 1)! a!}{n!} (\mu(A \cup i) - \mu(A)).$$

A capacity μ is *k-additive* for some $2 \leq k \leq n$ if $m^\mu(A) = 0$ for all A such that $|A| > k$, and there exists some $A \subseteq X$, $|A| = k$, such that $m^\mu(A) \neq 0$. Note that due to the definition of the interaction transform, the above definition can be equivalently written with I^μ instead of m^μ.

4.1 Choquet integral based aggregation functions

The Choquet integral was introduced in 1953 by Choquet [**?**]. It is a generalization of the Lebesgue integral where the (classical) measure is replaced by a capacity.

Definition 30. Let μ be a capacity on $X = \{1, \dots, n\}$ and $\mathbf{x} \in [0, \infty[^n$. The *Choquet integral* of $\mathbf{x}$ with respect to μ is defined by

$$\mathcal{C}_\mu(\mathbf{x}) := \sum_{i=1}^{n} (x_{\sigma(i)} - x_{\sigma(i-1)})\mu(A_{\sigma(i)})$$

with σ a permutation on $\{1, \dots, n\}$ such that $x_{\sigma(1)} \leqslant x_{\sigma(2)} \leqslant \cdots \leqslant x_{\sigma(n)}$, with the convention $x_{\sigma(0)} := 0$, and $A_{\sigma(i)} := \{\sigma(i), \dots, \sigma(n)\}$.

We will often use the more compact notation $x_{(i)}$ instead of $x_{\sigma(i)}$, which was already introduced in Section 3.

It is straightforward to see that an equivalent formula is

$$\mathcal{C}_\mu(\mathbf{x}) = \sum_{i=1}^{n} x_{(i)}\big(\mu(A_{(i)}) - \mu(A_{(i+1)})\big),$$

with $A_{(n+1)} := \varnothing$. It can be shown that $\mathcal{C}_\mu$ is a continuous aggregation function, for any capacity μ. However, if $\mathcal{C}_\mu$ is defined on $\mathbb{I}$ which is a bounded interval $[0, a]$, it is necessary that the capacity is normalized, otherwise the bounds of interval will not be preserved.

If $\mathbf{x} \in \mathbb{R}^n$, there exist two ways of defining $\mathcal{C}_\mu(\mathbf{x})$, according to how the symmetry is done. The usual one is the following:

$$\mathcal{C}_\mu(\mathbf{x}) = \mathcal{C}_\mu(\mathbf{x}^+) - \mathcal{C}_{\mu^d}(\mathbf{x}^-)$$

for all $\mathbf{x} \in \mathbb{R}^n$, where $x_i^+ := \mathbf{x}_i \vee 0$ for all $i \in X$, and $\mathbf{x}^- := (-\mathbf{x})^+$. For the sake of concision, we do not detail further this topic here and refer the reader to [**?**]. The Choquet integral with $\mathbb{I} = \mathbb{R}$ has a very simple form in terms of the Möbius transform:

$$\mathcal{C}_\mu(\mathbf{x}) = \sum_{A \subseteq X} \Big(m^\mu(A) \cdot \bigwedge_{i \in A} x_i\Big).$$

A last remark before studying properties of the Choquet integral is that this aggregation function when $\mathbb{I} = [0, 1]$ can be obtained as the unique linear parcimonious interpolation on the vertices of the hypercube $[0, 1]^n$. This remarkable property comes from the fact that for any capacity μ, we have $\mathcal{C}_\mu(\mathbf{1}_A) = \mu(A)$ for any subset $A \subseteq X$, where $\mathbf{1}_A$ is the characteristic vector of A. Indeed, vertices of $[0, 1]^n$ correspond to the vectors $\mathbf{1}_A$, $A \subseteq X$, and for a given $\mathbf{x} \in [0, 1]^n$ which is not a vertex, the interpolation is done with the vertices of the canonical simplex

$$[0, 1]^n_\sigma := \{\mathbf{x} \in [0, 1]^n \mid x_{\sigma(1)} \leqslant \cdots \leqslant x_{\sigma(n)}\}$$

with σ any permutation such that $\mathbf{x} \in [0, 1]^n_\sigma$.

