Fonctions d'agrégation pour la décision*

Jean-Luc Marichal

Service de Mathématiques Appliquées, Université du Luxembourg

162A, avenue de la Faïencerie, L-1511 Luxembourg

Email: jean-luc.marichal[at]uni.lu

Version révisée, 30 juin 2005

1 Introduction

Les fonctions d'agrégation sont généralement définies et utilisées pour combiner et résumer plusieurs valeurs numériques en une seule, de telle sorte que le résultat final de l'agrégation prenne en compte, d'une manière prescrite, toutes les valeurs individuelles. De telles fonctions sont largement utilisées dans de nombreuses disciplines bien connues comme la statistique, l'économie, la finance, l'informatique, etc.

Par exemple, supposons que plusieurs personnes forment des jugements quantifiables sur la mesure d'un objet (poids, longueur, surface, hauteur, importance ou autres attributs) ou même sur le ratio de deux telles mesures (combien plus lourd, plus long, plus grand, plus important un objet est-il par rapport à un autre). Pour atteindre un consensus sur ces jugements, des fonctions d'agrégation classiques ont été proposées : la moyenne arithmétique, la moyenne géométrique, la médiane et bien d'autres encore.

En aide à la décision multicritère, les valeurs à agréger sont généralement des *préférences* (d'une alternative par rapport à une autre) ou des *degrés de satisfaction* (d'une alternative) relatifs à des critères. En aide à la décision face à l'incertain, les valeurs à agréger représentent les *conséquences* d'une action relatives à des états de la nature.

Nous supposerons que les valeurs à agréger appartiennent à des échelles numériques, qui peuvent être de type ordinal ou cardinal. Sur une échelle ordinale, les nombres n'ont d'autres significations que de définir une relation d'ordre sur l'échelle, et les distances ou différences entre les valeurs ne peuvent pas être interprétées. Sur une échelle cardinale, les distances entre les valeurs ne sont pas arbitraires. En fait, il y a plusieurs sortes d'échelles cardinales : sur une échelle d'intervalle, où la position du zéro est purement conventionnelle, les valeurs sont définies à une transformation

*Chapitre 4 de *Concepts et Méthodes pour l'Aide à la Décision* (Hermès Science Publications).

linéaire positive près, c'est-à-dire $\phi(x) = rx + s$, avec $r > 0$ et $s \in \mathbb{R}$ (par exemple des températures exprimées sur l'échelle Celsius) ; sur une échelle de ratio, où un zéro réel existe, les valeurs sont définies à une similarité près, c'est-à-dire $\phi(x) = rx$, avec $r > 0$ (par exemple des longueurs exprimées en pouces). Nous reviendrons sur ces aspects du mesurage à la sous-section 2.2.

Une fois que les valeurs à agréger sont définies, nous pouvons les fusionner en une seule valeur au moyen d'une fonction d'agrégation. Mais une telle opération peut s'effectuer de nombreuses façons selon ce qui est attendu de la fonction d'agrégation, selon la nature des valeurs à agréger, et selon le type des échelles qui sont utilisées. Ainsi, pour un problème donné, le choix d'une fonction d'agrégation doit être fait avec soin et l'utilisation de telle ou telle fonction doit toujours être justifiée.

Pour aider le praticien à choisir une fonction d'agrégation appropriée au problème qu'il traite, il est utile et même convenable d'adopter une approche axiomatique. Cette approche consiste à classer et choisir les fonctions d'agrégation selon des propriétés qu'elles vérifient. Ainsi, un catalogue de propriétés "souhaitables" est établi et, lorsque c'est possible, une description de la famille des fonctions d'agrégation satisfaisant un ensemble donné de propriétés est fourni. C'est le principe même de l'axiomatisation.

Proposer une caractérisation axiomatique intéressante n'est généralement pas une tâche facile. La plupart du temps, une même famille de fonctions d'agrégation peut être caractérisée par différent ensembles de propriétés. Néanmoins, toutes les caractérisations possibles ne sont pas également importantes. Certaines impliquent des conditions purement techniques sans interprétation claire et le résultat devient inintéressant. D'autres impliquent des conditions qui contiennent explicitement le résultat et la caractérisation devient triviale. A l'opposé, il y a des caractérisations ne faisant intervenir que des propriétés naturelles, facilement interprétables. En fait, c'est le seul cas où le résultat peut être considéré comme une contribution importante. En effet, il améliore notre compréhension des fonctions d'agrégation concernées et fournit des arguments forts pour justifier ou rejeter leur utilisation dans un contexte donné.

Le but principal de ce chapitre est de présenter, sur une base axiomatique, les familles de fonctions d'agrégation les plus utilisées en aide à la décision. Nous nous limiterons cependant aux fonctions d'agrégation qui associent une valeur numérique à chaque profil de n valeurs, lesquels représentent des objets ou des alternatives. Nous ne traiterons pas des fonctions d'utilité qui, de façon plus générale, permettent de ranger les alternatives sans leur assigner des valeurs précises. Ainsi par exemple, les procédures de type 'leximin' ou 'discrimin' sont des procédures de rangement, plutôt que des fonctions d'agrégation à proprement parler.

L'organisation de ce chapitre est la suivante : Dans la section 2, nous donnons la liste des principales propriétés que nous utiliserons. Cette liste est divisée en trois grandes parties : (1) les propriétés élémentaires (continuité, symétrie, etc.), (2) les propriétés liées aux types d'échelles utilisées pour représenter les données, et (3) certaines propriétés algébriques comme l'associativité. Dans la section 3, nous présentons le concept de moyenne et ses différentes définitions. La définition qui est peut-être la plus commune est celle des moyennes quasi-arithmétiques avec une axiomatique très naturelle due à Kolmogoroff et Nagumo. A la section 4, nous présentons les fonctions associatives, qui sont à l'origine de la théorie des semi-groupes. Ces fonctions ont permis de développer le concept de connecteurs flous tels que les t-normes, les t-conormes et les uninormes. A la section 5, nous présentons une branche importante de la théorie des fonctions d'agrégation, à savoir les intégrales non-additives de Choquet et de Sugeno. Ces intégrales permettent de généraliser les modes d'agrégation classiques, comme la moyenne arithmétique pondérée et la médiane, à des fonctions prenant en compte les interactions possibles parmi les attributs considérés. Enfin, aux sections 6 et 7, nous présentons des fonctions particulières conçues pour l'agrégation en présence d'échelles d'intervalle, d'échelles de ratio et d'échelles ordinales.

Nous terminons cette introduction en précisant quelques notations qui seront souvent utilisées dans ce chapitre.

D'une façon générale, nous noterons une fonction d'agrégation à n variables par $A : E^n \to \mathbb{R}$ où E est un intervalle réel, borné ou non. E° désignera l'intérieur de E. Nous considérerons parfois des suites de fonctions $(A^{(n)} : E^n \to \mathbb{R})_{n\geq 1}$, l'exposant $^{(n)}$ ne servant qu'à préciser le nombre d'arguments de la fonction $A^{(n)}$.

Nous utiliserons également N pour désigner l'ensemble des indices $\{1, \ldots, n\}$ et 2^N pour désigner l'ensemble de ses parties. Π_N sera aussi utilisé pour désigner l'ensemble des permutations sur N. Enfin, pour tout $S \subseteq N$, le vecteur caractéristique de S dans $\{0, 1\}^n$ sera noté $\mathbf{1}_S$.

Il existe également des notations relativement standards pour certaines fonctions d'agrégation. Voici les plus courantes :

- La *moyenne arithmétique* est définie par

$$\mathrm{AM}(x) = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

- Pour tout vecteur de poids $\omega = (\omega_1, \ldots, \omega_n) \in [0, 1]^n$ tel que $\sum_i \omega_i = 1$, la *moyenne arithmétique pondérée* et la *fonction moyenne ordonnée* sont définies respectivement par

$$\begin{aligned} \mathrm{WAM}_\omega(x) &= \sum_{i=1}^{n} \omega_i x_i, \\ \mathrm{OWA}_\omega(x) &= \sum_{i=1}^{n} \omega_i x_{(i)}, \end{aligned}$$

  où $(\cdot)$ représente une permutation sur N telle que $x_{(1)} \leq \cdots \leq x_{(n)}$.

- Pour tout $k \in N$, la *projection* et la *statistique d'ordre* associées au $k$ème argument sont respectivement définies par

$$\begin{aligned} \mathrm{P}_k(x) &= x_k, \\ \mathrm{OS}_k(x) &= x_{(k)}. \end{aligned}$$

- Pour tout $S \subseteq N$, $S \neq \varnothing$, les fonctions *minimum partiel* et *maximum partiel* associés à S sont respectivement définis par

$$\min_S(x) = \min_{i\in S} x_i,$$

$$\max_S(x) \quad = \quad \max_{i\in S} x_i.$$

Dans ce chapitre, les opérations min et max seront parfois notées $\wedge$ et $\vee$, respectivement.

2 Propriétés pour l'agrégation

Comme nous venons de le dire dans l'introduction, pour choisir un mode d'agrégation raisonnable et satisfaisant, il est utile d'adopter une approche axiomatique et sélectionner ainsi les fonctions d'agrégation qui vérifient certaines propriétés. De telles propriétés peuvent être dictées par la nature des valeurs à agréger. Par exemple, dans un problème classique d'analyse multicritère, un des objectifs est d'évaluer le score global d'une alternative à partir de scores partiels obtenus sur différents critères. Dans ce cas, il ne serait pas très naturel de donner au score global une valeur inférieure au plus petit des scores partiels ou supérieure au plus grand des scores partiels. Ainsi, seule une fonction de type "interne" (une moyenne) peut être utilisée. Pour donner un autre exemple, supposons que l'on souhaite agréger des opinions dans une procédure de vote. Si les votants sont anonymes, la fonction d'agrégation doit être symétrique.

Dans cette section, nous présentons quelques propriétés qui peuvent être vues comme souhaitables ou non en fonction du problème considéré. Bien sûr, toutes ces propriétés ne sont pas requises avec la même intensité et peuvent avoir des objectifs très différents. Certaines représentent des conditions impératives dont la violation conduirait à des modes d'agrégation contre-intuitifs. D'autres sont plus techniques et ont pour seul but de faciliter la représentation ou le calcul des fonctions d'agrégation. Enfin, il y a aussi des propriétés plutôt facultatives qui ne s'appliquent que dans des circonstances particulières et qui ne sont pas universellement acceptées.

2.1 Propriétés mathématiques élémentaires

Définition 2.1 $A : E^n \to \mathbb{R}$ *est symétrique si, pour tout* $\pi \in \Pi_N$*, on a*

$$A(x_1, \ldots, x_n) = A(x_{\pi(1)}, \ldots, x_{\pi(n)}) \qquad (x \in E^n).$$

La propriété de symétrie signifie que l'ordre des x_i est sans importance pour l'agrégation. Ceci est requis notamment lorsque l'on combine des critères d'importances égales ou des opinions d'experts anonymes.

Définition 2.2 $A : E^n \to \mathbb{R}$ *est continu s'il est continu au sens habituel.*

L'avantage d'une fonction continue est qu'elle ne présente aucun saut brusque suite à de faibles variations des valeurs partielles.

Définition 2.3 $A : E^n \to \mathbb{R}$ *est*

– *non décroissant si, pour tous* $x, x' \in E^n$, *on a*

$$x \leq x' \quad \Rightarrow \quad A(x) \leq A(x'),$$

– *strictement croissant s'il est non décroissant et si , pour tous* $x, x' \in E^n$, *on a*

$$x \leq x' \text{ et } x \neq x' \quad \Rightarrow \quad A(x) < A(x'),$$

– *unanimement croissant s'il est non décroissant et si , pour tous* $x, x' \in E^n$, *on a*

$$x < x' \quad \Rightarrow \quad A(x) < A(x').$$

Une fonction non décroissante présente un comportement non négatif à tout accroissement des arguments. En d'autres termes, l'accroissement d'une valeur partielle ne fait pas décroître le résultat. La fonction est strictement croissante si, en plus, elle réagit positivement à tout accroissement d'au moins une valeur partielle. Enfin, la fonction est unanimement croissante si elle est non décroissante et présente une réaction positive chaque fois que tous les arguments croissent. Par exemple, nous observons que, sur $[0, 1]^n$, la fonction maximum $A(x) = \max x_i$ est unanimement croissante, alors que la somme bornée $A(x) = \min(\sum_{i=1}^{n} x_i, 1)$ ne l'est pas.

Définition 2.4 $A : E^n \to \mathbb{R}$ *est idempotent si* $A(x, \ldots, x) = x$ *pour tout* $x \in E$.

Définition 2.5 $A : [a, b]^n \to \mathbb{R}$ *est faiblement idempotent si* $A(a, \ldots, a) = a$ *et* $A(b, \ldots, b) = b$.

Dans de nombreuses applications, il est requis que la fonction d'agrégation vérifie la propriété d'idempotence : si tous les x_i sont identiques, $M(x_1, \ldots, x_n)$ restitue la valeur commune.

Définition 2.6 $A : E^n \to \mathbb{R}$ *est*
– *conjonctif si* $A(x) \leq \min x_i$ *pour tout* $x \in E^n$,
– *disjonctif si* $\max x_i \leq A(x)$ *pour tout* $x \in E^n$,
– *interne si* $\min x_i \leq A(x) \leq \max x_i$ *pour tout* $x \in E^n$.

Les fonctions conjonctives combinent les valeurs comme si elles étaient reliées par un opérateur logique "ET". En d'autres termes, le résultat de l'agrégation n'est élevé que si toutes les valeurs partielles sont élevées. Les t-normes sont des fonctions qui se comportent de cette manière (voir la sous-section 4.5). A l'opposé, les fonctions disjonctives combinent les valeurs comme un opérateur logique "OU", de telle sorte que le résultat de la combinaison est élevé si au moins l'une des valeurs partielles est élevée. Les fonctions disjonctives les plus connues sont les t-conormes.

Entre ces deux situations extrêmes se trouvent les fonctions internes, situées entre le min et le max. Dans ce type de fonctions, une valeur partielle faible peut être compensée par une autre plus élevée. Par definition, les fonctions de type "moyennes" sont des fonctions internes (voir Section 3).

2.2 Propriétés de stabilité liées aux types d'échelles

Selon le type d'échelle qui est utilisé, les opération autorisées sur les valeurs sont plus ou moins limitées. Par exemple, une agrégation sur des échelles ordinales doit nécessairement se restreindre aux opérations n'utilisant rien d'autre que des comparaisons, telles que les statistiques d'ordre.

Une *échelle de mesurage* est une application qui associe un nombre réel à chaque objet mesuré. Le *type* d'une échelle, ainsi défini par Stevens [100, 101], est la donnée d'une classe de *transformations admissibles*, transformations permettant de passer d'une échelle acceptable à une autre. Par exemple, une échelle sera appelée *échelle de ratio* si la classe des transformations admissibles consiste en les similarités $\phi(x) = rx$, avec $r > 0$. Dans ce cas, les valeurs sont déterminées au choix de l'unité près. La masse est un exemple d'échelle de ratio. La conversion de kilogrammes en livres est donnée par la transformation admissible $\phi(x) = 2.2x$. La longueur (centimètres, pouces) et les intervalles de temps (années, secondes) sont deux autres exemples d'échelles de ratio. Une échelle et dite *échelle d'intervalle* si la classe des transformations admissibles consiste en les transformations linéaires positives $\phi(x) = rx + s$, avec $r > 0$ et $s \in \mathbb{R}$. Les valeurs sont alors déterminées au choix de l'unité près mais aussi de la position du zéro. La température (sauf lorsqu'il y a un zéro absolu) définit une échelle d'intervalle. Ainsi par exemple, la transformation permettant de passer des degrés Celsius aux degrés Fahrenheit est donnée par $\phi(x) = 9x/5 + 32$. Une échelle est dite *échelle ordinale* si la classe des transformations admissibles consiste en les bijections ϕ strictement croissantes. Dans ce cas, les valeurs sont déterminées à l'ordre près. Par exemple, l'échelle de la qualité de l'air utilisée dans plusieurs grandes villes est une échelle ordinale. Elle associe la valeur 1 à de l'air irrespirable, 2 à de l'air insatisfaisant, 3 à de l'air acceptable, 4 à de l'air de bonne qualité, 5 à de l'air excellent. Pour définir une telle échelle, on aurait pu utiliser les nombres 1.2, 6.5, 8.7, 205.6, 750, ou n'importe quelle série de nombres qui préserve l'ordre défini. D'autres définitions de types d'échelle peuvent être trouvées dans le livre de Roberts [91] sur la théorie du mesurage ; voir également Roberts [92, 93]. Le lecteur trouvera aussi plus détails sur le mesurage dans le chapitre **??** de ce volume.

Une proposition impliquant des échelles de mesurage est dite *signifiante* si le fait qu'elle soit vraie ou fausse est invariant lorsque les échelles sont remplacées par des versions acceptables [91, p. 59]. Par exemple, une méthode de rangement est signifiante si le rangement des alternatives induit par l'agrégation sous-jacente ne dépend pas des transformations admissibles d'échelles.