Definition 31. Let $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$. We say that $\mathbf{x}, \mathbf{x}'$ are *comonotonic* if there exists a permutation σ on $\{1, \ldots, n\}$ such that $x_{\sigma(1)} \leqslant x_{\sigma(2)} \leqslant \cdots \leqslant x_{\sigma(n)}$ and $x'_{\sigma(1)} \leqslant x'_{\sigma(2)} \leqslant \cdots \leqslant x'_{\sigma(n)}$ (equivalently if there is no $i, j \in \{1, \ldots, n\}$ such that $x_i > x_j$ and $x'_i < x'_j$).

Proposition 13. The Choquet integral satisfies the following properties:

(i) The Choquet integral is linear with respect to the capacity: for any capacities μ_1, μ_2 on X, any $\lambda_1, \lambda_2 \geqslant 0$,

$$\mathcal{C}_{\lambda_1\mu_1+\lambda_2\mu_2} = \lambda_1\mathcal{C}_{\mu_1} + \lambda_2\mathcal{C}_{\mu_2}.$$

If one restricts to normalized capacities, then the condition $\lambda_1 + \lambda_2 = 1$ is needed.

(ii) The Choquet integral satisfies *comonotonic additivity*, i.e., for any comonotonic vectors $\mathbf{x}, \mathbf{x}' \in [0, \infty[^n$, and any capacity μ,

$$\mathcal{C}_\mu(\mathbf{x} + \mathbf{x}') = \mathcal{C}_\mu(\mathbf{x}) + \mathcal{C}_\mu(\mathbf{x}').$$

(iii) Let μ, μ' be two capacities on X. Then $\mu \leqslant \mu'$ if and only if $\mathcal{C}_\mu \leqslant \mathcal{C}_{\mu'}$.

(iv) If μ is a 0-1 capacity, then

$$\mathcal{C}_\mu(\mathbf{x}) = \bigvee_{\substack{A \subseteq \{1,\ldots,n\} \\ \mu(A)=1}} \bigwedge_{i \in A} x_i, \quad \forall \mathbf{x} \in [0,1]^n,$$

.

(v) The Choquet integral $\mathcal{C}_\mu$ is symmetric if and only if μ is symmetric.

(vi) The Choquet integral on $\mathbb{R}^n$ is invariant to positive affine transformation (interval scale change), that is

$$\mathcal{C}_\mu(c\mathbf{x} + a\mathbf{1}_X) = c \cdot \mathcal{C}_\mu(\mathbf{x}) + a$$

for any $c > 0$ and $a \in \mathbb{R}$.

(vii) For any capacity μ on X, we have $(\mathcal{C}_\mu)^d = \mathcal{C}_{\mu^d}$, i.e., the dual of the Choquet integral is the Choquet integral. w.r.t. its dual capacity.

The next theorem gives a characterization of the Choquet integral.

Theorem 9. Let $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}$ be a given function. Then there exists a unique normalized capacity μ such that $\mathsf{F} = \mathcal{C}_\mu$ if and only if F fulfills the following properties:

(i) Comonotonic additivity;

(ii) Nondecreasing monotonicity;

(iii) $\mathsf{F}(\mathbf{1}_{\{1,\ldots,n\}}) = 1$, $\mathsf{F}(\mathbf{0}) = 0$.

Moreover, μ is defined by $\mu(A) := \mathsf{F}(\mathbf{1}_A)$.

This result was shown by De Campos and Bolaños [**?**] in the case of $\mathbb{I} = [0, \infty[$, assuming in addition positive homogeneity, which can be deduced from (i) and (ii). The proof in the continuous case is due to Schmeidler [**?**].

We give the relation of the Choquet integral with other aggregation functions.

Proposition 14. Let μ be a normalized capacity and consider $\mathbb{I} = \mathbb{R}$. The following holds

(i) $\mathcal{C}_\mu = \mathsf{Min}$ if and only if $\mu = \mu_{\min}$.

(ii) $\mathcal{C}_\mu = \mathsf{Max}$ if and only if $\mu = \mu_{\max}$.

(iii) $\mathcal{C}_\mu = \mathsf{OS}_k$ (k-order statistics) if and only if μ is the threshold measure τ_{n-k+1}.

(iv) $\mathcal{C}_\mu = \mathsf{P}_k$ (k-th projection) if and only if μ is the Dirac measure δ_k.

(v) $\mathcal{C}_\mu = \mathsf{WAM}_{\mathbf{w}} = \check{\mathcal{C}}_\mu$ if and only if μ is additive, with $\mu(\{i\}) = w_i$, $\forall i \in \{1, \dots, n\}$.

(vi) $\mathcal{C}_\mu = \mathsf{OWA}_{\mathbf{w}}$ if and only if μ is symmetric, with $w_i = \mu(A_{n-i+1}) - \mu(A_{n-i})$, $i = 2, \dots, n$, and $w_1 = 1 - \sum_{i=2}^n w_i$, where A_i is any subset of X with $|A_i| = i$ (equivalently, $\mu(A) = \sum_{j=0}^{i-1} w_{n-j}$, $\forall A, |A| = i$).

Let us come back to k-additive capacities. The Choquet integral has an interesting expression in terms of the interaction transform when the capacity is 2-additive:

$$\mathcal{C}_\mu(\mathbf{x}) = \sum_{\substack{i,j\in X\\ I_{ij}>0}} (x_i \wedge x_j) I_{ij} + \sum_{\substack{i,j\in X\\ I_{ij}<0}} (x_i \vee x_j)|I_{ij}| + \sum_{i\in X}\Big(\phi_i x_i - \frac{1}{2}\sum_{j\neq i}|I_{ij}|\Big),$$

with $I_{ij} := I^\mu(\{i,j,\})$ and $\phi_i := I^\mu(\{i\})$. It is important to note that the above expression is a convex sum of at most $\frac{n(n+1)}{2}$ terms, grouped in three parts: a conjunctive one, a disjunctive one and an additive one.

We end this section by considering multilevel Choquet integrals. An interesting question is the following: what new aggregation function family do we obtain by combining the output of several different Choquet integrals by another Choquet integral, possibly iterating this process on several levels? For example, we may consider a 3-level aggregation of $\mathbf{x} \in \mathbb{R}^3$:

$$\mathsf{A}(\mathbf{x}) := \mathcal{C}_{\mu_3}(\mathcal{C}_{\mu_{11}}(x_1,x_2), \mathcal{C}_{\mu_2}(\mathcal{C}_{\mu_{11}}(x_1,x_2), \mathcal{C}_{\mu_{12}}(x_1,x_2,x_3)), x_3).$$

This question has been solved by Murofushi and Narukawa [**?**], and it is based on a result by Ovchinnikov on piecewise linear functions [**?**]. The answer is simply that one gets nothing new after the second level.

First we give a formal definition of a multilevel Choquet integral.

Definition 32. Let $\Gamma \subseteq \mathbb{R}^n$. For any $i \in [n]$, the projection P_i is a *0-level Choquet integral*. Let us consider $\mathsf{F}_i : \Gamma \to \mathbb{R}$, $i \in M := \{1, \dots, m\}$, being k_i-level Choquet integrals, and a capacity μ on M. Then

$$\mathsf{F}(\mathbf{x}) := \mathcal{C}_\mu(\mathsf{F}_1(\mathbf{x}), \dots, \mathsf{F}_m(\mathbf{x}))$$

is a k-*level Choquet integral*, with $k := \mathsf{Max}(k_1, \ldots, k_m) + 1$. A *multilevel Choquet integral* is a function that is a k-level Choquet integral for some integer $k > 1$.

The result is the following.

Theorem 10. Let $\Gamma \subseteq \mathbb{R}^n$ be a convex closed n-dimensional set, and $\mathsf{F} : \Gamma \to \mathbb{R}$. The following are equivalent.

(i) F is a multilevel Choquet integral.

(ii) F is a 2-level Choquet integral, with all capacities of the first level being additive, and the capacity of the second level being 0-1 valued.

(iii) F is nondecreasing, positively homogeneous, and continuous piecewise linear.