En 1959, Luce [61] observa que la forme générale d'une relation fonctionnelle entre des variables est relativement restreinte lorsqu'on connait le type d'échelle utilisé pour les variables. Ces restrictions peuvent être déterminées par la formulation d'une équation fonctionnelle basée sur les transformations admissibles. La méthode de Luce est basée sur le principe qu'une transformation admissible des variables indépendantes peut conduire à une transformation admissible de la variable dépendante. Par exemple, supposons que $f(a) = A(f_1(a), \ldots, f_n(a))$, où f et $f_1, \ldots, f_n$ sont toutes des échelles de ratio dont les unités sont indépendantes les unes des autres. Dans ce cas, nous obtenons l'équation fonctionnelle

$$A(r_1x_1, \ldots, r_nx_r) = R(r_1, \ldots, r_n)A(x_1, \ldots, x_n),$$
$$r_i > 0, \quad R(r_1, \ldots, r_n) > 0.$$

Aczél et col. [9] ont alors montré que les solutions de cette équation sont données par

$$A(x) = a \prod_{i=1}^{n} g_i(x_i), \quad \text{avec } a > 0,\ g_i > 0,$$

et où les fonctions g_i vérifient

$$g_i(x_iy_i) = g_i(x_i)g_i(y_i) \quad \text{pour tous } x_i, y_i \in \mathbb{R}.$$

Dans cette sous-section, nous présentons quelques équations fonctionelles relatives à certains types d'échelle. Le lecteur intéressé pourra trouver de plus amples détails dans [8, 9] et un état de l'art dans [93].

Définition 2.7 *$A : \mathbb{R}^n \to \mathbb{R}$ est*

- *signifiant pour les mêmes échelles de ratio entrées-sorties si, pour tout $r > 0$, on a*

$$A(rx_1, \ldots, rx_n) = rA(x_1, \ldots, x_n) \qquad (x \in \mathbb{R}^n),$$

- *signifiant pour les mêmes échelles de ratio entrées si, pour tout $r > 0$, il existe $R_r > 0$ tel que*

$$A(rx_1, \ldots, rx_n) = R_rA(x_1, \ldots, x_n) \qquad (x \in \mathbb{R}^n),$$

- *signifiant pour les mêmes échelles d'intervalle entrées-sorties si, pour tous $r > 0$ et $s \in \mathbb{R}$, on a*

$$A(rx_1 + s, \ldots, rx_n + s) = rA(x_1, \ldots, x_n) + s \qquad (x \in \mathbb{R}^n),$$

- *signifiant pour les mêmes échelles d'intervalle entrées si, pour tous $r > 0$ et $s \in \mathbb{R}$, il existe $R_{r,s} > 0$ et $S_{r,s} \in \mathbb{R}$ tels que*

$$A(rx_1 + s, \ldots, rx_n + s) = R_{r,s}A(x_1, \ldots, x_n) + S_{r,s} \qquad (x \in \mathbb{R}^n),$$

- *signifiant pour les mêmes échelles ordinales entrées-sorties si, pour toute bijection strictement croissante $\phi : \mathbb{R} \to \mathbb{R}$, on a*

$$A(\phi(x_1), \ldots, \phi(x_n)) = \phi(A(x_1, \ldots, x_n)) \qquad (x \in \mathbb{R}^n),$$

- *signifiant pour les mêmes échelles ordinales entrées si, pour toute bijection strictement croissante $\phi : \mathbb{R} \to \mathbb{R}$, il existe une fonction strictement croissante $\psi_\phi : \mathbb{R} \to \mathbb{R}$ telle que*

$$A(\phi(x_1), \ldots, \phi(x_n)) = \psi_\phi(A(x_1, \ldots, x_n)) \qquad (x \in \mathbb{R}^n).$$

2.3 Propriétés algébriques

Les propriétés qui suivent se rapportent aux procédures d'agrégation qui peuvent se "décomposer" en agrégations partielles c'est-à-dire pour lesquelles il est possible de partitionner l'ensemble des attributs en sous-groupes disjoints, de construire une agrégation partielle pour chaque sous-groupe et ensuite de combiner ces résultats partiels pour obtenir une valeur globale. Une telle décomposition peut prendre plusieurs formes. Peut-être une des plus "restrictives" de ces décompositions est l'associativité, bien connue des algébristes. Nous présentons également deux autres formulations plus faibles : décomposabilité et bisymétrie.

Présentons tout d'abord l'associativité pour les fonctions à deux arguments.

Définition 2.8 *$A : E^2 \to E$ est associatif si, pour tout $x \in E^3$, on a*

$$A(A(x_1, x_2), x_3) = A(x_1, A(x_2, x_3)).$$

Une vaste littérature est consacrée à l'équation fonctionnelle d'associativité. Pour une liste de références, voir [4, §6.2].

Cette propriété s'étend aux suites de fonctions comme suit :

Définition 2.9 *La suite $(A^{(n)} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ est associative si $A^{(1)}(x) = x$ pour tout $x \in E$ et*

$$A^{(n)}(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = A^{(n)}(A^{(k)}(x_1, \ldots, x_k), A^{(n-k)}(x_{k+1}, \ldots, x_n))$$

pour tout $x \in E^n$ et tous $k, n \in \mathbb{N}$ tels que $1 \leq k \leq n$.

Ce qui est implicite dans la définition d'une suite associative, c'est la manière de passer très facilement d'une agrégation de n valeurs à une agrégation de $n+1$ valeurs. En effet, de la définition, on déduit la formule

$$A^{(n+1)}(x_1, \ldots, x_{n+1}) = A^{(2)}(A^{(n)}(x_1, \ldots, x_n), x_{n+1}),$$

pour tout $n \in \mathbb{N} \setminus \{0\}$.

Passons à présent à la propriété de décomposabilité. Dans ce but, nous introduisons la notation suivante : pour tout $k \in \mathbb{N} \setminus \{0\}$ et tout $x \in \mathbb{R}$, nous posons $k \cdot x = x, \ldots, x$ (k fois). Par exemple,

$$A(3 \cdot x, 2 \cdot y) = A(x, x, x, y, y).$$

Définition 2.10 *La suite $(A^{(n)} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ est décomposable si $A^{(1)}(x) = x$ pour tout $x \in E$ et*

$$A^{(n)}(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = A^{(n)}(k \cdot A^{(k)}(x_1, \ldots, x_k), (n-k) \cdot A^{(n-k)}(x_{k+1}, \ldots, x_n))$$

pour tout $x \in E^n$ et tous $k, n \in \mathbb{N}$ tels que $1 \leq k \leq n$.

La définition ici est la même que celle de l'associativité, excepté que les agrégations partielles sont dupliquées un nombre de fois égal au nombre de valeurs agrégées.

Cette propriété de décomposabilité a été introduite sous le nom d'*associativité des moyennes* par Bemporad [14, p. 87] dans une caractérisation de la moyenne arithmétique. Elle a aussi été utilisée par Kolmogoroff [59] et Nagumo [83] pour caractériser les moyennes quasi-arithmétiques. Plus récemment, Marichal et Roubens [71] ont proposé d'appeler cette propriété "décomposabilité" pour ne pas la confondre avec l'associativité classique.

La propriété de *bisymétrie*, qui résulte simultanément de l'associativité et la symétrie, est définie pour les fonctions à n variables comme suit :

Définition 2.11 *$A : E^n \to E$ est bisymétrique si*

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

pour toute matrice carrée $(x_{ij}) \in E^{n\times n}$.

Pour des fonctions à deux variables, cette propriété à été étudiée d'un point de vue algébrique en l'utilisant principalement dans des structures privées de la propriété d'associativité. Pour une liste de références, voir [4, §6.4] et [6, Chapitre 17].

Pour une suite de fonctions, cette propriété devient :

Définition 2.12 *La suite $(A^{(n)} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ est bisymétrique si $A^{(1)}(x) = x$ pour tout $x \in E$ et*

$$\begin{aligned} & A^{(p)}(A^{(n)}(x_{11},\ldots,x_{1n}),\ldots,A^{(n)}(x_{n1},\ldots,x_{pn})) \\ = \; & A^{(n)}(A^{(p)}(x_{11},\ldots,x_{p1}),\ldots,A^{(p)}(x_{1n},\ldots,x_{pn})) \end{aligned}$$

pour tous $n, p \in \mathbb{N} \setminus \{0\}$ et toute matrice $(x_{ij}) \in E^{p\times n}$.

3 Moyennes

Il ne serait pas convenable de proposer un chapitre sur les fonctions d'agrégation sans traiter des fonctions de type moyenne. Déjà bien connu et étudié par les Grecs de l'Antiquité (voir par exemple [12, Chapitre 3]), le concept de moyenne a donné lieu aujourd'hui à un champs d'étude très vaste avec une variété impressionnante d'applications. En fait, une abondante littérature sur les propriétés de plusieurs moyennes (tels que la moyenne arithmétique, géométrique, etc.) a déjà été écrite, surtout depuis le 19ème siècle, et continue à se developper aujourd'hui. Un excellent panorama du domaine peut être trouvé dans Frosini [44]. Voir aussi le remarquable ouvrage de Bullen *et al.* [18].

La première définition moderne de la moyenne est probablement due à Cauchy [19], qui considérait en 1821 la moyenne de n variables indépendantes $x_1, \ldots, x_n$ comme une fonction $M(x_1, \ldots, x_n)$ qui devrait être interne (cf. Definition 2.6) à l'ensemble des valeurs des x_i.

Le concept de moyenne en tant qu'*égaliseur numérique* est habituellement attribué à Chisini [20], qui donna en 1929 la définition suivante (p. 108) :

> Soit $y = g(x_1, \ldots, x_n)$ une fonction de n variables indépendantes $x_1, \ldots, x_n$ représentant des quantités homogènes. Une moyenne de $x_1, \ldots, x_n$ par rapport à la fonction g est un nombre M tel que, si tous les x_i sont remplacés par M, la valeur de la fonction reste inchangée, c'est-à-dire,
>
> $$g(M, \ldots, M) = g(x_1, \ldots, x_n).$$

Lorsque g est la somme, le produit, la somme des carrés, la somme des inverses, ou encore la somme des exponentielles, la solution de l'équation de Chisini correspond respectivement à la moyenne arithmétique, la moyenne géométrique, la moyenne quadratique, la moyenne harmonique, et la moyenne exponentielle. Malheureusement, comme l'a remarqué de Finetti [26, p. 378], la définition de Chisini est si générale qu'elle n'implique même pas que la "moyenne"—en supposant qu'il existe une solution réelle à l'équation de Chisini—soit une fonction interne au sens de Cauchy.

La citation suivante de Ricci [90, p. 39] pourrait également être considérée comme une autre critique possible de l'approche de Chisini :

> ... lorsque toutes les valeurs deviennent égales, la moyenne devient cette valeur commune. La proposition inverse n'est pas vraie. Si une fonction de plusieurs variables prend leur valeur commune lorsque toutes les variables coïncident, ce n'est pas une condition suffisante pour appeler cette fonction une moyenne. Par exemple, la fonction
>
> $$g(x_1, x_2, \ldots, x_n) = x_n + (x_n - x_1) + (x_n - x_2) + \cdots + (x_n - x_{n-1})$$
>
> vaut x_n lorsque $x_1 = \cdots = x_n$, mais est même supérieure à x_n chaque fois que x_n est supérieur à n'importe laquelle des autres variables.

En 1930, Kolmogoroff [59] et Nagumo [83] considéraient que la moyenne devrait être beaucoup plus que simplement une fonction interne ou un égaliseur numérique. Ils ont alors défini une *valeur moyenne* comme une suite décomposable (cf. Definition 2.10) de fonctions

$$M^{(1)}(x_1) = x_1, M^{(2)}(x_1, x_2), \ldots, M^{(n)}(x_1, \ldots, x_n), \ldots$$

qui sont continues, symétriques, strictement croissantes, et idempotentes. Ils ont ensuite démontré, indépendamment l'un de l'autre, que ces conditions sont nécessaires et suffisantes pour la quasi-arithméticité de la moyenne, c'est-à-dire, pour l'existence d'une fonction f continue et strictement monotone telle que $M^{(n)}$ soit de la forme

$$M^{(n)}(x_1, \ldots, x_n) = f^{-1}\Big[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big] \tag{1}$$

pour tout $n \in \mathbb{N} \setminus \{0\}$.

Les *moyennes quasi-arithmétiques* (1) comprennent la plupart des moyennes algébriques connues ; voir Table 1. Cependant, certaines moyennes, comme la médiane, ne font pas partie de cette catégorie.

$f(x)$	$M^{(n)}(x_1,\ldots,x_n)$	nom
x	$\frac{1}{n}\sum_{i=1}^{n} x_i$	arithmétique
x^2	$\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)^{1/2}$	quadratique
$\log x$	$\left(\prod_{i=1}^{n} x_i\right)^{1/n}$	géométrique
x^{-1}	$\dfrac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}}$	harmonique
x^α $(\alpha \in \mathbb{R}\setminus\{0\})$	$\left(\frac{1}{n}\sum_{i=1}^{n} x_i^\alpha\right)^{1/\alpha}$	puissance
$e^{\alpha x}$ $(\alpha \in \mathbb{R}\setminus\{0\})$	$\frac{1}{\alpha}\ln\left(\frac{1}{n}\sum_{i=1}^{n} e^{\alpha x_i}\right)$	exponentielle

Tab. 1 – Exemples de moyennes quasi-arithmétiques

Les propriétés ci-dessus, définissant une valeur moyenne, sont assez naturelles. Par exemple, on peut facilement voir que, pour les moyennes non décroissantes, l'idempotence est équivalente à l'internalité de Cauchy, et ces deux propriétés sont acceptées par tous les statisticiens comme des conditions minimales pour définir une moyenne.

La propriété de décomposabilité des moyennes est assez naturelle. Lorsqu'elle est associée à l'idempotence, elle peut s'écrire

$$\begin{gathered} M^{(k)}(x_1,\ldots,x_k) = M^{(k)}(x'_1,\ldots,x'_k) \\ \Downarrow \\ M^{(n)}(x_1,\ldots,x_k,x_{k+1},\ldots,x_n) = M^{(n)}(x'_1,\ldots,x'_k,x_{k+1},\ldots,x_n) \end{gathered}$$

ce qui signifie que la moyenne ne change pas lorsqu'on modifie certaines valeurs sans modifier leur moyenne partielle.

L'objectif de cette section n'est pas de présenter un état de l'art de tous les résultats connus de ce vaste royaume des moyennes. Nous ne faisons ici qu'effleurer la surface du sujet en mettant en évidence des caractérisations axiomatiques pour les familles de moyennes les plus connues et les plus souvent utilisées.

Les médianes et, plus généralement, les statistiques d'ordre, qui sont des moyennes particulières, construites pour agréger des valeurs ordinales, seront brièvement presentées à la section 7.

3.1 Moyennes quasi-arithmétiques

Comme nous venons de le mentionner, les moyennes quasi-arithmétiques ont été introduites à l'aide d'une axiomatique très naturelle. Dans cette sous-section, nous étudions ces moyennes en tant que fonctions à n variables, mais aussi en tant que suites de fonctions. Des résultats sur cette classe de moyennes peuvent aussi être trouvés dans [18, Chapitre 4].

Il a été demontré par Aczél [2] (voir aussi [4, §6.4] et [6, Chapitre 17]) que les moyennes quasi-arithmétiques sont les seules fonctions $M : E^n \to E$ qui soient symétriques, continues, strictement croissantes, idempotentes et bisymétriques. L'énoncé de ce résultat peut être formulé comme suit :

Théorème 3.1 *$M : E^n \to E$ est une fonction symétrique, continue, strictement croissante, idempotente et bisymétrique si et seulement s'il existe une fonction $f : E \to \mathbb{R}$ continue et strictement monotone telle que*

$$M(x) = f^{-1}\Big[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big] \qquad (x \in E^n). \tag{2}$$

Les *moyennes quasi-arithmétiques* (2) sont des fonctions d'agrégation internes et couvrent un large spectre de moyennes, comprenant les moyennes arithmétiques, quadratiques, géométriques, et harmoniques ; voir Table 1.

La fonction f apparaissant dans (2) est appelée *générateur* de M. On peut montrer que f est déterminé à une transformation linéaire près : avec $f(x)$, toute fonction

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

définit le même M, et uniquement les fonctions de cette forme.