4.2 Sugeno integral based aggregation functions

The Sugeno integral was introduced by Sugeno in 1972 [**?**], independently of the work of Choquet. Yet, there are striking similarities between the two definitions, up to the fact that the usual arithmetic operations of the Choquet integral are replaced by the lattice operations $\vee, \wedge$. However, the introduction of these two lattices operations implies many fundamental differences in their properties, and makes the Sugeno integral very close to the lattice polynomial functions.

We will see that the definition of the Sugeno integral makes sense only if $\mathbb{I} = [0, \mu(X)]$.

Definition 33. Let μ be a capacity on $X = \{1, \ldots, n\}$, and $\mathbf{x} \in [0, \mu(X)]^n$. The *Sugeno integral* of $\mathbf{x}$ with respect to μ is defined by

$$\mathcal{S}_\mu(\mathbf{x}) := \bigvee_{i=1}^{n} \big(x_{\sigma(i)} \wedge \mu(A_{\sigma(i)})\big)$$

with the same notations as in Definition 30. It can be shown that the Sugeno integral is a continuous aggregation function. The condition $\mathbb{I} = [0, \mu(X)]$ is necessary to fulfill the boundary conditions. As for the Choquet integral, we have $\mathcal{S}_\mu(\mathbf{1}_A) = \mu(A)$ for all $A \subseteq X$.

It can be shown that two other equivalent expressions are:

$$\mathcal{S}_\mu(\mathbf{x}) = \bigwedge_{i=1}^{n} \big(x_{\sigma(i)} \vee \mu(A_{\sigma(i+1)})\big) \tag{13}$$

$$= \mathsf{Med}(x_1, \ldots, x_n, \mu(A_{\sigma(2)}), \ldots, \mu(A_{\sigma(n)})), \tag{14}$$

with $A_{\sigma(n+1)} := \varnothing$, and Med is the classical median function.

The Sugeno integral has a close relation with weighted lattice polynomial functions. They are inductively defined as follows: (i) for any $k \in X$ and any $c \in \mathbb{I}$, the projection P_k and the constant function c are weighted lattice polynomial functions; (ii) if p, q are weighted lattice polynomial functions, then $p \vee q$ and $p \wedge q$ are weighted lattice polynomial functions; every weighted lattice polynomial function is formed by finitely many applications of rules (i) and (ii).

Each weighted lattice polynomial function $p : \mathbb{I}^n \to \mathbb{I}$ can be written both in conjunctive and disjunctive normal forms [**?**]:

$$p(\mathbf{x}) = \bigvee_{A \subseteq X} \Big(\alpha(A) \wedge \bigwedge_{i \in A} x_i\Big) = \bigwedge_{A \subseteq X} \Big(\beta(A) \vee \bigvee_{i \in A} x_i\Big), \tag{15}$$

where $\alpha, \beta : 2^X \to \mathbb{I}$ are some set functions. Based on (15), we have the next representation of the Sugeno integral.

Proposition 15. For any $\mathbf{x} \in [0, \mu(X)]^n$ and any capacity μ on X, the Sugeno integral of $\mathbf{x}$ with respect to μ can be written as

$$\mathcal{S}_\mu(\mathbf{x}) = \bigvee_{A \subseteq X} \left(\bigwedge_{i \in A} x_i \wedge \mu(A)\right)$$

$$\mathcal{S}_\mu(\mathbf{x}) = \bigwedge_{A \subseteq X} \left(\bigvee_{i \in A} x_i \vee \mu(X \setminus A)\right).$$

The following result gives the exact relation between weighted lattice polynomial functions and the Sugeno integral.

Theorem 11. Let $\mathsf{F} : [0,1]^n \to [0,1]$ be a function. The following assertions are equivalent.

(i) There exists a unique normalized capacity μ such that $\mathsf{F} = \mathcal{S}_\mu$;

(ii) F is an idempotent weighted lattice polynomial function;

(iii) F is an endpoint-preserving weighted lattice polynomial function.

A consequence of this theorem is that the multilevel Sugeno integral is the Sugeno integral: one does not get anything new by combining Sugeno integrals.

The Sugeno integral possesses several particular properties given in the next propositions.

Proposition 16. The Sugeno integral satisfies the following properties:

(i) The Sugeno integral commutes with max-min combinations of capacities: for any nonnegative games μ_1, μ_2 on X, any $\lambda_1, \lambda_2 \in [0, \infty[$,

$$\mathcal{S}_{(\lambda_1 \wedge \mu_1) \vee (\lambda_2 \wedge \mu_2)} = (\lambda_1 \wedge \mathcal{S}_{\mu_1}) \vee (\lambda_2 \wedge \mathcal{S}_{\mu_2})$$

$$\mathcal{S}_{(\lambda_1 \vee \mu_1) \wedge (\lambda_2 \vee \mu_2)} = (\lambda_1 \vee \mathcal{S}_{\mu_1}) \wedge (\lambda_2 \vee \mathcal{S}_{\mu_2}),$$

with the convention $((\lambda_1 \vee \mu_1) \wedge (\lambda_2 \vee \mu_2))(\varnothing) := 0$.

(ii) The Sugeno integral satisfies *comonotonic maxitivity* and *comonotonic minitivity*: for any comonotonic vectors $\mathbf{x}, \mathbf{x}' \in [0, \mu(X)]^n$, and any capacity μ,

$$\mathcal{S}_\mu(\mathbf{x} \vee \mathbf{x}') = \mathcal{S}_\mu(\mathbf{x}) \vee \mathcal{S}_\mu(\mathbf{x}')$$

$$\mathcal{S}_\mu(\mathbf{x} \wedge \mathbf{x}') = \mathcal{S}_\mu(\mathbf{x}) \wedge \mathcal{S}_\mu(\mathbf{x}').$$

(iii) Let μ, μ' be two nonnegative games on X. Then $\mu \leqslant \mu'$ if and only if $\mathcal{S}_\mu \leqslant \mathcal{S}_{\mu'}$.

(iv) $\mathcal{C}_\mu = \mathcal{S}_\mu$ if and only if μ is a $0-1$ capacity.

(v) For any normalized capacity μ and any $\mathbf{x} \in [0,1]^n$, $|\mathcal{C}_\mu(\mathbf{x}) - \mathcal{S}_\mu(\mathbf{x})| \leqslant \frac{1}{4}$.

(vi) For any capacity μ on X, we have $(\mathcal{S}_\mu)^d = \mathcal{S}_{\mu^d}$.

Proposition 17. The Sugeno integral $\mathcal{S}_\mu$ satisfies the following properties:

(i) Symmetry (or neutrality, commutativity) if and only if μ is symmetric;

(ii) Additivity if and only if μ is a 0-1 additive capacity (Dirac measure);

(iii) Maxitivity if and only if μ is a maxitive capacity (possibility measure);

(iv) Minitivity if and only if μ is a minitive capacity (necessity measure).

The next characterization is due to Marichal [**?**]. Still others can be found in this reference.

Theorem 12. Let $\mathsf{F} : [0,1]^n \to [0,1]$. Then there exists a capacity μ on X such that $\mathsf{F} = \mathcal{S}_\mu$ if and only if F satisfies (i) nondecreasingness, (ii) $\mathcal{S}_\mu(\alpha \vee \mathbf{x}) = \alpha \vee \mathcal{S}_\mu(\mathbf{x})$ ($\vee$-homogeneity), and (iii) $\mathcal{S}_\mu(\alpha \wedge \mathbf{x}) = \alpha \wedge \mathcal{S}_\mu(\mathbf{x})$ ($\wedge$-homogeneity).

Now we show the relation of the Sugeno integral with other aggregation functions.

We begin by introducing several aggregation functions.