En plus de ce résultat d'Aczél, nous avons également celui de Kolmogoroff-Nagumo que nous rappelons ici :

Théorème 3.2 *La suite $(M^{(n)} : E^n \to E)_{n\geq 1}$ est une suite décomposable de fonctions symétriques, continues, strictement croissantes et idempotentes si et seulement s'il existe une fonction $f : E \to \mathbb{R}$ continue et strictement monotone telle que*

$$M^{(n)}(x) = f^{-1}\Big[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big] \qquad (x \in E^n).$$

Nagumo [83] a étudié certaines sous-familles de la classe des moyennes quasi-arithmétiques. Il a démontré le résultat suivant (voir aussi [5, §4] et [6, Chapitre 15]) :

Proposition 3.1 *Supposons $E =]0, \infty[$ ou un sous-intervalle.*
(i) $M : E^n \to E$ est une moyenne quasi-arithmétique signifiante pour les mêmes échelles de ratio entrées-sorties si et seulement si

– *soit M est la moyenne géométrique :*

$$M(x) = \Big(\prod_{i=1}^{n} x_i\Big)^{\frac{1}{n}} \qquad (x \in E^n),$$

– *ou M est la moyenne puissance : il existe $\alpha \in \mathbb{R} \setminus \{0\}$ tel que*

$$M(x) = \Big(\frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha}\Big)^{\frac{1}{\alpha}} \qquad (x \in E^n). \tag{3}$$

(ii) $M : E^n \to E$ est une moyenne quasi-arithmétique signifiante pour les mêmes échelles d'intervalle entrées-sorties si et seulement si M est la moyenne arithmétique.

Notons $M_{(\alpha)}$ la moyenne puissance (3) générée par $\alpha \in \mathbb{R} \setminus \{0\}$. Il est bien connu [13, §16] que, si $\alpha_1 < \alpha_2$ alors $M_{(\alpha_1)}(x) \leq M_{(\alpha_2)}(x)$ pour tout $x \in]0, +\infty[^n$ (égalité si et seulement si tous les x_i sont égaux).

Cette famille particulière de moyennes a été étudiée par Dujmović [35, 36] et plus tard par Dyckhoff et Pedrycz [37]. Elle comprend la plupart des moyennes traditionnelles : la moyenne arithmétique $M_{(1)}$, la moyenne harmonique $M_{(-1)}$, la moyenne quadratique $M_{(2)}$, et trois cas limites : la moyenne géométrique $M_{(0)}$, le minimum $M_{(-\infty)}$ et le maximum $M_{(+\infty)}$ (voir par exemple [1]).

En revenant au Théorème 3.1, notons qu'Aczél [2] a aussi étudié le cas où la symétrie et l'idempotence sont omises (voir aussi [4, §6.4] et [6, Chapitre 17]). Il a obtenu le résultat suivant :

Théorème 3.3 *(i) $M : E^n \to E$ est une fonction continue, strictement croissante, idempotente, et bisymétrique si et seulement s'il existe une fonction $f : E \to \mathbb{R}$ continue et strictement monotone et des nombres réels $\omega_1, \ldots, \omega_n > 0$ vérifiant $\sum_i \omega_i = 1$ tels que*

$$M(x) = f^{-1}\Big[\sum_{i=1}^{n} \omega_i\, f(x_i)\Big] \qquad (x \in E^n). \tag{4}$$

(ii) $M : E^n \to E$ est une fonction continue, strictement croissante, et bisymétrique si et seulement s'il existe une fonction $f : E \to \mathbb{R}$ continue et strictement monotone et des nombres réels $p_1, \ldots, p_n > 0$ et $q \in \mathbb{R}$ tels que

$$M(x) = f^{-1}\Big[\sum_{i=1}^{n} p_i\, f(x_i) + q\Big] \qquad (x \in E^n). \tag{5}$$

Les *moyennes quasi-linéaires* (4) et les *fonctions quasi-linéaires* (5) sont des fonctions d'agrégation pondérées. L'unicité vis-à-vis de f est discutée en détail dans [4, §6.4]. Quelques cas particulier de moyennes quasi-linéaires sont présentés dans la Table 2.

3.2 Moyennes lagrangiennes et moyennes de Cauchy

Considérons le point intermédiaire M dans la formule classique du théorème des accroissements finis de Lagrange

$$F(y) - F(x) = F'(M)(y - x) \qquad (x, y \in E), \tag{6}$$

$f(x)$	$M(x)$	nom de la moyenne pondérée
x	$\sum_{i=1}^{n} \omega_i\, x_i$	arithmétique
x^2	$\left(\sum_{i=1}^{n} \omega_i\, x_i^2\right)^{1/2}$	quadratique
$\log x$	$\prod_{i=1}^{n} x_i^{\omega_i}$	géométrique
x^α $(\alpha \in \mathbb{R} \setminus \{0\})$	$\left(\sum_{i=1}^{n} \omega_i\, x_i^\alpha\right)^{1/\alpha}$	puissance

Tab. 2 – Exemples de moyennes quasi-linéaires

comme une fonction des variables x, y, avec la convention $M(x, x) = x$, où $F : E \to \mathbb{R}$ est une fonction continûment dérivable et strictement convexe ou concave. En reformulant cette définition en termes d'intégrales au lieu de dérivées, nous pouvons réécrire (6) comme

$$M(x,y) = \begin{cases} f^{-1}\left(\dfrac{1}{y-x}\displaystyle\int_x^y f(\xi)d\xi\right), & \text{si } x \neq y, \\ x, & \text{si } x = y, \end{cases} \tag{7}$$

pour $x, y \in E$, où $f : E \to \mathbb{R}$ est une fonction continue et strictement monotone. Cette fonction $M(x, y)$ est appelée la *moyenne lagrangienne* associée à f ; voir par exemple [15] et [18, p. 343]. L'unicité du générateur est la même que pour les moyennes quasi-arithmétiques, c'est-à-dire, défini à une même transformation linéaire près ; voir [15, Corollaire 7] et [75, Théorème 1].

Plusieurs moyennes classiques sont lagrangiennes. Les moyennes arithmétique et géométrique correspondent à prendre respectivement $f(x) = x$ et $f(x) = 1/x^2$ dans (7). Cependant la moyenne harmonique n'est pas lagrangienne.

En général, certaines des moyennes les plus communes sont à la fois quasi-arithmétiques et lagrangiennes, mais il y a des moyennes quasi-arithmétiques, comme la moyenne harmonique, qui ne sont pas lagrangiennes. Inversement, la moyenne logarithmique

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

est un exemple de moyenne lagrangienne ($f(x) = 1/x$) qui n'est pas quasi-arithmétique.

Considérons à présent le théorème de la valeur moyenne de Cauchy, qui s'énonce de la manière suivante : pour toutes fonctions F et g, continues sur un intervalle

$[x, y]$ et dérivables sur $]x, y[$, il existe $M \in]a, b[$ tel que

$$\frac{F(y) - F(x)}{g(y) - g(x)} = \frac{F'(M)}{g'(M)}$$

Si les fonctions g et $f := F'/g'$ sont strictement monotone sur $]x, y[$, la valeur moyenne $M(x, y)$ est unique et peut être écrite comme

$$M(x,y) = \begin{cases} f^{-1}\left(\dfrac{1}{g(y) - g(x)} \displaystyle\int_x^y f(\xi) dg(\xi)\right), & \text{si } x \neq y, \\ x, & \text{si } x = y, \end{cases}$$

pour $x, y \in E$. Elle est alors appelée la *moyenne de Cauchy associée au couple* (f, g) ; voir [16]. Une telle moyenne est continue, symétrique, idempotente et strictement croissante.

Lorsque $g = f$ (resp. g est la fonction identité), on retrouve la moyenne quasi-arithmétique (resp. lagrangienne) générée par f. La *moyenne anti-lagrangienne* [16] est obtenue lorsque f est la fonction identité. Par exemple, la moyenne harmonique est une moyenne anti-lagrangienne générée par la fonction $g = 1/x^2$. Les générateurs d'une même moyenne anti-lagrangienne sont définis à une même transformation linéaire près.

4 Fonctions d'agrégation associatives

Avant de présenter des axiomatiques sur les fonctions associatives, nous rappelons quelques concepts bien utiles. Un *semi-groupe* (E, A) est un ensemble E muni d'une opération associative $A : E^2 \to E$. Comme précédemment, nous supposerons que E est un intervalle réel, borné ou non. Un élément $e \in E$ est

a) une *identité* pour A si $A(e, x) = A(x, e) = x$ pour tout $x \in E$,

b) un *zéro* (ou *annihilateur*) pour A si $A(e, x) = A(x, e) = e$ pour tout $x \in E$,

c) un *idempotent* pour A si $A(e, e) = e$.

Pour tout semi-groupe (E, A), il est clair qu'il y a au plus une identité et au plus un zéro pour A dans E, et les deux sont idempotents.

Nous introduisons également le concept de *somme ordinale*, bien connu en théorie des semi-groupes (voir par exemple [22, 60]).

Définition 4.1 *Soit K un ensemble totalement ordonné et soit $\{(E_k, A_k) \mid k \in K\}$ une collection de semi-groupes disjoints indexés par K. Alors la* somme ordinale *de $\{(E_k, A_k) \mid k \in K\}$ est définie par l'union $\cup_{k \in K} E_k$ sous l'opération binaire suivante :*

$$A(x,y) = \begin{cases} A_k(x,y), & \text{si } \exists\, k \in K \text{ tel que } x, y \in E_k \\ \min(x,y), & \text{si } \exists\, k_1, k_2 \in K, k_1 \neq k_2 \text{ tels que } x \in E_{k_1} \text{ et } y \in E_{k_2}. \end{cases}$$

La somme ordinale est un semi-groupe sous l'opération définie ci-dessus.

4.1 Fonctions strictement croissantes

En étudiant les solutions continues et strictement croissantes sur E^2 de l'équation fonctionnelle d'associativité (cf. Définition 2.8), Aczél [3] fut a l'origine du résultat suivant (voir aussi [4, Sect. 6.2]).

Théorème 4.1 *Soit E un intervalle réel, borné ou non, ouvert sur une extrémité. $A : E^2 \to E$ est continu, strictement croissant et associatif si et seulement s'il existe une fonction $f : E \to \mathbb{R}$ continue et strictement monotone telle que*

$$A(x,y) = f^{-1}[f(x) + f(y)] \qquad ((x,y) \in E^2). \tag{8}$$

Il a aussi été démontré que la fonction f apparaissant dans (8) est unique à une constante multiplicative près, c'est-à-dire, avec $f(x)$ toute fonction $g(x) = r\, f(x)$ ($r \in \mathbb{R} \setminus \{0\}$) représente le même A, et uniquement les fonctions de ce type.

De plus, la fonction f est telle que, si $e \in E$ alors

$$A(e,e) = e \quad \Leftrightarrow \quad f(e) = 0. \tag{9}$$

De là, et vu la stricte monotonie de f, il y a au plus un idempotent pour A (qui est l'identité en fait) et donc A ne peut être idempotent. Ainsi, il n'existe aucune fonction qui soit simultanément continue, strictement croissante, idempotente et associative. Cependant, on peut remarquer que toutes les fonctions continues, strictement croissantes et associatives sont symétriques. La somme ($f(x) = x$) et le produit ($f(x) = \log x$) sont des exemples bien connus de fonctions continues, strictement croissantes et associatives.

Selon Ling [60], tout semi-groupe (E, M) vérifiant les hypothèses du Théorème 4.1 est dit *Aczélien*.

Puisque toute suite associative de fonctions $(A^{(n)} : E^n \to E)_{n\geq 1}$ est univoquement déterminée par sa fonction à deux variables, nous avons immédiatement le résultat suivant :

Corollaire 4.1 *Soit E un interval réel, borné ou non, ouvert sur une extrémité. $(A^{(n)} : E^n \to E)_{n\geq 1}$ est une suite associative de fonctions continues et strictement croissantes si et seulement s'il existe une fonction $f : E \to \mathbb{R}$ continue et strictement monotone telle que, pour tout $n \in \mathbb{N} \setminus \{0\}$,*

$$A^{(n)}(x) = f^{-1}\Big[\sum_{i=1}^{n} f(x_i)\Big] \qquad (x \in E^n).$$

4.2 Semi-groupes Archimédiens

Certains auteurs ont tenté de généraliser le Théorème 4.1 en relâchant la stricte croissance en la non décroissance. Il semble cependant que la classe des fonctions continues, non décroissantes et associatives n'ait pas encore été décrite. Toutefois, sous certaines conditions, des résultats ont été obtenus.

D'abord, nous énonçons une représentation qui est souvent attribuée à Ling [60]. A vrai dire, son principal théorème peut se déduire facilement de résultats connus précédemment sur la topologie des semi-groupes ; voir Faucett [38] et Mostert et Shields [79]. Néanmoins, l'avantage de l'approche de Ling est double : (i) le traitement de deux cas différents par une approche unifiée et (ii) des démonstrations élémentaires.

Théorème 4.2 *Soit $E = [a, b]$. $A : E^2 \to E$ est continu, non décroissant, associatif et*

$$A(b, x) = x \qquad (x \in E) \tag{10}$$

$$A(x, x) < x \qquad (x \in E^\circ) \tag{11}$$

si et seulement s'il existe une fonction $f : E \to [0, +\infty]$ continue et strictement décroissante, avec $f(b) = 0$, telle que

$$A(x, y) = f^{-1}[\min(f(x) + f(y), f(a))] \qquad (x, y \in E). \tag{12}$$

Le fait que E soit fermé n'est pas réellement une restriction. Si E est un intervalle réel, borné ou non, avec b pour extrémité supérieure (b peut être $+\infty$), alors nous pouvons remplacer la condition (10) par

$$\lim_{t \to b^-} A(t, t) = b, \quad \lim_{t \to b^-} A(t, x) = x \qquad (x \in E).$$

Toute fonction f solutionnant l'équation (12) est appelé *générateur additif* (ou simplement *générateur*) de M. De plus, nous pouvons facilement voir que toute fonction A de la forme (12) est symétrique et conjonctive.

La condition (10) exprime que b est une *identité à gauche* pour M. Il s'avère, de (12), que b agit comme une identité, et a comme un zéro. La condition (11) exprime simplement qu'il n'y a pas d'idempotent pour A dans $]a, b[$. En effet, par la non décroissance et (10), nous avons toujours $A(x, x) \leq A(b, x) = x$ pour tout $x \in [a, b]$.

Selon que $f(a)$ est fini ou infini (rappelons que $f(a) \in [0, +\infty]$), A prend une forme bien définie (voir Fodor et Roubens [43, §1.3] et Schweizer et Sklar [99]) :

- $f(a) < +\infty$ si et seulement si A a des *diviseurs de zéro* (c'est-à-dire $\exists\, x, y \in\,]a, b[$ tel que $A(x, y) = a$). Dans ce cas, il existe une fonction $g : [a, b] \to [0, 1]$ continue et strictement croissante, avec $g(a) = 0$ et $g(b) = 1$, telle que

  $$A(x, y) = g^{-1}[\max(g(x) + g(y) - 1, 0)] \qquad (x, y \in [a, b]). \tag{13}$$

  Pour le voir, il suffit de poser $g(x) := 1 - f(x)/f(a)$.
  Pour les suite associatives $(A^{(n)} : [a, b]^n \to [a, b])_{n \geq 1}$, (13) devient

  $$A^{(n)}(x) = g^{-1}\Big[\max\Big(\sum_{i=1}^{n} g(x_i) - n + 1, 0\Big)\Big] \qquad (x \in [a, b]^n,\ n \in \mathbb{N} \setminus \{0\}).$$

- $\lim_{t \to a^+} f(x) = +\infty$ si et seulement si A est strictement croissant sur $]a, b[$. Dans ce cas, il existe une fonction $g : [a, b] \to [0, 1]$ continue et strictement croissante, avec $g(a) = 0$ et $g(b) = 1$, telle que

  $$A(x, y) = g^{-1}[g(x)\, g(y)] \qquad (x, y \in [a, b]), \tag{14}$$

Pour le voir, il suffit de poser $g(x) := \exp(-f(x))$.
Pour les suite associatives $(A^{(n)} : [a,b]^n \to [a,b])_{n\geq 1}$, (14) devient

$$A^{(n)}(x) = g^{-1}\Big[\prod_{i=1}^{n} g(x_i)\Big] \qquad (x \in [a,b]^n,\ n \in \mathbb{N}\setminus\{0\}).$$

Bien sûr, le Théorème 4.2 peut aussi être écrit sous une forme duale comme suit :

Théorème 4.3 *Soit $E = [a,b]$. $A : E^2 \to E$ est continu, non décroissant, associatif et*

$$\begin{aligned} A(a,x) &= x \qquad (x \in E) \\ A(x,x) &> x \qquad (x \in E^\circ) \end{aligned}$$

si et seulement s'il existe une fonction $f : E \to [0,+\infty]$ continue et strictement croissante, avec $f(a) = 0$, telle que

$$A(x,y) = f^{-1}[\min(f(x)+f(y), f(b))] \qquad (x,y \in E). \tag{15}$$

Ici encore, E peut être un intervalle quelconque, éventuellement non borné. Les fonctions A de la forme (15) sont symétriques et disjonctives. Il n'y a aucun idempotent intérieur. L'extrémité inférieure a agit comme une identité et l'extrémité supérieure b agit comme un zéro.

Un fois encore, deux cas peuvent être examinés :

- $f(b) < +\infty$ si et seulement si A a des diviseurs de zéro (c'est-à-dire $\exists\, x,y \in\,]a,b[$ tel que $A(x,y) = b$). Dans ce cas, il existe une fonction $g : [a,b] \to [0,1]$ continue et strictement croissante, avec $g(a) = 0$ et $g(b) = 1$, telle que

$$A(x,y) = g^{-1}[\min(g(x)+g(y), 1)] \qquad (x,y \in [a,b]). \tag{16}$$

  Pour le voir, il suffit de poser $g(x) := f(x)/f(b)$.
  Pour les suite associatives $(A^{(n)} : [a,b]^n \to [a,b])_{n\geq 1}$, (16) devient

$$A^{(n)}(x) = g^{-1}\Big[\min\Big(\sum_{i=1}^{n} g(x_i), 1\Big)\Big] \qquad (x \in [a,b]^n,\ n \in \mathbb{N}\setminus\{0\}).$$

- $\lim_{t\to b^-} f(x) = +\infty$ si et seulement si A est strictement croissant sur $]a,b[$. Dans ce cas, il existe une fonction $g : [a,b] \to [0,1]$ continue et strictement croissante, avec $g(a) = 0$ et $g(b) = 1$, telle que

$$A(x,y) = g^{-1}[1 - (1-g(x))\,(1-g(y))] \qquad (x,y \in [a,b]), \tag{17}$$

  Pour le voir, il suffit de poser $g(x) := 1 - \exp(-f(x))$.
  Pour les suite associatives $(A^{(n)} : [a,b]^n \to [a,b])_{n\geq 1}$, (17) devient

$$A^{(n)}(x) = g^{-1}\Big[1 - \prod_{i=1}^{n}(1-g(x_i))\Big] \qquad (x \in [a,b]^n,\ n \in \mathbb{N}\setminus\{0\}).$$

Tout semi-groupe vérifiant les hypothèses du Théorème 4.2 ou 4.3 est appelé *Archimédien*, voir Ling [60]. En d'autres mots, tout semi-groupe (E, A) est dit *Archimédien* si A est continu, non décroissant et associatif, une extrémité de E étant une identité pour A, et il n'y a pas d'idempotent pour M à l'intérieur de E. Nous pouvons faire la distinction entre les semi-groupes Archimédiens conjonctifs ou disjonctifs selon que l'identité se trouve à l'extrémité supérieure ou inférieure de E, respectivement. Un semi-groupe Archimédien est dit *proprement Archimédien* ou *Aczélien* si tout générateur additif est non borné, sinon il est *improprement Archimédien*.

Ling [60, §6] a démontré que tout semi-groupe Archimédien peut être obtenu comme limite de groupes Aczéliens.

4.3 Une classe de fonctions associatives non décroissantes

Nous présentons maintenant une description des fonctions $A : [a, b]^2 \to [0, 1]$ qui sont continues, non décroissantes, faiblement idempotentes et associatives. Pour tout $\theta \in [a, b]$, nous définissons $\mathcal{A}_{a,b,\theta}$ comme l'ensemble des fonctions $A : [a, b]^2 \to [0, 1]$ continues, non décroissantes, faiblement idempotentes, associatives et telles que $A(a, b) = A(b, a) = \theta$. Les cas extrêmes $\mathcal{A}_{a,b,a}$ et $\mathcal{A}_{a,b,b}$ joueront un rôle important dans la suite. Les résultats de cette sous-section peuvent être trouvés dans [65].

Théorème 4.4 *$A : [a, b]^2 \to [0, 1]$ est continu, non décroissant, faiblement idempotent et associatif si et seulement s'il existe $\alpha, \beta \in [a, b]$ et deux fonctions $A_{a,\alpha\wedge\beta,\alpha\wedge\beta} \in \mathcal{A}_{a,\alpha\wedge\beta,\alpha\wedge\beta}$ et $A_{\alpha\vee\beta,b,\alpha\vee\beta} \in \mathcal{A}_{\alpha\vee\beta,b,\alpha\vee\beta}$ tels que, pour tous $x, y \in [a, b]$,*

$$A(x, y) = \begin{cases} A_{a,\alpha\wedge\beta,\alpha\wedge\beta}(x, y), & \text{si } x, y \in [a, \alpha \wedge \beta] \\ A_{\alpha\vee\beta,b,\alpha\vee\beta}(x, y), & \text{si } x, y \in [\alpha \vee \beta, b] \\ (\alpha \wedge x) \vee (\beta \wedge y) \vee (x \wedge y), & \text{sinon.} \end{cases}$$

Passons à présent à la description de $\mathcal{A}_{a,b,a}$. Mostert et Shields [79, p. 130, Théorème B] ont démontré le théorème suivant :

Théorème 4.5 *$A : [a, b]^2 \to [a, b]$ est continu, associatif et tel que a agit comme un zéro et b comme une identité si et seulement si*

- *soit*
$$A(x, y) = \min(x, y) \qquad (x, y \in [a, b]),$$
- *ou il existe une fonction continue et strictement décroissante $f : [a, b] \to [0, +\infty]$, avec $f(b) = 0$, tel que*
$$A(x, y) = f^{-1}[\min(f(x) + f(y), f(a))] \qquad (x, y \in [a, b])$$
(semi-groupe Archimédien conjonctif),
- *ou il existe un ensemble dénombrable d'indices $K \subseteq \mathbb{N}$, une famille de sous-intervalles ouverts disjoints $\{]a_k, b_k[\,|\, k \in K\}$ de $[a, b]$ et une famille $\{f_k \,|\, k \in$*

$K\}$ de fonctions continues et strictement décroissantes $f_k : [a_k, b_k] \to [0, +\infty]$, avec $f_k(b_k) = 0$, tels que, pour tous $x, y \in [a, b]$,

$$A(x,y) = \begin{cases} f_k^{-1}[\min(f_k(x) + f_k(y), f_k(a_k))], & \textit{si } \exists\, k \in K \textit{ tel que } x, y \in [a_k, b_k], \\ \min(x,y), & \textit{sinon,} \end{cases}$$

(somme ordinale de semi-groupes Archimédiens conjonctifs).

On peut montrer que $\mathcal{A}_{a,b,a}$ est la famille des fonctions $A : [a, b]^2 \to [a, b]$ continues, non décroissantes, associatives et telles que a agit comme un zéro et b comme une identité. En conséquence, la description de la famille $\mathcal{A}_{a,b,a}$ est aussi donnée par le Théorème 4.5. De plus, il s'avère que toutes les fonctions vérifiant les hypothèses de ce résultat sont symétriques, non décroissantes et conjonctives.

Le Théorème 4.5 peut aussi être écrit sous une forme duale comme suit :

Théorème 4.6 *$A : [a, b]^2 \to [a, b]$ est continu, associatif et tel que a agit comme une identité et b comme un zéro si et seulement si*

- *soit*

$$A(x, y) = \max(x, y) \qquad (x, y \in [a, b]),$$

- *ou il existe une fonction continue et strictement croissante $f : [a, b] \to [0, +\infty]$, avec $f(a) = 0$, tel que*

$$A(x, y) = f^{-1}[\min(f(x) + f(y), f(b))] \qquad (x, y \in [a, b])$$

  (semi-groupe Archimédien disjonctif),

- *ou il existe un ensemble dénombrable d'indices $K \subseteq \mathbb{N}$, une famille de sous-intervalles ouverts disjoints $\{]a_k, b_k[\mid k \in K\}$ de $[a, b]$ et une famille $\{f_k \mid k \in K\}$ de fonctions continues et strictement croissantes $f_k : [a_k, b_k] \to [0, +\infty]$, avec $f_k(a_k) = 0$, tels que, pour tous $x, y \in [a, b]$,*

$$A(x,y) = \begin{cases} f_k^{-1}[\min(f_k(x) + f_k(y), f_k(b_k))], & \textit{si } \exists\, k \in K \textit{ tel que } x, y \in [a_k, b_k], \\ \max(x,y), & \textit{sinon,} \end{cases}$$

  (somme ordinale de semi-groupes Archimédiens disjonctifs).

Comme précédemment, $\mathcal{A}_{a,b,b}$ est la famille des fonctions $A : [a, b]^2 \to [a, b]$ continues, non décroissantes, associatives et telles que a agit comme une identité et b comme un zéro. La description de la famille $\mathcal{A}_{a,b,b}$ est donc donnée par le Théorème 4.6. De plus, toutes les fonctions vérifiant les hypothèses de ce résultat sont symétriques, non décroissantes et disjonctives.

Les Théorèmes 4.4, 4.5 et 4.6 réunis fournissent une description complète de la famille des fonctions $A : [a, b]^2 \to [a, b]$ continues, non décroissantes, faiblement idempotentes et associatives. En imposant des conditions supplémentaires, on obtient les corollaires qui suivent :

Corollaire 4.2 *$A : [a, b]^2 \to [a, b]$ est continu, strictement croissant, faiblement idempotent et associatif si et seulement s'il existe une fonction $g : [a, b] \to [0, 1]$ continue et strictement croissante, avec $g(a) = 0$ et $g(b) = 1$ tel que*

– *soit*

$$A(x,y) = g^{-1}[g(x)\, g(y)] \qquad (x,y \in [a,b]),$$

– *ou*

$$A(x,y) = g^{-1}[g(x) + g(y) - g(x)\, g(y)] \qquad (x,y \in [a,b]).$$

Corollaire 4.3 *$A : [a,b]^2 \to [a,b]$ est symétrique, continu, non décroissant, faiblement idempotent et associatif si et seulement s'il existe $\alpha \in [a,b]$ et deux fonctions $A_{a,\alpha,\alpha} \in \mathcal{A}_{a,\alpha,\alpha}$ et $A_{\alpha,b,\alpha} \in \mathcal{A}_{\alpha,b,\alpha}$ tels que, pour tous $x,y \in [a,b]$,*

$$A(x,y) = \begin{cases} A_{a,\alpha,\alpha}(x,y), & \text{si } x,y \in [a,\alpha] \\ A_{\alpha,b,\alpha}(x,y), & \text{si } x,y \in [\alpha,b] \\ \alpha, & \text{sinon.} \end{cases}$$

Corollaire 4.4 *$A : [a,b]^2 \to [a,b]$ est continu, non décroissant, faiblement idempotent, associatif et a exactement une identité dans $[a,b]$ si et seulement si $A \in \mathcal{A}_{a,b,a} \cup \mathcal{A}_{a,b,b}$.*

4.4 Fonctions associatives internes

Passons maintenant aux fonction associatives internes ou, en quelque sorte, aux moyennes associatives. Comme ces fonctions sont toutes idempotentes, nous étudions les fonctions associatives idempotentes. Bien que nous ayons déjà observé qu'il n'existe aucune fonction continue, strictement croissante, idempotente et associative, la classe des fonctions continues, non décroissantes, idempotentes et associatives n'est pas vide et peut être décrite à partir du Théorème 4.4. Cependant, Fodor [42] avait déjà obtenu cette description dans un cadre plus général. Le théorème est le suivant :

Théorème 4.7 *Soit E un intervalle réel, borné ou non. $A : E^2 \to E$ est continu, non décroissant, idempotent et associatif si et seulement s'il existe $\alpha, \beta \in E$ tel que*

$$A(x,y) = (\alpha \wedge x) \vee (\beta \wedge y) \vee (x \wedge y) \qquad ((x,y) \in E^2). \tag{18}$$

Notons que, par la distributivité de $\wedge$ et $\vee$, A peut aussi être écrit sous la forme équivalente :

$$A(x,y) = (\beta \vee x) \wedge (\alpha \vee y) \wedge (x \vee y) \qquad ((x,y) \in E^2).$$

Pour les suites associatives de fonctions, le résultat peut être formulé comme suit :

Théorème 4.8 *Soit E un intervalle réel, borné ou non. $(A^{(n)} : E^n \to E)_{n\geq 1}$ est une suite associative de fonctions continues, non décroissantes et idempotentes si et seulement s'il existe $\alpha, \beta \in E$ tel que*

$$A^{(n)}(x) = (\alpha \wedge x_1) \vee \Big(\bigvee_{i=2}^{n-1} (\alpha \wedge \beta \wedge x_i)\Big) \vee (\beta \wedge x_n) \vee \Big(\bigwedge_{i=1}^{n} x_i\Big) \qquad (x \in E^n,\ n \in \mathbb{N}\setminus\{0\}).$$

Avant Fodor [42], le cas des fonctions symétriques avait été obtenu par Fung and Fu [45] et d'une manière plus concise par Dubois et Prade [31]. Le résultat peut être formulé comme ceci :

Théorème 4.9 *Soit E un intervalle réel, borné ou non.*
i) $A : E^2 \to E$ est symétrique, continu, non décroissant, idempotent et associatif si et seulement s'il existe $\alpha \in E$ tel que

$$A(x, y) = \text{médiane}(x, y, \alpha) \qquad (x, y \in E).$$

ii) $(A^{(n)} : E^n \to E)_{n\geq 1}$ est une suite associative de fonctions symétriques, continues, non décroissantes et idempotentes si et seulement s'il existe $\alpha \in E$ tel que

$$A^{(n)}(x) = \text{médiane}\Big(\bigwedge_{i=1}^{n} x_i, \bigvee_{i=1}^{n} x_i, \alpha\Big) \qquad (x \in E^n,\ n \in \mathbb{N} \setminus \{0\}). \tag{19}$$

Les trois théorèmes précédents montrent que l'idempotence est rarement consistente avec l'associativité. Par exemple, la moyenne associative (19) n'est pas très décisive puisqu'elle conduit à une valeur prédéfinie α dès qu'il existe $x_i \leq \alpha$ et $x_j \geq \alpha$.

Czogała et Drewniak [23] ont examiné le cas où A possède une identité $e \in E$. Ils ont obtenu le résultat suivant :

Théorème 4.10 *Soit E un intervalle réel, borné ou non.*
i) Si $A : E^2 \to E$ est non décroissant, idempotent, associatif et a une identité $e \in E$, alors il existe une fonction décroissante $g : E \to E$ avec $g(e) = e$ telle que, pour tous $x, y \in E$,

$$A(x, y) = \begin{cases} x \wedge y, & \text{si } y < g(x) \\ x \vee y, & \text{si } y > g(x) \\ x \wedge y \text{ ou } x \vee y, & \text{si } y = g(x). \end{cases}$$

ii) Si $A : E^2 \to E$ est continu, non décroissant, idempotent, associatif et a une identité $e \in E$, alors $A = \min$ ou $\max$.

4.5 t-normes, t-conormes, uninormes

Dans la théorie des ensembles flous, un des principaux sujets consiste à définir des connecteurs logiques flous qui sont des extensions appropriées des connecteurs logiques 'ET', 'OU' et 'NON' dans le cas où l'ensemble des valeurs est l'intervalle unité $[0, 1]$ au lieu de la paire $\{0, 1\}$.

Les connecteurs flous modélisant les 'ET' et 'OU' sont appelés *normes triangulaires* (ou t-normes) et *conormes triangulaires* (t-conormes), respectivement ; voir [11, 99].

Définition 4.2 *i) Une t-norme est une fonction $T : [0, 1]^2 \to [0, 1]$ symétrique, non décroissante, associative et ayant 1 comme identité.*

ii) Une t-conorme est une fonction $S : [0, 1]^2 \to [0, 1]$ symétrique, non décroissante, associative et ayant 0 comme identité.

L'étude de ces fonctions a commencé avec Schweizer et Sklar [97, 98] et Ling [60] ; voir aussi Dubois et Prade [32]. Aujourd'hui, il existe une littérature très abondante sur le domaine ; voir le livre de Klement, Mesiar et Pap [58].

Bien sûr, la famille des t-normes continues n'est rien d'autre que la classe $\mathcal{A}_{0,1,0}$, et la famille des t-conormes continue est la classe $\mathcal{A}_{0,1,1}$. Ces deux classes ont été complètement décrites dans cette section. De plus, le Corollaire 4.4 fournit une caractérisation de leur union.

Corollaire 4.5 $A : [0,1]^2 \to [0,1]$ *est continu, non décroissant, faiblement idempotent, associatif et a exactement une identité dans* $[0,1]$ *si et seulement si* A *est une* t*-norme continue ou une* t*-conorme continue.*

Il est bien connu que les t-normes et t-conormes sont largement étudiées en théorie des ensembles flous, surtout dans la modélisation des connecteurs flous et des implications floues (voir [105]). Les applications à des problèmes pratiques requièrent l'utilisation des t-normes ou t-conormes les plus appropriées. Sur ce sujet, Fodor [40] a présenté une méthode de construction de nouvelles t-normes à partir de t-normes.

A noter aussi que certaines propriétés des t-normes, telles que l'associativité, ne jouent pas de rôle essentiel dans la modélisation des préférences et la théorie du choix. Récemment, certains auteurs [10, 37, 110] ont étudié des opérations binaires non associatives sur $[0,1]$ dans différents contextes. Ces opérateurs peuvent être vus comme des généralisations de t-normes et t-conormes dans le sens qu'ils incluent ces derniers. De plus, Fodor [41] a défini et étudié le concept de t-normes faibles. Ses résultats ont été utilement appliqués dans le cadre des relations de préférence floues strictes.

Depuis peu, d'autres fonctions associatives ont été introduites et étudiées : les *t-opérateurs* [72] et les *uninormes* [108] (voir aussi [73, 74]), qui se sont avérées utiles dans les systèmes experts, les réseaux neuronaux et la théorie des quantificateurs flous.