Definition 34. Let $\mathbf{w} \in [0,1]^n$ satisfying $\bigvee_{i=1}^n w_i = 1$. Then, for any $\mathbf{x} \in [0,1]^n$:

(i) The *weighted maximum* with respect to $\mathbf{w}$ is the aggregation function defined by

$$\mathsf{WMax}_{\mathbf{w}}(\mathbf{x}) := \bigvee_{i=1}^{n} (w_i \wedge x_i), \quad \forall \mathbf{x} \in \mathbb{I}^n.$$

(ii) The *weighted minimum* with respect to $\mathbf{w}$ is the aggregation function defined by

$$\mathsf{WMin}_{\mathbf{w}}(\mathbf{x}) := \bigwedge_{i=1}^{n} ((1-w_i) \vee x_i), \quad \forall \mathbf{x} \in \mathbb{I}^n.$$

(iii) The *ordered weighted maximum* with respect to $\mathbf{w}$ is the aggregation function defined by

$$\mathsf{OWMax}_{\mathbf{w}}(\mathbf{x}) := \bigvee_{i=1}^{n} (w_i \wedge x_{(i)}), \quad \forall \mathbf{x} \in \mathbb{I}^n,$$

with $x_{(1)} \leqslant \cdots \leqslant x_{(n)}$, and $w_1 \geqslant w_2 \geqslant \cdots \geqslant w_n$.

(iv) The *ordered weighted minimum* with respect to $\mathbf{w}$ is the aggregation function defined by

$$\mathsf{OWMin}_{\mathbf{w}}(\mathbf{x}) := \bigwedge_{i=1}^{n} ((1-w_i) \vee x_{(i)}), \quad \forall \mathbf{x} \in \mathbb{I}^n.$$

with $x_{(1)} \leqslant \cdots \leqslant x_{(n)}$, and $w_1 \leqslant w_2 \leqslant \cdots \leqslant w_n$.

Proposition 18. Let μ be a capacity. The following holds.

(i) $\mathcal{S}_\mu = \mathsf{Min}$ if and only if $\mu = \mu_{\min}$.

(ii) $\mathcal{S}_\mu = \mathsf{Max}$ if and only if $\mu = \mu_{\max}$.

(iii) $\mathcal{S}_\mu = \mathsf{OS}_k$ if and only if μ is the threshold capacity τ_{n-k+1}.

(iv) $\mathcal{S}_\mu = \mathsf{P}_k$ if and only if μ is the Dirac measure δ_k.

(v) $\mathcal{S}_\mu = \mathsf{WMax}_\mathbf{w}$ if and only if μ is a normalized maxitive capacity, with $\mu(\{i\}) = w_i$, for all $i \in \{1, \ldots, n\}$.

(vi) $\mathcal{S}_\mu = \mathsf{WMin}_\mathbf{w}$ if and only if μ is a normalized minitive capacity, with $\mu(X \setminus \{i\}) = \mu(X) - w_i$, for all $i \in X$.

(vii) $\mathcal{S}_\mu = \mathsf{OWMax}_\mathbf{w}$ if and only if μ is a normalized symmetric capacity such that $\mu(A) = w_{n-|A|+1}$, for any $A \subseteq X$, $A \neq \varnothing$.

(viii) $\mathcal{S}_\mu = \mathsf{OWMin}_\mathbf{w}$ if and only if μ is a normalized symmetric capacity such that $\mu(A) = 1 - w_{n-|A|}$, for any $A \subsetneq X$.

(ix) The set of Sugeno integrals with respect to 0-1 capacities coincides with the set of lattice polynomial functions.

5 Concluding remarks

In this Part I we have focused on internal aggregation functions. We have discussed some of their properties, some construction methods (nonadditive integrals, for example) and some representation theorems. Much more details can be found in monographs of Bullen [**?**], and in our monograph [**?**]. Some of internal, i.e., idempotent aggregation functions will be discussed in Part II due to construction methods discussed there. As a typical example recall the weighted median, which for integer weights can be defined straightforwardly as

$$\mathsf{Med}_\mathbf{w}(x_1, \ldots, x_n) = \mathsf{Med}(w_1 \cdot x_1, \ldots, w_n \cdot x_n),$$

while for real (nonnegative) weights it is linked to the minimization problem of expression $\sum_{i=1}^n w_1 |x_i - r|$.

In the nonadditive integral domain we have discussed only the Choquet and the Sugeno integrals, though there are several other types of nonadditive integrals which may be interested (see, e.g., section 5.6 in [**?**]). Without going deeper into details, we recall briefly two of them, based on t-conorms, uninorms and copulas (these particular aggregation functions on $[0, 1]$ are discussed in Part II). The (S, U)-integral based on continuous t-conorm S and a uninorm U satisfying the restricted distributivity relation

$$\mathsf{U}(x, \mathsf{S}(y, z)) = \mathsf{S}(\mathsf{U}(x, y), \mathsf{U}(x, z))$$

for all $x, y, z \in [0, 1]$ such that $\mathsf{S}(y, z) < 1$ is given, for any $\mathbf{x} \in [0, 1]^n$ and S-additive fuzzy measure μ on X, $\mu(A \cup B) = \mathsf{S}(\mu(A), \mu(B))$ whenever A and B are disjoint, by

$$(\mathsf{S}, \mathsf{U})_\mu(\mathbf{x}) = \mathsf{S}_{j=1}^k \left(\mathsf{S}_{i=1}^m (\mathsf{U}(x_i, \mu(\{i\})))\right).$$

As an example take $\mathsf{S} = \mathsf{S}_L$ the bounded sum and $\mathsf{U} = \prod$ the product. Then $(\mathsf{S}, \mathsf{U})_\mu$ coincide with the Weber integral [**?**]. Any S_L-additive fuzzy measure on X is determined by the values of μ on singletons, $w_i = \mu(\{i\})$, and $\mu(A) = \min\left(\sum_{i\in A} w_i, 1\right)$. Note that necessarily $\sum_{i=1}^n w_i \geqslant 1$, and

$$(\mathsf{S}, \mathsf{U})_\mu(\mathbf{x}) = \min\left(\sum_{i=1}^n w_i x_i, 1\right).$$

Obviously, if $\sum_{i=1}^n w_i = 1$, the weighted arithmetic mean $\mathsf{WAM}_\mathbf{w}$ is recovered.

Copulas C are linked to the probability measures P_C on $([0,1]^2, \mathcal{B}([0,1]^2)$ with uniform marginals, $\mathsf{C}(x, y) = P_\mathsf{C}([0, x] \times [0, y])$. A copula based integral (see Imaoka [**?**] for special copulas and Klement at al. [**?**] for general copulas) is given for any $\mathbf{x} \in [0,1]^n$ and any fuzzy measure μ on X, by two equivalent formulas

$$\begin{aligned} I_\mathsf{C}(\mathbf{x}, \mu) &= \sum_{i=1}^n \left(\mathsf{C}(x_{\sigma(i)}, \mu(A_{\sigma(i)})) - \mathsf{C}(x_{\sigma(i-1)}, \mu(A_{\sigma(i)}))\right) \\ &= \sum_{i=1}^n \left(\mathsf{C}(x_{\sigma(i)}, \mu(A_{\sigma(i)})) - \mathsf{C}(x_{\sigma(i)}, \mu(A_{\sigma(i+1)}))\right), \end{aligned}$$

with the convention $A_{\sigma(n+1)} = \varnothing$, and $x_{\sigma(0)} = 0$. Observe that $I_{\prod}(\cdot, \mu) = \mathcal{C}_\mu$ is the Choquet integral, while $I_{\mathsf{Min}}(\cdot, \mu) = \mathcal{S}_\mu$ is the Sugeno integral.

In Part II, we will discussed conjunctive and disjunctive aggregation functions, and also some mixed aggregation functions related to both conjunctive and disjunctive aggregation functions. Moreover, several construction methods for aggregation functions will be introduced.

Acknowledgment

The authors gratefully acknowledge the support of their respective institutions during the long period of writing of the manuscript. During this period, we benefitted from the support of various projects which facilitated communication and collaboration among us, in particular the bilateral project between France and Serbia *Pavle Savić* N° 11092SF supported by EGIDE, Paris, and bilateral project between Slovakia and Serbia SK-SRB-19, the internal research project *Decision Mathematics and Operations Research* supported by the University of Luxembourg, project APVV-0012-07 supported by the Slovak grant agency, the national grants MNTRS (Serbia, Project 144012), Provincial Secretariat for Science and Technological Development of Vojvodina.