Définition 4.3 $i)$ *Un* t*-operateur est une fonction* $F : [0,1]^2 \to [0,1]$ *symétrique, non décroissante, associative, ayant* 0 *et* 1 *comme éléments idempotents et telle que les sections* $x \mapsto F(x,0)$ *et* $x \mapsto F(x,1)$ *sont continues sur* $[0,1]$.

$ii)$ *Une uninorme est une fonction* $U : [0,1]^2 \to [0,1]$ *symétrique, non décroissante, associative et ayant une identité.*

Il est clair qu'une uninorme devient une t-norme (resp. t-cornorme) lorsque l'identité est 1 (resp. 0).

Nous n'insisterons pas sur ce sujet des t-normes, t-conormes, et uninormes. Le lecteur intéressé consultera le remarquable ouvrage de Klement, Mesiar et Pap [58].

Pour des résultats encore plus récents, signalons un excellent article sur les fonctions associatives par Sander [94].

5 Intégrales non additives

De nombreuses fonctions d'agrégation peuvent être vues comme des intégrales discrètes non additives par rapport à des mesures non additives. Dans cette section, nous introduisons principalement les intégrales de Choquet et de Sugeno. Le lecteur trouvera plus de détails sur ce domaine dans le chapitre **??** de ce volume.

5.1 Motivations

Un des aspects significatifs dans les problèmes d'agrégation est la prise en compte de l'importance des attributs ou critères considérés, laquelle est habituellement modélisée par l'utilisation de poids. Puisque ces poids doivent être pris en compte durant la phase d'agrégation, il est nécessaire d'utiliser des fonctions d'agrégation pondérées, abandonnant ainsi la propriété de symétrie. Jusqu'à récemment, les fonctions d'agrégation pondérées les plus utilisées étaient des fonctions de type moyennes, tels que les moyennes quasi-linéaires (4).

Cependant, les moyennes arithmétiques pondérées et, plus généralement, les moyennes quasi-linéaires présentent certaines faiblesses. Aucune de ces fonctions n'est capable de modéliser une quelconque interaction parmi les attributs. En effet, il est bien connu en théorie de l'utilité multiattribut (MAUT) que ces fonctions conduisent à l'*indépendance préférentielle mutuelle* (voir par exemple [39]) parmi les attributs, qui exprime, dans un certain sens, l'indépendance des attributs. Comme ces fonctions ne sont pas appropriées en présence d'attributs dépendants, la tendance a été de construire des attributs censés être indépendants, ce qui entraînait souvent des erreurs dans les évaluations.

Dans le but d'obtenir une représentation flexible des phenomènes complexes d'interaction parmi les attributs ou critères (par exemple, une synergie positive ou négative entre certains critères), il s'est avéré utile de remplacer le vecteur poids par une fonction d'ensemble non additive, permettant ainsi de définir un poids non seulement sur chaque critère, mais aussi sur chaque sous-ensemble de critères.

C'est dans ce but que l'utilisation des *mesures floues* a été proposée par Sugeno [102] pour généraliser les mesures additives. Il est maintenant bien connu que, dans de nombreuses situations du monde réel, l'additivité n'est pas une propriété appropriée pour les fonctions d'ensemble, à cause de l'absence d'additivité dans de nombreuses facettes du raisonnement humain. Pour pouvoir exprimer la subjectivité humaine, Sugeno proposa de remplacer la propriété d'additivité des fonctions d'ensemble par la monotonie et appela ces mesures monotone non additive des mesures floues.

Considérons l'ensemble des n indices $N = \{1, \ldots, n\}$. Selon les applications considérées, ces indices peuvent représenter des attributs, des critères, des juges, des experts, des votants, etc. Pour souligner le fait que N a n éléments, nous écrirons parfois N_n.

Définition 5.1 *Une mesure floue sur N est une fonction d'ensemble $\mu : 2^N \to [0,1]$ qui est monotone, c'est-à-dire $\mu(S) \leq \mu(T)$ chaque fois que $S \subseteq T$, et vérifie les conditions limites $\mu(\varnothing) = 0$ et $\mu(N) = 1$.*

Dans les problèmes d'analyse multicritère, le coefficient $\mu(S)$, pour un $S \subseteq N$ donné, est généralement interprété comme le poids ou l'importance de la combinaison S de critères. Ainsi, en plus des poids usuels sur les critères pris séparément, des poids sur toute combinaison de critères sont également définis. La monotonie signifie alors simplement que le fait d'ajouter un nouveau critère à une combinaison ne peut faire décroître son importance. Dans ce chapitre, l'ensembles des mesures floues sur N sera noté $\mathcal{F}_N$.

A partir d'une telle mesure floue, on peut construire une fonction d'agrégation permettant de calculer une sorte de valeur moyenne en prenant en compte les coefficients de la mesure floue. Une telle fonction d'agrégation est une *intégrale floue*, concept introduit par Sugeno [102, 103].

Les intégrales floues sont des intégrales d'une fonction par rapport à une mesure floue, par analogie à l'intégrale de Lebesgue qui est définie par rapport à une mesure ordinaire (additive). Comme l'intégrale d'une fonction représente généralement sa valeur moyenne, une intégrale floue peut être considérée comme un cas particulier de fonction d'agrégation.

Contrairement aux moyennes arithmétiques pondérées, les intégrales floues sont capables de prendre en compte les interaction éventuelles parmi les attributs ou critères. C'est une des raisons pour lesquelles ces intégrales ont été largement étudiées dans les problèmes d'aide multicritère à la décision [48, 50, 51, 52].

Il existe plusieurs classes d'intégrales floues, parmi lesquelles les plus représentatives sont celles de Choquet et Sugeno. Dans cette section, nous étudions de près ces deux types d'intégrales en tant que fonctions d'agrégation. En particulier, nous présentons des caractérisations axiomatiques de ces intégrales. La différence principale entre ces deux intégrales est que la première est appropriée pour agréger des valeurs définies sur une échelle d'intervalle, alors que la seconde est plutôt conçue pour agréger des valeurs définies sur une échelle ordinale.

5.2 L'intégrale de Choquet

L'intégrale de Choquet a été introduite en théorie des capacités [21]. Elle a ensuite été utilisée dans plusieurs contextes, notamment en théorie de l'utilité non additive [46, 95, 96, 104], en théorie des mesures et intégrales floues [25, 56, 80, 81] (voir également l'excellent volume édité [52]), mais aussi en finance [28] et en théorie des jeux [29].

Comme cette intégrale est considérée ici comme une fonction d'agrégation à n variables, nous adopterons la notation d'une simple fonction plutôt que la forme intégrale usuelle, et l'intégrande sera un ensemble ordonné de n valeurs réelles, noté $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Définition 5.2 *Soit $\mu \in \mathcal{F}_N$. L'intégrale de Choquet (discrète) de $x \in \mathbb{R}^n$ par rapport à μ est définie par*

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

où $(\cdot)$ *indique une permutation sur* N *telle que* $x_{(1)} \leq \ldots \leq x_{(n)}$*. D'autre part,* $A_{(i)} = \{(i), \ldots, (n)\}$*, et* $A_{(n+1)} = \varnothing$*.*

Par exemple, si $x_3 \leq x_1 \leq x_2$, nous avons

$$\begin{aligned} \mathcal{C}_\mu(x_1, x_2, x_3) &= x_3\,[\mu(\{3,1,2\}) - \mu(\{1,2\})] \\ &\quad + x_1\,[\mu(\{1,2\}) - \mu(\{2\})] \\ &\quad + x_2\,\mu(\{2\}). \end{aligned}$$

Ainsi, l'intégrale de Choquet discrète est une expression linéaire, à un réarrangement près des arguments. Elle est étroitement liée à l'intégrale de Lebesgue discrète (moyenne arithmétique pondérée) puisque ces deux intégrales coïncident lorsque la mesure floue est additive :

$$\mathcal{C}_\mu(x) = \sum_{i=1}^{n} \mu(i)\, x_i \qquad (x \in \mathbb{R}^n).$$

Dans ce sens, l'intégrale de Choquet est une généralisation de l'intégrale de Lebesgue.

Passons à present aux axiomatiques de l'intégrale de Choquet. Tout d'abord, comme on peut le voir, cette fonction d'agrégation vérifie un certain nombre de propriétés naturelles : elle est continue, non décroissante, unanimement croissante, idempotente, interne, signifiante pour les mêmes échelles d'intervalle entrées-sorties ; voir par exemple [50]. Elle vérifie aussi la propriété d'additivité comonotone [27, 96], c'est-à-dire,

$$f(x_1 + x'_1, \ldots, x_n + x'_n) = f(x_1, \ldots, x_n) + f(x'_1, \ldots, x'_n)$$

pour tous vecteurs comonotones $x, x' \in \mathbb{R}^n$, où deux vecteurs $x, x' \in \mathbb{R}^n$ sont comonotone s'il existe une permutation σ sur N telle que

$$x_{\sigma(1)} \leq \cdots \leq x_{\sigma(n)} \quad \text{et} \quad x'_{\sigma(1)} \leq \cdots \leq x'_{\sigma(n)}.$$

Une justification de cette propriété en aide à la décision multicritère peut être trouvée dans [77, 78].

Le résultat suivant [69, Proposition 4.1] donne une caractérisation de l'intégrale de Choquet à deux variables d'une manière très naturelle :

Proposition 5.1 $f : \mathbb{R}^2 \to \mathbb{R}$ *est non décroissant et signifiant pour les échelles d'intervalle entrées-sorties si et seulement s'il existe* $\mu \in \mathcal{F}_2$ *tel que* $f = \mathcal{C}_\mu$*.*

La classe des intégrales de Choquet à n variables a d'abord été caractérisée par Schmeidler [96], en utilisant l'additivité comonotone ; voir aussi [25], [24], [47], et [53, Théorème 8.6]. Notons que ce résultat avait été énoncé et démontré dans le cas continu (infini) plutôt que dans le cas discret.

Théorème 5.1 $f : \mathbb{R}^n \to \mathbb{R}$ *est non décroissant, additif comonotone, et vérifie l'identité* $f(\mathbf{1}_N) = 1$ *si et seulement s'il existe* $\mu \in \mathcal{F}_N$ *tel que* $f = \mathcal{C}_\mu$*.*

Comme l'intégrale de Choquet est définie à partir d'une mesure floue, il est parfois utile de considérer, pour un ensemble N donné, la famille des intégrales de Choquet sur N comme un ensemble de fonctions

$$\{f_\mu : \mathbb{R}^n \to \mathbb{R} \mid \mu \in \mathcal{F}_N\}$$

ou, de façon équivalente, comme une fonction $f : \mathbb{R}^n \times \mathcal{F}_N \to \mathbb{R}$.

Citons une première caractérisation de la famille des intégrales de Choquet sur N ; voir Groes et col. [55]. Pour tout $S \subseteq N$, $S \neq \varnothing$, notons μ_S la mesure floue sur N définie par $\mu_S(T) = 1$ si $T \supseteq S$ et 0 sinon.

Théorème 5.2 *La classe des fonctions $\{f_\mu : \mathbb{R}^n \to \mathbb{R} \mid \mu \in \mathcal{F}_N\}$ vérifie les propriétés suivantes :*
- *pour tous $\mu, \nu \in \mathcal{F}_N$ et tout $\lambda \in \mathbb{R}$ tels que $\lambda\mu + (1-\lambda)\nu \in \mathcal{F}_N$ on a*

$$f_{\lambda\mu+(1-\lambda)\nu} = \lambda f_\mu + (1-\lambda) f_\nu\,,$$

- *pour tout $S \subseteq N$, on a $f_{\mu_S} = \min_S$,*

si et seulement si $f_\mu = \mathcal{C}_\mu$ pour tout $\mu \in \mathcal{F}_N$.

Une seconde caractérisation, obtenue par l'auteur [62, 63], s'énonce comme suit :

Théorème 5.3 *La classe des fonctions $\{f_\mu : \mathbb{R}^n \to \mathbb{R} \mid \mu \in \mathcal{F}_N\}$ vérifie les propriétés suivantes :*
- *toute fonction f_μ est une expression linéaire de μ, c'est-à-dire qu'il existe 2^n fonctions $g_T : \mathbb{R}^n \to \mathbb{R}$ $(T \subseteq N)$ telles que $f_\mu = \sum_{T\subseteq N} g_T\, \mu(T)$ pour tout $\mu \in \mathcal{F}_N$,*
- *pour tout $\mu \in \mathcal{F}_N$ et tout $S \subseteq N$, on a $f_\mu(\mathbf{1}_S) = \mu(S)$,*
- *pour tout $\mu \in \mathcal{F}_N$, la fonction f_μ est non décroissante et signifiante pour les échelles d'intervalle entrées-sorties,*

si et seulement si $f_\mu = \mathcal{C}_\mu$ pour tout $\mu \in \mathcal{F}_N$.

Ces deux caractérisations sont assez naturelles et, en réalité, assez proches l'une de l'autre. La condition de linéarité proposée dans la seconde caractérisation est utile si l'on veut conserver un modèle d'agrégation aussi simple que possible. Techniquement, cette condition est équivalente à la condition de superposition :

$$f_{\lambda_1\mu+\lambda_2\nu} = \lambda_1 f_\mu + \lambda_2 f_\nu$$

pour tous $\mu, \nu \in \mathcal{F}_N$ et tous $\lambda_1, \lambda_2 \in \mathbb{R}$ tels que $\lambda_1\mu + \lambda_2\nu \in \mathcal{F}_N$. Bien sûr, la linéarité implique la première condition de la première caractérisation. De plus, sous cette condition de linéarité, les autres conditions sont équivalentes. En fait, dans la preuve de la seconde caractérisation [62, 63] l'auteur a remplacé la condition $f_{\mu_S} = \min_S$ par ces trois conditions : $f_\mu(\mathbf{1}_S) = \mu(S)$, non décroissance et signifiance pour les échelles d'intervalle entrées-sorties de f_μ.

Nous avons également les trois résultats suivants [62, §4.2.3] :

Proposition 5.2 *Une intégrale de Choquet* $\mathcal{C}_\mu : \mathbb{R}^n \to \mathbb{R}$ *est bisymétrique si et seulement si*

$$\mathcal{C}_\mu \in \{\min_S, \max_S \mid S \subseteq N\} \cup \{\mathrm{WAM}_\omega \mid \omega \in [0,1]^n\}.$$

Proposition 5.3 *Une suite d'intégrales de Choquet* $\mathcal{C} := (\mathcal{C}^{(n)}_{\mu^{(n)}} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ *est bisymétrique si et seulement si*

- *soit, pour tout* $n \in \mathbb{N} \setminus \{0\}$*, il existe* $S \subseteq N_n$ *tel que* $\mathcal{C}^{(n)}_{\mu^{(n)}} = \min_S$*,*
- *ou, pour tout* $n \in \mathbb{N} \setminus \{0\}$*, il existe* $S \subseteq N_n$ *tel que* $\mathcal{C}^{(n)}_{\mu^{(n)}} = \max_S$*,*
- *ou, pour tout* $n \in \mathbb{N} \setminus \{0\}$*, il existe* $\omega \in [0,1]^n$ *tel que* $\mathcal{C}^{(n)}_{\mu^{(n)}} = \mathrm{WAM}_\omega$*,*

Proposition 5.4 *Une suite d'intégrales de Choquet* $\mathcal{C} := (\mathcal{C}^{(n)}_{\mu^{(n)}} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ *est decomposable si et seulement si*

- *soit* $\mathcal{C} = (\min^{(n)})_{n\geq 1}$*,*
- *ou* $\mathcal{C} = (\max^{(n)})_{n\geq 1}$*,*
- *ou il existe* $\theta \in [0,1]$ *tel que, pour tout* $n \in \mathbb{N} \setminus \{0\}$*, on a* $\mathcal{C}^{(n)}_{\mu^{(n)}} = \mathrm{WAM}_\omega$*, avec*

$$\omega_i = \frac{(1-\theta)^{n-i}\theta^{i-1}}{\sum_{j=1}^{n}(1-\theta)^{n-j}\theta^{j-1}} \qquad (i \in N_n).$$

Proposition 5.5 *Une suite d'intégrales de Choquet* $\mathcal{C} := (\mathcal{C}^{(n)}_{\mu^{(n)}} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ *est associative si et seulement si*

$$\mathcal{C} = (\min^{(n)})_{n\geq 1} \text{ ou } (\max^{(n)})_{n\geq 1} \text{ ou } (P_1^{(n)})_{n\geq 1} \text{ ou } (P_n^{(n)})_{n\geq 1}.$$

Venons-en à présent à certains cas particuliers de l'intégrale de Choquet, à savoir : les moyennes arithmétiques pondérées (WAM) et les fonctions moyennes ordonnées (OWA).

La *moyenne arithmétique pondérée* WAM_ω est une intégrale de Choquet définie à partir d'une mesure additive. Elle vérifie la propriété classique d'additivité :

$$f(x_1 + x_1', \ldots, x_n + x_n') = f(x_1, \ldots, x_n) + f(x_1', \ldots, x_n')$$

pour tous vecteurs $x, x' \in \mathbb{R}^n$. Plus exactement, nous avons les résultats suivant (voir [62, §4.2.4] et [82])

Proposition 5.6 *L'intégrale de Choquet* $\mathcal{C}_\mu : \mathbb{R}^n \to \mathbb{R}$ *est additive si et seulement s'il existe* $\omega \in [0,1]^n$ *tel que* $\mathcal{C}_\mu = \mathrm{WAM}_\omega$*.*

Proposition 5.7 $A : \mathbb{R}^n \to \mathbb{R}$ *est non décroissant, signifiant pour les échelles d'intervalle entrées-sorties et additif si et seulement s'il existe* $\omega \in [0,1]^n$ *tel que* $A = \mathrm{WAM}_\omega$*.*

La *fonction moyenne ordonnée* OWA_ω a été proposée en 1988 par Yager [107]. Depuis son introduction, cette fonction d'agrégation a été utilisée dans de nombreux domaines tels que les réseaux neuronaux, les sytèmes de bases de données, les contrôleurs de logique floue, et l'aide à la décision multicritère. Un panorama sur cette fonction peut être trouvé dans le livre édité [106] ; voir aussi [52].

Le résultat suivant, attribué à Grabisch [49] (voir [67] pour une démonstration concise), montre que la fonction OWA n'est rien d'autre qu'une intégrale de Choquet par rapport à une mesure floue *cardinale*, c'est-à-dire qui ne dépend que du cardinal des sous-ensembles.

Proposition 5.8 *Soit $\mu \in \mathcal{F}_N$. Les assertions suivantes sont équivalentes :*

i) *Pour tous $S, S' \subseteq N$ tels que $|S| = |S'|$, on a $\mu(S) = \mu(S')$.*

ii) *Il existe un vecteur de poids ω tel que $\mathcal{C}_\mu = \mathrm{OWA}_\omega$.*

iii) *$\mathcal{C}_\mu$ est une fonction symétrique.*

La mesure floue μ associée à OWA_ω est donnée par

$$\mu(S) = \sum_{i=n-s+1}^{n} \omega_i \qquad (S \subseteq N,\ \ S \neq \varnothing).$$

Inversement, les poids associés à OWA_ω sont donnés par

$$\omega_{n-s} = \mu(S \cup i) - \mu(S) \qquad (i \in N,\ \ S \subseteq N \setminus i).$$

La classe des fonctions OWA comprend une sous-famille importante, à savoir : les statistiques d'ordre

$$\mathrm{OS}_k(x) = x_{(k)},$$

lorsque $\omega_k = 1$ pour un certain $k \in N$. Dans ce cas, on a, pour tout $S \subseteq N$,

$$\mu(S) = \begin{cases} 1, & \text{si } s \geq n-k+1, \\ 0, & \text{sinon.} \end{cases}$$

Cette sous-famille contient elle-même le minimum, le maximum, et la médiane.

Des axiomatiques de la classe des fonctions OWA se déduisent immédiatement de celles de l'intégrale de Choquet et de la Proposition 5.8.

5.3 L'intégrale de Sugeno

L'intégrale de Sugeno [102, 103] a été introduite comme intégrale floue, c'est-à-dire une intégrale définie à partir d'une mesure floue. Cette intégrale a ensuite fait l'objet d'une recherche très étendue et a été utilisée dans plusieurs domaines (un panorama peut être trouvé dans l'article [30] et le volume édité [52]).

Comme pour l'intégrale de Choquet, nous donnons ici la définition de la version discrète (finie) de l'intégrale de Sugeno qui n'est rien d'autre qu'une fonction d'agrégation de $[0,1]^n$ dans $[0,1]$.

Définition 5.3 *Soit $\mu \in \mathcal{F}_N$. L'intégrale de Sugeno (discrète) de $x \in [0,1]^n$ par rapport à μ est définie par*

$$\mathcal{S}_\mu(x) := \bigvee_{i=1}^{n} [x_{(i)} \wedge \mu(A_{(i)})],$$

où $(\cdot)$ indique une permutation sur N telle que $x_{(1)} \leq \ldots \leq x_{(n)}$. D'autre part, $A_{(i)} = \{(i), \ldots, (n)\}$, et $A_{(n+1)} = \varnothing$.

Comme dans la définition de l'intégrale de Choquet, le "coefficient" attaché à chaque variable x_i est fixé uniquement par la permutation $(\cdot)$. Par exemple, si $x_3 \leq x_1 \leq x_2$, nous avons

$$\mathcal{S}_\mu(x_1, x_2, x_3) = [x_3 \wedge \mu(\{3,1,2\})] \vee [x_1 \wedge \mu(\{1,2\})] \vee [x_2 \wedge \mu(\{2\})].$$

De la définition, nous pouvons immédiatement déduire que

$$\mathcal{S}_\mu(x) \in \{x_1, \ldots, x_n\} \cup \{\mu(S) \mid S \subseteq N\} \qquad (x \in [0,1]^n).$$

De plus, comme pour l'intégrale de Choquet, nous avons

$$\mathcal{S}_\mu(\mathbf{1}_S) = \mu(S) \qquad (S \subseteq N),$$

ce qui montre que l'intégrale de Sugeno est complètement déterminée par ses valeurs sur les sommets de l'hypercube $[0,1]^n$.

Il a aussi été démontré [54, 64, 102] que l'intégrale de Sugeno peut aussi être mis sous la forme suivante, qui ne nécessite pas le rangement des variables :

$$\mathcal{S}_\mu(x) = \bigvee_{T \subseteq N} \Big[\mu(T) \wedge (\bigwedge_{i \in T} x_i)\Big] \qquad (x \in [0,1]^n).$$

Il a aussi été démontré [57] que l'intégrale de Sugeno est une sorte de médiane pondérée :

$$\mathcal{S}_\mu(x) = \text{médiane}[x_1, \ldots, x_n, \mu(A_{(2)}), \mu(A_{(3)}), \ldots, \mu(A_{(n)})] \qquad (x \in [0,1]^n).$$

Par exemple, si $x_3 \leq x_1 \leq x_2$ alors

$$\mathcal{S}_\mu(x_1, x_2, x_3) = \text{médiane}[x_1, x_2, x_3, \mu(1,2), \mu(2)].$$

Le résultat suivant [66] montre que l'intégrale de Sugeno est un concept assez naturel et, contrairement à l'intégrale de Choquet, est appropriée pour une agrégation dans un contexte ordinal.

Proposition 5.9 *Toute fonction $A : [0,1]^n \to [0,1]$ faiblement idempotente et construite à partir de variables $x_1, \ldots, x_n \in [0,1]$, de constantes $r_1, \ldots, r_m \in [0,1]$, des opérations $\wedge = \min$ et $\vee = \max$, et des parenthèses est une intégrale de Sugeno (et inversement).*

Passons à présent aux axiomatiques de l'intégrale de Sugeno. On peut facilement voir que l'intégrale de Sugeno est une fonction continue, non décroissante, unanimement croissante, indempotente et interne. Elle vérifie aussi les propriétés de *minitivité comonotone* et *maxitivité comonotone* [25], c'est-à-dire

$$\begin{aligned} f(x_1 \wedge x_1', \ldots, x_n \wedge x_n') &= f(x_1, \ldots, x_n) \wedge f(x_1', \ldots, x_n') \\ f(x_1 \vee x_1', \ldots, x_n \vee x_n') &= f(x_1, \ldots, x_n) \vee f(x_1', \ldots, x_n') \end{aligned}$$

pour tous vecteurs comonotones $x, x' \in [0,1]^n$. Plus particulièrement, elle est *faiblement minitive* et *faiblement maxitive*, c'est-à-dire qu'elle vérifie

$$\begin{aligned} f(x_1 \wedge r, \ldots, x_n \wedge r) &= f(x_1, \ldots, x_n) \wedge r \\ f(x_1 \vee r, \ldots, x_n \vee r) &= f(x_1, \ldots, x_n) \vee r \end{aligned}$$

pour tout vecteur $x \in [0,1]^n$ et tout $r \in [0,1]$. Plus particulièrement encore, en remplaçant x par le vecteur booléen $\mathbf{1}_S$ dans les deux dernières équations, on voit qu'elle est aussi *non compensatoire*, c'est-à-dire qu'elle vérifie

$$f(r\mathbf{1}_S) \in \{f(\mathbf{1}_S), r\} \quad \text{et} \quad f(\mathbf{1}_S + r\mathbf{1}_{N\setminus S}) \in \{f(\mathbf{1}_S), r\}$$

pour tout $S \subseteq N$ et tout $r \in [0,1]$.

La minitivité comonotone et la maxitivité comonotone ont été justifiées dans le contexte de l'agrégation de sous-ensembles flous par Ralescu et Ralescu [88]. La non compensation a été justifiée en aide à la décision face à l'incertain dans [30].

Les principales axiomatiques de l'intégrale de Sugeno en tant que fonction d'agrégation sont résumées dans le résultat qui suit ; voir [62, 64] :

Théorème 5.4 *Soit* $A : [0,1]^n \to [0,1]$. *Les assertions suivantes sont équivalentes :*
- *A est non décroissant, idempotent et non compensatoire,*
- *A est non décroissant, faiblement minitif et faiblement maxitif,*
- *A est non décroissant, idempotent, minitif comonotone et maxitif comonotone,*
- *il existe $\mu \in \mathcal{F}_N$ tel que $A = \mathcal{S}_\mu$.*

L'intégrale de Sugeno à deux variables peut être caractérisée d'une façon très naturelle au moyen de la propriété d'associativité. En effet, le Théorème 4.7 peut se réécrire comme suit :

Proposition 5.10 $A : [0,1]^2 \to [0,1]$ *est continu, non décroissant, idempotent et associatif si et seulement s'il existe $\mu \in \mathcal{F}_2$ tel que $A = \mathcal{S}_\mu$.*

En considérant des suites associatives ou décomposables, nous avons ce qui suit ; voir [62, p. 113] :

Proposition 5.11 *Soit* $A := (A^{(n)} : [0,1]^n \to [0.1])_{n \geq 1}$ *une suite de fonctions. Alors les assertions suivantes sont équivalentes :*
- *A est une suite associative d'intégrales de Sugeno,*
- *A est une suite décomposable d'intégrales de Sugeno,*

– *A est une suite associative de fonctions continues, non décroissantes et idempotentes,*
– *il existe $\alpha, \beta \in [0,1]$ tels que*

$$A^{(n)}(x) = (\alpha\wedge x_1)\vee\Big(\bigvee_{i=2}^{n-1}(\alpha\wedge\beta\wedge x_i)\Big)\vee(\beta\wedge x_n)\vee\Big(\bigwedge_{i=1}^{n} x_i\Big) \qquad (x \in [0,1]^n,\ n \in \mathbb{N}\setminus\{0\}).$$

De même que l'intégrale de Choquet inclut deux grandes sous-classes, à savoir la moyenne arithmétique pondérée et la fonction moyenne ordonnée, l'intégrale de Sugeno inclut, entre autres choses, les minimum et maximum pondérés et les minimum et maximum ordonnés pondérés. Ces fonctions ont été introduites et étudiées respectivement dans [33] et [34].

Pour tout vecteur $\omega = (\omega_1, \ldots, \omega_n) \in [0,1]^n$ tel que $\bigvee_{i=1}^n \omega_i = 1$, le *maximum pondéré* associé à ω est défini par

$$\mathrm{pmax}_\omega(x) = \bigvee_{i=1}^{n}(\omega_i \wedge x_i) \qquad (x \in [0,1]^n).$$

Pour tout vecteur $\omega = (\omega_1, \ldots, \omega_n) \in [0,1]^n$ tel que $\bigwedge_{i=1}^n \omega_i = 0$, le *minimum pondéré* associé à ω est défini par

$$\mathrm{pmin}_\omega(x) = \bigwedge_{i=1}^{n}(\omega_i \vee x_i) \qquad (x \in [0,1]^n).$$

Les fonctions pmax_ω et pmin_ω peuvent être caracterisées comme suit ; voir [33, 62, 89] :

Proposition 5.12 *Soit $\mu \in \mathcal{F}_N$. Les assertions suivantes sont équivalentes :*
– *μ est une mesure de possibilité, c'est-à-dire telle que*

$$\mu(S \cup T) = \mu(S) \vee \mu(T) \qquad (S, T \subseteq N),$$

– *il existe $\omega \in [0,1]^n$ tel que $\mathcal{S}_\mu = \mathrm{pmax}_\omega$,*
– *on a*

$$\mathcal{S}_\mu(x_1 \vee x_1', \ldots, x_n \vee x_n') = \mathcal{S}_\mu(x_1, \ldots, x_n) \vee \mathcal{S}_\mu(x_1', \ldots, x_n') \qquad (x, x' \in [0,1]^n).$$

Proposition 5.13 *Soit $\mu \in \mathcal{F}_N$. Les assertions suivantes sont équivalentes :*
– *μ est une mesure de nécessité, c'est-à-dire telle que*

$$\mu(S \cap T) = \mu(S) \wedge \mu(T) \qquad (S, T \subseteq N),$$

– *il existe $\omega \in [0,1]^n$ tel que $\mathcal{S}_\mu = \mathrm{pmin}_\omega$,*
– *on a*

$$\mathcal{S}_\mu(x_1 \wedge x_1', \ldots, x_n \wedge x_n') = \mathcal{S}_\mu(x_1, \ldots, x_n) \wedge \mathcal{S}_\mu(x_1', \ldots, x_n') \qquad (x, x' \in [0,1]^n).$$

Pour tout vecteur $\omega = (\omega_1, \ldots, \omega_n) \in [0,1]^n$ tel que $\bigvee_{i=1}^n \omega_i = 1$, le *maximum ordonné pondéré* associé à ω est défini par

$$\mathrm{opmax}_\omega(x) = \bigvee_{i=1}^n (\omega_i \wedge x_{(i)}) \qquad (x \in [0,1]^n).$$

Pour tout vecteur $\omega = (\omega_1, \ldots, \omega_n) \in [0,1]^n$ tel que $\bigwedge_{i=1}^n \omega_i = 0$, le *minimum pondéré* associé à ω est défini par

$$\mathrm{opmin}_\omega(x) = \bigwedge_{i=1}^n (\omega_i \vee x_{(i)}) \qquad (x \in [0,1]^n).$$

Assez curieusement, la classe des minima ordonnés pondérés coïncide avec celle des maxima ordonnés pondérés et s'identifie aux intégrales de Sugeno symétriques. Le résultat est le suivant ; voir [49, 62] :

Proposition 5.14 *Soit $\mu \in \mathcal{F}_N$. Les assertions suivantes sont équivalentes :*
- *μ depend uniquement de la cardinalité des sous-ensembles,*
- *il existe $\omega \in [0,1]^n$ tel que $\mathcal{S}_\mu = \mathrm{opmax}_\omega$,*
- *il existe $\omega \in [0,1]^n$ tel que $\mathcal{S}_\mu = \mathrm{opmin}_\omega$,*
- *$\mathcal{S}_\mu$ est une fonction symétrique.*

En se servant du fait que l'intégrale de Sugeno est aussi une médiane pondérée, on peut écrire

$$\begin{aligned} \mathrm{opmax}_\omega(x) &= \mathrm{médiane}(x_1, \ldots, x_n, \omega_2, \ldots, \omega_n), \\ \mathrm{opmin}_\omega(x) &= \mathrm{médiane}(x_1, \ldots, x_n, \omega_1, \ldots, \omega_{n-1}). \end{aligned}$$

Une dernière sous-classe intéressante est celle des polynômes latticiels, qui ne sont rien d'autre que des intégrales de Sugeno définies à partir de mesures floues prenant leurs valeurs dans $\{0,1\}$. Nous caractériserons ces fonctions dans la dernière section.

6 Agrégation sur des échelles de ratio et d'intervalles

Dans cette section, nous présentons les familles de fonctions d'agrégation qui sont signifiantes pour les échelles de ratio et les échelles d'intervalle (voir Définition 2.7).

Tout d'abord, concernant les échelles de ratio, on a les deux résultats suivants ; voir [6, Chapitre 20], [7, p. 439], et [9, case#2] :

Théorème 6.1 $A :]0,\infty[^n \to]0,\infty[$ *est signifiant pour les mêmes échelles de ratio entrées-sorties si et seulement si*

$$A(x) = x_1\, F\Big(\frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1}\Big) \qquad (x \in]0,\infty[^n),$$

avec $F :]0,\infty[^{n-1} \to]0,\infty[$ arbitraire (F = constant si $n = 1$).

Théorème 6.2 *$A :]0,\infty[^n \to]0,\infty[$ est signifiant pour les mêmes échelles de ratio entrées si et seulement si*

$$A(x) = g(x_1)\, F\Big(\frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1}\Big) \qquad (x \in]0,\infty[^n),$$

avec $F :]0,\infty[^{n-1} \to]0,\infty[$ arbitraire (F = constant si $n = 1$) et $g :]0,\infty[\to]0,\infty[$ tel que $g(xy) = g(x)g(y)$ pour tous $x, y \in]0,\infty[$. $g(x) = x^c$ si A is continu (c arbitraire).

Concernant les échelles d'intervalle, on a les résultats suivants ; voir [9, case#5] et [62, §3.4.1] :

Théorème 6.3 *$A : \mathbb{R}^n \to \mathbb{R}$ est signifiant pour les mêmes échelles d'intervalle entrées-sorties si et seulement si*

$$A(x) = \begin{cases} \mathrm{S}(x)\, F\Big(\frac{x_1 - \mathrm{AM}(x)}{\mathrm{S}(x)}, \ldots, \frac{x_n - \mathrm{AM}(x)}{\mathrm{S}(x)}\Big) + \mathrm{AM}(x), & \text{si } \mathrm{S}(x) \neq 0, \\ x_1, & \text{si } \mathrm{S}(x) = 0, \end{cases}$$

où $\mathrm{S}(x) = \sqrt{\sum_{i=1}^n (x_i - \mathrm{AM}(x))^2}$ et $F : \mathbb{R}^n \to \mathbb{R}$ arbitraire ($A(x) = x$ si $n = 1$).

Théorème 6.4 *$A : \mathbb{R}^n \to \mathbb{R}$ est signifiant pour les mêmes échelles d'intervalle entrées si et seulement si*

$$A(x) = \begin{cases} \mathrm{S}(x)\, F\Big(\frac{x_1 - \mathrm{AM}(x)}{\mathrm{S}(x)}, \ldots, \frac{x_n - \mathrm{AM}(x)}{\mathrm{S}(x)}\Big) + a\,\mathrm{AM}(x) + b, & \text{si } \mathrm{S}(x) \neq 0, \\ a\, x_1 + b, & \text{si } \mathrm{S}(x) = 0, \end{cases}$$

ou

$$A(x) = \begin{cases} g(\mathrm{S}(x))\, F\Big(\frac{x_1 - \mathrm{AM}(x)}{\mathrm{S}(x)}, \ldots, \frac{x_n - \mathrm{AM}(x)}{\mathrm{S}(x)}\Big) + b, & \text{si } \mathrm{S}(x) \neq 0, \\ b, & \text{si } \mathrm{S}(x) = 0, \end{cases}$$

où $a, b \in \mathbb{R}$, $\mathrm{S}(x) = \sqrt{\sum_{i=1}^n (x_i - \mathrm{AM}(x))^2}$, $F : \mathbb{R}^n \to \mathbb{R}$ arbitraire ($A(x) = ax + b$ si $n = 1$), et $g : \mathbb{R} \to]0,\infty[$ tel que $g(xy) = g(x)g(y)$ pour tous $x, y \in \mathbb{R}$.

La restriction de ces familles aux fonctions non décroissantes et aux fonctions strictement croissantes a été discutée dans Aczél et col. [7].

Dans le reste de cette section, nous présentons des axiomatisations de quelques sous-familles de fonctions qui sont signifiantes pour les mêmes échelles d'intervalle entrées-sorties (ces résultats sont extraits de l'article [69]). Par exemple, nous avons déjà remarqué à la sous-section 5.2 que l'intégrale de Choquet discrète vérifie cette propriété. Plus généralement, il est évident que toute fonction d'agrégation obtenue en composant un nombre arbitraire d'intégrales de Choquet discrètes est encore signifiante pour les mêmes échelles d'intervalle entrées-sorties. Ces fonctions, appelées *intégrales de Choquet composées* ont partiellement été étudiées et font encore aujourd'hui l'objet d'une importante recherche ; voir par exemple [84].

Si l'on se restreint aux fonctions bisymétriques, nous avons les résultats suivants :

Proposition 6.1 *$A : \mathbb{R}^n \to \mathbb{R}$ est non décroissant, signifiant pour les mêmes échelles d'intervalle entrées-sorties et bisymétrique si et seulement si*

$$A \in \{\min_S, \max_S \mid S \subseteq N\} \cup \{\mathrm{WAM}_\omega \mid \omega \in [0,1]^n\}.$$

Corollaire 6.1 *$A : \mathbb{R}^n \to \mathbb{R}$ est symétrique, non décroissant, signifiant pour les mêmes échelles d'intervalle entrées-sorties et bisymétrique si et seulement si*

$$A \in \{\min, \max, \mathrm{AM}\}.$$

Proposition 6.2 *$(A^{(n)} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ est une suite bisymétrique de fonctions non décroissantes et signifiantes pour les mêmes échelles d'intervalle entrées-sorties si et seulement si*

- *soit, pour tout $n \in \mathbb{N} \setminus \{0\}$, il existe $S \subseteq N_n$ tel que $M^{(n)} = \min_S$,*
- *ou, pour tout $n \in \mathbb{N} \setminus \{0\}$, il existe $S \subseteq N_n$ tel que $M^{(n)} = \max_S$,*
- *ou, pour tout $n \in \mathbb{N} \setminus \{0\}$, il existe $\omega \in [0,1]^n$ tel que $M^{(n)} = \mathrm{WAM}_\omega$.*

Corollaire 6.2 *$A := (A^{(n)} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ est une suite bisymétrique de fonctions symétriques, non décroissantes et signifiantes pour les mêmes échelles d'intervalle entrées-sorties si et seulement si*

$$A = (\min{}^{(n)})_{n\geq 1} \text{ ou } (\max{}^{(n)})_{n\geq 1} \text{ ou } (\mathrm{AM}^{(n)})_{n\geq 1}.$$

Passons à présent aux suites décomposables et associatives de fonctions d'agrégation. Nous avons les résultats suivants :

Proposition 6.3 *$A := (A^{(n)} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ est une suite décomposable de fonctions non décroissantes et signifiantes pour les mêmes échelles d'intervalle entrées-sorties si et seulement si*

- *soit $A = (\min^{(n)})_{n\geq 1}$,*
- *ou $A = (\max^{(n)})_{n\geq 1}$,*
- *ou il existe $\theta \in [0,1]$ tel que, pour tout $n \in \mathbb{N} \setminus \{0\}$, on a $A^{(n)} = \mathrm{WAM}_\omega$, avec*

$$\omega_i = \frac{(1-\theta)^{n-i}\theta^{i-1}}{\sum_{j=1}^n (1-\theta)^{n-j}\theta^{j-1}} \qquad (i \in N_n).$$

Corollaire 6.3 *$A := (A^{(n)} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ est une suite décomposable de fonctions symétriques, non décroissantes et signifiantes pour les mêmes échelles d'intervalle entrées-sorties si et seulement si*

$$A = (\min{}^{(n)})_{n\geq 1} \text{ ou } (\max{}^{(n)})_{n\geq 1} \text{ ou } (\mathrm{AM}^{(n)})_{n\geq 1}.$$

Proposition 6.4 *$A := (A^{(n)} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ est une suite associative de fonctions non décroissantes et signifiantes pour les mêmes échelles d'intervalle entrées-sorties si et seulement si*

$$A = (\min{}^{(n)})_{n\geq 1} \text{ ou } (\max{}^{(n)})_{n\geq 1} \text{ ou } (\mathrm{P}_1^{(n)})_{n\geq 1} \text{ ou } (\mathrm{P}_n^{(n)})_{n\geq 1}.$$

Corollaire 6.4 *$A := (A^{(n)} : \mathbb{R}^n \to \mathbb{R})_{n\geq 1}$ est une suite associative de fonctions symétriques, non décroissantes et signifiantes pour les mêmes échelles d'intervalle entrées-sorties si et seulement si*

$$A = (\min{}^{(n)})_{n\geq 1} \text{ ou } (\max{}^{(n)})_{n\geq 1}.$$

7 Agrégation sur des échelles ordinales

Nous terminons notre panorama des fonctions d'agrégation par les fonctions signifiantes pour les mêmes échelles ordinales entrées-sorties.

La description de ces fonctions n'est pas immédiate et nécessite le concept d'ensemble invariant. Notons Φ l'ensemble des bijections strictement croissantes de $\mathbb{R}$.

Définition 7.1 *Un sous-ensemble non vide $I \subseteq \mathbb{R}^n$ is dit* invariant *si*

$$x \in I \quad \Rightarrow \quad \phi(x) \in I \qquad (\phi \in \Phi).$$

Un tel ensemble est dit minimal *s'il ne contient aucun sous-ensemble invariant propre.*

La famille $\mathcal{I}$ de tous les sous-ensembles invariants de $\mathbb{R}^n$ fournit une partition de $\mathbb{R}^n$ en classes d'équivalence, où $x, y \in \mathbb{R}^n$ sont équivalents s'il existe $\phi \in \Phi$ tel que $y = \phi(x)$. En fait, on peut montrer que tout sous-ensemble invariant est de la forme

$$I = \{x \in \mathbb{R}^n \mid x_{\pi(1)} \lhd_1 \cdots \lhd_{n-1} x_{\pi(n)}\},$$

où $\pi \in \Pi_N$ et $\lhd_i \in \{<, \leq\}$ pour $i = 1, \ldots, n-1$.

Les fonctions signifiantes pour les mêmes échelles ordinales entrées-sorties ont été étudiées par plusieurs auteurs [68, 71, 76, 87]. La description de ces fonctions est la suivante [87] :

Théorème 7.1 *$A : \mathbb{R}^n \to \mathbb{R}$ est signifiant pour les mêmes échelles ordinales entrées-sorties si et seulement si, pour tout $I \in \mathcal{I}$, il existe $i \in N$ tel que $A|_I = P_i|_I$ est la $i$ème projection.*

Les fonctions signifiantes pour les mêmes échelles ordinales entrées ont également été largement étudiées [68, 70, 85, 86, 109]. La description de ces fonctions est la suivante [70], où "Im" signifie "Image" :

Théorème 7.2 *$A : \mathbb{R}^n \to \mathbb{R}$ est signifiant pour les mêmes échelles ordinales entrées si et seulement si, pour tout $I \in \mathcal{I}$, il existe $i_I \in N$ et une fonction constante ou strictement monotone $g_I : P_{i_I}(I) \to \mathbb{R}$ tels que*

$$A|_I = g_I \circ P_{i_I}|_I,$$

où, pour tous $I, J \in \mathcal{I}$, soit $g_I = g_J$, ou $\mathrm{Im}(g_I) = \mathrm{Im}(g_J)$ est un singleton, ou $\mathrm{Im}(g_I) < \mathrm{Im}(g_J)$, ou encore $\mathrm{Im}(g_I) > \mathrm{Im}(g_J)$.

Ainsi, nous voyons que les fonctions signifiantes pour les mêmes échelles ordinales entrées-sorties se réduisent à des projections sur chaque sous-ensemble invariant et les fonctions signifiantes pour les mêmes échelles ordinales entrées se réduisent à des constantes ou des projections transformées sur ces mêmes sous-ensembles invariants.

La restriction de ces fonctions aux fonctions non décroissantes et/ou continues a aussi été étudiée. Pour décrire ces sous-familles, nous avons besoin du concept de polynôme latticiel.

Définition 7.2 *Un polynôme latticiel à n variables est une expression impliquant n variables $x_1, \ldots, x_n$ liées par les opérations latticielles $\wedge = \min$ et $\vee = \max$ dans une combinaison arbitraire de parenthèses.*

Par exemple, $L(x) = (x_1 \vee x_2) \wedge x_3$ est un polynôme latticiel à trois variables.

On peut montrer (voir [17, Chapitre 2, §5]) que tout polynôme latticiel de n variables peut s'écrire sous forme disjonctive comme

$$L_\gamma(x) = \bigvee_{\substack{S \subseteq N \\ \gamma(S)=1}} \bigwedge_{i \in S} x_i \qquad (x \in \mathbb{R}^n),$$

où $\gamma : 2^N \to \{0,1\}$ est une mesure floue binaire (c'est-à-dire à valeurs 0 ou 1). Nous noterons Γ_N la famille de ces mesures floues.

Il a aussi été démontré [66] que la classe des polynômes latticiels restreints au domaine $[0,1]^n$ s'identifie à l'intersection entre la famille des intégrales de Choquet sur $[0,1]^n$ et la famille des intégrales de Sugeno.

Concernant les fonctions non décroissantes, nous avons les descriptions suivantes [68, 70] :

Proposition 7.1 *$A : \mathbb{R}^n \to \mathbb{R}$ est non décroissant et signifiant pour les mêmes échelles ordinales entrées-sorties si et seulement s'il existe $\gamma \in \Gamma_N$ tel que $A = L_\gamma$.*

Proposition 7.2 *$A : \mathbb{R}^n \to \mathbb{R}$ est non décroissant et signifiant pour les mêmes échelles ordinales entrées si et seulement s'il existe $\gamma \in \Gamma_N$ et une fonction $g : \mathbb{R} \to \mathbb{R}$ constante ou strictement croissante tels que $A = g \circ L_\gamma$.*

Il apparaît que les fonctions des deux théorèmes précédents sont continues, à des discontinuités près de la fonction g.

Pour les fonctions continues, nous avons ce qui suit [68] :

Corollaire 7.1 *$A : \mathbb{R}^n \to \mathbb{R}$ est continu et signifiant pour les mêmes échelles ordinales entrées-sorties si et seulement s'il existe $\gamma \in \Gamma_N$ tel que $A = L_\gamma$.*

Corollaire 7.2 *$A : \mathbb{R}^n \to \mathbb{R}$ est continu et signifiant pour les mêmes échelles ordinales entrées si et seulement s'il existe $\gamma \in \Gamma_N$ et une fonction $g : \mathbb{R} \to \mathbb{R}$ constante ou strictement monotone et continue tels que $A = g \circ L_\gamma$.*

Les polynômes latticiels sont idempotents, mais pas nécessairement symétriques. En fait, les polynômes latticiels symétriques sont exactement les statistiques d'ordre, lesquels contiennent la médiane classique. En ajoutant la symétrie et/ou l'idempotence aux résultats précédents, nous avons les corollaires suivants :

Corollaire 7.3 *$A : \mathbb{R}^n \to \mathbb{R}$ est symétrique, non décroissant (ou continu) et signifiant pour les mêmes échelles ordinales entrées-sorties si et seulement s'il existe $k \in N$ tel que $A = \mathrm{OS}_k$.*

Corollaire 7.4 *$A : \mathbb{R}^n \to \mathbb{R}$ est idempotent, non décroissant (ou continu) et signifiant pour les mêmes échelles ordinales entrées si et seulement s'il existe $\gamma \in \Gamma_N$ tel que $A = L_\gamma$.*

Corollaire 7.5 *$A : \mathbb{R}^n \to \mathbb{R}$ est symétrique, non décroissant et signifiant pour les mêmes échelles ordinales entrées si et seulement s'il existe $k \in N$ et une fonction $g : \mathbb{R} \to \mathbb{R}$ constante ou strictement croissante tels que $A = g \circ \mathrm{OS}_k$.*

Corollaire 7.6 *$A : \mathbb{R}^n \to \mathbb{R}$ est symétrique, continu et signifiant pour les mêmes échelles ordinales entrées si et seulement s'il existe $k \in N$ et une fonction $g : \mathbb{R} \to \mathbb{R}$ constante ou strictement monotone et continue tels que $A = g \circ \mathrm{OS}_k$.*

8 Conclusion

Danc ce chapitre, nous avons passé en revue les fonctions d'agrégation les plus classiques qui sont utilisées en aide à la décision. Un classement convenable de ces fonctions en un catalogue ne peut se faire que via une approche axiomatique qui consiste à lister une série de propriétés raisonnables et à classer ou mieux, caractériser, les fonctions d'agrégation en fonction des ces propriétés.

Etant donné le besoin grandissant de définir des agrégateurs appropriés répondants à des critères très précis face à des situations de plus en plus variées, il n'est pas surprenant qu'un tel catalogue de fonctions d'agrégation, qui est déjà très volumineux, ne cesse de se remplir et fait aujourd'hui l'objet d'une importante recherche.

Nous avons ici simplement écrémé la surface d'un domaine qui est en pleine expansion et qui est de plus en plus représenté dans de nombreuses conférences internationales comme IFSA, IEEE, IPMU, EUSFLAT, EUROFUSE, FSTA, AGOP, etc.

Références

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series.* For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[2] J. Aczél. On mean values. *Bull. Amer. Math. Soc.*, 54 :392–400, 1948.

[3] J. Aczél. Sur les opérations définies pour nombres réels. *Bull. Soc. Math. France*, 76 :59–64, 1948.

[4] J. Aczél. *Lectures on functional equations and their applications.* Academic Press, New York, 1966.

[5] J. Aczél and C. Alsina. Synthesizing judgements : a functional equations approach. *Math. Modelling*, 9(3-5) :311–320, 1987. The analytic hierarchy process.

[6] J. Aczél and J. Dhombres. *Functional equations in several variables.* Cambridge University Press, Cambridge, 1989. With applications to mathematics, information theory and to the natural and social sciences.

[7] J. Aczél, D. Gronau, and J. Schwaiger. Increasing solutions of the homogeneity equation and of similar equations. *J. Math. Anal. Appl.*, 182(2) :436–464, 1994.

[8] J. Aczél and F. S. Roberts. On the possible merging functions. *Math. Social Sci.*, 17(3) :205–243, 1989.

[9] J. Aczél, F. S. Roberts, and Z. Rosenbaum. On scientific laws without dimensional constants. *J. Math. Anal. Appl.*, 119(1-2) :389–416, 1986.

[10] C. Alsina, G. Mayor, M. S. Tomás, and J. Torrens. A characterization of a class of aggregation functions. *Fuzzy Sets and Systems*, 53(1) :33–38, 1993.

[11] C. Alsina, E. Trillas, and L. Valverde. On some logical connectives for fuzzy sets theory. *J. Math. Anal. Appl.*, 93(1) :15–26, 1983.

[12] C. Antoine. *Les moyennes*, volume 3383 of *Que Sais-Je ? [What Do I Know ?]*. Presses Universitaires de France, Paris, 1998.

[13] E. F. Beckenbach and R. Bellman. *Inequalities.* Second revised printing. Ergebnisse der Mathematik und ihrer Grenzgebiete. Neue Folge, Band 30. Springer-Verlag, New York, Inc., 1965.

[14] G. Bemporad. Sul principio della media aritmetica. (Italian). *Atti Accad. Naz. Lincei*, 3(6) :87–91, 1926.

[15] L. R. Berrone and J. Moro. Lagrangian means. *Aequationes Math.*, 55(3) :217–226, 1998.

[16] L. R. Berrone and J. Moro. On means generated through the Cauchy mean value theorem. *Aequationes Math.*, 60(1-2) :1–14, 2000.

[17] G. Birkhoff. *Lattice theory.* Third edition. American Mathematical Society Colloquium Publications, Vol. XXV. American Mathematical Society, Providence, R.I., 1967.

[18] P. S. Bullen, D. S. Mitrinović, and P. M. Vasić. *Means and their inequalities*, volume 31 of *Mathematics and its Applications (East European Series).* D. Reidel Publishing Co., Dordrecht, 1988. Translated and revised from the Serbo-Croatian.

[19] A. L. Cauchy. *Cours d'analyse de l'Ecole Royale Polytechnique, Vol. I. Analyse algébrique.* Debure, Paris, 1821.

[20] O. Chisini. Sul concetto di media. (Italian). *Periodico di matematiche*, 9(4) :106–116, 1929.

[21] G. Choquet. Theory of capacities. *Ann. Inst. Fourier, Grenoble*, 5 :131–295 (1955), 1953–1954.

[22] A. C. Climescu. Sur l'équation fonctionnelle de l'associativité. *Bull. École Polytech. Jassy [Bul. Politehn. Gh. Asachi. Iaşi]*, 1 :211–224, 1946.

[23] E. Czogała and J. Drewniak. Associative monotonic operations in fuzzy set theory. *Fuzzy Sets and Systems*, 12(3) :249–269, 1984.

[24] L. M. de Campos and M. Jorge. Characterization and comparison of sugeno and choquet integrals. *Fuzzy Sets and Systems*, 52(1) :61–67, 1992.

[25] L. M. de Campos, M. T. Lamata, and S. Moral. A unified approach to define fuzzy integrals. *Fuzzy Sets and Systems*, 39(1) :75–90, 1991.

[26] B. de Finetti. Sul concetto di media. (Italian). *Giorn. Ist. Ital. Attuari*, 2(3) :369–396, 1931.

[27] C. Dellacherie. Quelques commentaires sur les prolongements de capacités. In *Séminaire de Probabilités, V (Univ. Strasbourg, année universitaire 1969-1970)*, pages 77–81. Lecture Notes in Math., Vol. 191. Springer, Berlin, 1971.

[28] J. Dow and S. R. d. C. Werlang. Uncertainty aversion, risk aversion, and the optimal choice of portfolio. *Econometrica*, 60(1) :197–204, 1992.

[29] J. Dow and S. R. d. C. Werlang. Nash equilibrium under Knightian uncertainty : breaking down backward induction. *J. Econom. Theory*, 64(2) :305–324, 1994.

[30] D. Dubois, J. L. Marichal, H. Prade, M. Roubens, and R. Sabbadin. The use of the discrete Sugeno integral in decision-making : a survey. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 9(5) :539–561, 2001.

[31] D. Dubois and H. Prade. Criteria aggregation and ranking of alternatives in the framework of fuzzy set theory. In *Fuzzy sets and decision analysis*, volume 20 of *Stud. Management Sci.*, pages 209–240. North-Holland, Amsterdam, 1984.

[32] D. Dubois and H. Prade. A review of fuzzy set aggregation connectives. *Inform. Sci.*, 36(1-2) :85–121, 1985.

[33] D. Dubois and H. Prade. Weighted minimum and maximum operations in fuzzy set theory. *Inform. Sci.*, 39(2) :205–210, 1986.

[34] D. Dubois, H. Prade, and C. Testemale. Weighted fuzzy pattern matching. *Fuzzy Sets and Systems*, 28(3) :313–331, 1988. Mathematical modelling.

[35] J. J. Dujmović. Weighted conjunctive and disjunctive means and their application in system evaluation. *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz.*, (461-497) :147–158, 1974.

[36] J. J. Dujmović. Extended continuous logic and the theory of complex criteria. *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz.*, (498–541) :197–216, 1975.

[37] H. Dyckhoff and W. Pedrycz. Generalized means as model of compensative connectives. *Fuzzy Sets and Systems*, 14(2) :143–154, 1984.

[38] W. M. Faucett. Compact semigroups irreducibly connected between two idempotents. *Proc. Amer. Math. Soc.*, 6 :741–747, 1955.

[39] P. Fishburn and P. Wakker. The invention of the independence condition for preferences. *Management Sciences*, 41(7) :1130–1144, 1995.

[40] J. C. Fodor. A remark on constructing t-norms. *Fuzzy Sets and Systems*, 41(2) :195–199, 1991.

[41] J. C. Fodor. Strict preference relations based on weak t-norms. *Fuzzy Sets and Systems*, 43(3) :327–336, 1991. Aggregation and best choices of imprecise opinions (Brussels, 1989).

[42] J. C. Fodor. An extension of Fung-Fu's theorem. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 4(3) :235–243, 1996.

[43] J. C. Fodor and M. Roubens. *Fuzzy preference modelling and milticriteria decision support*. Kluwer, Dordrecht, 1994.

[44] V. Frosini. Averages. In *Italian Contributions to the Methodology of Statistics*, pages 1–17. Cleup, Padova, 1987.

[45] L. W. Fung and K. S. Fu. An axiomatic approach to rational decision making in a fuzzy environment. In *Fuzzy sets and their applications to cognitive and decision processes (Proc. U. S.-Japan Sem., Univ. Calif., Berkeley, Calif., 1974)*, pages 227–256. Academic Press, New York, 1975.

[46] I. Gilboa. Expected utility with purely subjective nonadditive probabilities. *J. Math. Econom.*, 16(1) :65–88, 1987.

[47] M. Grabisch. On the use of fuzzy integral as a fuzzy connective. In *Proc. of the Second IEEE International Conference on Fuzzy Systems*, pages 213–218, San Francisco, 1993.

[48] M. Grabisch. Fuzzy integral in multicriteria decision making. *Fuzzy Sets and Systems*, 69(3) :279–298, 1995.

[49] M. Grabisch. On equivalence classes of fuzzy connectives : the case of fuzzy integrals. *IEEE Trans. Fuzzy Syst.*, 3(1) :96–109, 1995.

[50] M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European J. Oper. Res.*, 89(3) :445–456, 1996.

[51] M. Grabisch. Fuzzy integral as a flexible and interpretable tool of aggregation. In *Aggregation and fusion of imperfect information*, volume 12 of *Stud. Fuzziness Soft Comput.*, pages 51–72. Physica, Heidelberg, 1998.

[52] M. Grabisch, T. Murofushi, and M. Sugeno, editors. *Fuzzy measures and integrals, Theory and applications*, volume 40 of *Studies in Fuzziness and Soft Computing*. Physica-Verlag, Heidelberg, 2000.

[53] M. Grabisch, H. T. Nguyen, and E. A. Walker. *Fundamentals of uncertainty calculi with applications to fuzzy inference*, volume 30 of *Theory and Decision Library. Series B : Mathematical and Statistical Methods*. Kluwer Academic Publishers, Dordrecht, 1995.

[54] G. H. Greco. On L-fuzzy integrals of measurable functions. *J. Math. Anal. Appl.*, 128(2) :581–585, 1987.

[55] E. Groes, J. Jacobsen, B. Sloth, and T. Tranæs. Axiomatic characterizations of the Choquet integral. *Econom. Theory*, 12(2) :441–448, 1998.

[56] U. Höhle. Integration with respect to fuzzy measures. In *Proc. IFAC Symposium on Theory and Applications of Digital Control*, pages 35–37, New Delhi, January 1982.

[57] A. Kandel and W. J. Byatt. Fuzzy sets, fuzzy algebra, and fuzzy statistics. *Proc. IEEE*, 66(12) :1619–1639, December 1978.

[58] E. P. Klement, R. Mesiar, and E. Pap. *Triangular norms*, volume 8 of *Trends in Logic—Studia Logica Library*. Kluwer Academic Publishers, Dordrecht, 2000.

[59] A. N. Kolmogoroff. Sur la notion de la moyenne. (French). *Atti Accad. Naz. Lincei*, 12(6) :388–391, 1930.

[60] C.-H. Ling. Representation of associative functions. *Publ. Math. Debrecen*, 12 :189–212, 1965.

[61] R. D. Luce. On the possible psychophysical laws. *Psych. Rev.*, 66 :81–95, 1959.

[62] J.-L. Marichal. *Aggregation operators for multicriteria decision aid.* PhD thesis, Institute of Mathematics, University of Liège, Liège, Belgium, December 1998.

[63] J.-L. Marichal. An axiomatic approach of the discrete choquet integral as a tool to aggregate interacting criteria. *IEEE Trans. Fuzzy Syst.*, 8(6) :800–807, 2000.

[64] J.-L. Marichal. On Sugeno integral as an aggregation function. *Fuzzy Sets and Systems*, 114(3) :347–365, 2000.

[65] J.-L. Marichal. On the associativity functional equation. *Fuzzy Sets and Systems*, 114(3) :381–389, 2000.

[66] J.-L. Marichal. An axiomatic approach of the discrete sugeno integral as a tool to aggregate interacting criteria in a qualitative framework. *IEEE Trans. Fuzzy Syst.*, 9(1) :164–172, 2001.

[67] J.-L. Marichal. Aggregation of interacting criteria by means of the discrete choquet integral. In *Aggregation operators : new trends and applications*, pages 224–244. Physica, Heidelberg, 2002.

[68] J.-L. Marichal. On order invariant synthesizing functions. *J. Math. Psych.*, 46(6) :661–676, 2002.

[69] J.-L. Marichal, P. Mathonet, and E. Tousset. Characterization of some aggregation functions stable for positive linear transformations. *Fuzzy Sets and Systems*, 102(2) :293–314, 1999.

[70] J.-L. Marichal, R. Mesiar, and T. Rückschlossová. A complete description of comparison meaningful functions. *Aequationes Math.*, 69(3) :309–320, 2005.

[71] J.-L. Marichal and M. Roubens. Characterization of some stable aggregation functions. In *Proc. 1st Int. Conf. on Industrial Engineering and Production Management (IEPM'93)*, pages 187–196, Mons, Belgium, June 1993.

[72] M. Mas, G. Mayor, and J. Torrens. t-operators. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 7(1) :31–50, 1999.

[73] M. Mas, G. Mayor, and J. Torrens. The modularity condition for uninorms and t-operators. *Fuzzy Sets and Systems*, 126(2) :207–218, 2002.

[74] M. Mas, M. Monserrat, and J. Torrens. On left and right uninorms. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 9(4) :491–507, 2001.

[75] J. Matkowski. Invariant and complementary quasi-arithmetic means. *Aequationes Math.*, 57(1) :87–107, 1999.

[76] R. Mesiar and T. Rückschlossová. Characterization of invariant aggregation operators. *Fuzzy Sets and Systems*, 142(1) :63–73, 2004.

[77] F. Modave, D. Dubois, M. Grabisch, and H. Prade. A choquet integral representation in multicriteria decision making. In *AAAI Fall Symposium on Frontiers in Soft Computing and Decisions Systems*, pages 30–39, Boston, MA, November 7–9 1997.

[78] F. Modave and M. Grabisch. Preference representation by the choquet integral : the commensurability hypothesis. In *Proc. 7th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'98)*, pages 164–171, Paris, 1998.

[79] P. S. Mostert and A. L. Shields. On the structure of semigroups on a compact manifold with boundary. *Ann. of Math. (2)*, 65 :117–143, 1957.

[80] 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(2) :201–227, 1989.

[81] T. Murofushi and M. Sugeno. A theory of fuzzy measures : representations, the Choquet integral, and null sets. *J. Math. Anal. Appl.*, 159(2) :532–549, 1991.

[82] T. Murofushi and M. Sugeno. Some quantities represented by the Choquet integral. *Fuzzy Sets Syst.*, 56(2) :229–235, 1993.

[83] M. Nagumo. Über eine klasse der mittelwerte. (German). *Japanese Journ. of Math.*, 7 :71–79, 1930.

[84] Y. Narukawa and T. Murofushi. The n-step Choquet integral on finite spaces. In *9th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2002)*, pages 539–543, Annecy, France, July 1-5 2002.

[85] A. I. Orlov. The connection between mean values and the admissible transformations of scale. *Math. Notes*, 30 :774–778, 1981.

[86] S. Ovchinnikov. Means on ordered sets. *Math. Social Sci.*, 32(1) :39–56, 1996.

[87] S. Ovchinnikov. Invariant functions on simple orders. *Order*, 14(4) :365–371, 1997/98.

[88] A. L. Ralescu and D. A. Ralescu. Extensions of fuzzy aggregation. *Fuzzy Sets and Systems*, 86(3) :321–330, 1997.

[89] D. A. Ralescu and M. Sugeno. Fuzzy integral representation. *Fuzzy Sets and Systems*, 84(2) :127–133, 1996.

[90] U. Ricci. Confronti tra medie. (Italian). *Giorn. Economisti e Rivista di Statistica*, 26 :38–66, 1915.

[91] F. S. Roberts. *Measurement theory*, volume 7 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1979. With applications to decisionmaking, utility and the social sciences, Advanced Book Program.

[92] F. S. Roberts. Merging relative scores. *J. Math. Anal. Appl.*, 147(1) :30–52, 1990.

[93] F. S. Roberts. Limitations on conclusions using scales of measurement. In *Operation Research and the Public Sector*, pages 621–671. Elsevier, Amsterdam, 1994.

[94] W. Sander. Associative aggregation operators. In *Aggregation operators*, volume 97 of *Stud. Fuzziness Soft Comput.*, pages 124–158. Physica, Heidelberg, 2002.

[95] R. Sarin and P. Wakker. A simple axiomatization of nonadditive expected utility. *Econometrica*, 60(6) :1255–1272, 1992.

[96] D. Schmeidler. Integral representation without additivity. *Proc. Amer. Math. Soc.*, 97(2) :255–261, 1986.

[97] B. Schweizer and A. Sklar. Associative functions and statistical triangle inequalities. *Publ. Math. Debrecen*, 8 :169–186, 1961.

[98] B. Schweizer and A. Sklar. Associative functions and abstract semigroups. *Publ. Math. Debrecen*, 10 :69–81, 1963.

[99] B. Schweizer and A. Sklar. *Probabilistic metric spaces*. North-Holland Series in Probability and Applied Mathematics. North-Holland Publishing Co., New York, 1983.

[100] S. S. Stevens. Mathematics, measurement, and psychophysics. In *Handbook of Experimental Psychology*, pages 1–49. Wiley, New York, 1951.

[101] S. S. Stevens. Measurement, psychophysics, and utility. In *Measurement : Definitions and Theories*, pages 18–63. Wiley, New York, 1959.

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

[103] M. Sugeno. Fuzzy measures and fuzzy integrals—a survey. In *Fuzzy automata and decision processes*, pages 89–102. North-Holland, New York, 1977.

[104] P. Wakker. Continuous subjective expected utility with nonadditive probabilities. *J. Math. Econom.*, 18(1) :1–27, 1989.

[105] S. Weber. A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms. *Fuzzy Sets and Systems*, 11(2) :115–134, 1983.

[106] R. Yager and J. Kacprzyk, editors. *The Ordered Weighted Averaging operators*. Kluwer Academic Publishers, USA, 1997. Theory and Applications.

[107] R. R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. Systems Man Cybernet.*, 18(1) :183–190, 1988.

[108] R. R. Yager and A. Rybalov. Uninorm aggregation operators. *Fuzzy Sets and Systems*, 80(1) :111–120, 1996.

[109] E. B. Yanovskaya. Group choice rules in problems with interpersonal preference comparisons. *Automat. Remote Control*, 50(6) :822–830, 1989.

[110] H.-J. Zimmermann and P. Zysno. Latent connectives in human decision making. *Fuzzy Sets and Systems*, 4 :37–51, 1980.