



HAL
open science

Sur l'apprentissage de capacités pour les intégrales de Sugeno avec des systèmes d'équations relationnelles floues

Ismaïl Baaj

► **To cite this version:**

Ismaïl Baaj. Sur l'apprentissage de capacités pour les intégrales de Sugeno avec des systèmes d'équations relationnelles floues. Rencontres Francophones sur la Logique Floue et ses Applications 2024, Nov 2024, Brest, France. pp.141-148. hal-04774946

HAL Id: hal-04774946

<https://hal.science/hal-04774946v1>

Submitted on 9 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sur l'apprentissage de capacités pour les intégrales de Sugeno avec des systèmes d'équations relationnelles floues

Ismaïl Baaj

Univ. Artois, CNRS, CRIL, F-62300 Lens, France

baaj@cril.fr

Résumé :

Dans cet article, nous introduisons une méthode d'apprentissage d'une capacité sous-jacente à une intégrale de Sugeno en fonction des données d'entraînement basées sur les systèmes d'équations relationnelles floues. Aux données d'entraînement, nous associons deux systèmes d'équations : un système de type $\max - \min$ et un système de type $\min - \max$. En résolvant ces deux systèmes (dans le cas où ils sont compatibles) à l'aide des résultats de Sanchez, nous montrons que nous pouvons obtenir directement les capacités extrémales représentant les données. En réduisant le système d'équations de type $\max - \min$ (resp. $\min - \max$) aux sous-ensembles de critères de cardinalité inférieure ou égale à q (resp. de cardinalité supérieure ou égale à $n - q$ où n est le nombre de critères), nous donnons une condition suffisante pour déduire, à partir de sa plus grande solution potentielle (resp. sa plus petite solution potentielle), une capacité q -maxitive (resp. q -minitive). Enfin, si ces deux systèmes réduits sont incompatibles, nous montrons comment obtenir la plus grande capacité q -maxitive approchée et la plus petite capacité q -minitive approchée, en utilisant des résultats récents pour traiter l'incompatibilité des systèmes d'équations relationnelles floues.

Mots-clés :

Intégrales de Sugeno, Systèmes d'équations relationnelles floues ; Apprentissage

Abstract:

In this article, we introduce a method for learning a capacity underlying a Sugeno integral according to training data based on systems of fuzzy relational equations. To the training data, we associate two systems of equations : a $\max - \min$ system and a $\min - \max$ system. By solving these two systems (in the case that they are consistent) using Sanchez's results, we show that we can directly obtain the extremal capacities representing the training data. By reducing the $\max - \min$ (resp. $\min - \max$) system of equations to subsets of criteria of cardinality less than or equal to q (resp. of cardinality greater than or equal to $n - q$), where n is the number of criteria, we give a sufficient condition for deducing, from its potential greatest solution (resp. its potential lowest solution), a q -maxitive (resp. q -minitive) capacity. Finally, if these two reduced systems of equations are inconsistent, we show how to obtain the greatest approximate q -maxitive capacity and the lowest approximate q -minitive capacity, using recent results to handle the inconsistency of systems of fuzzy relational equations.

Keywords:

Sugeno integrals, Systems of fuzzy relational equations ; Learning

1 Introduction

En théorie de la décision, l'intégrale de Sugeno [14] est couramment utilisée comme une fonction d'agrégation qualitative [8] à valeurs dans une échelle finie L . La définition de cette intégrale repose sur une fonction d'ensemble appelée capacité, qui modélise qualitativement l'importance (ou l'interaction) de sous-ensembles de critères, et est utilisée dans de nombreux domaines [9] tels que la modélisation de l'incertitude, l'agrégation multicritère ou en théorie des jeux.

Au cours des dernières années, de nombreuses approches, e.g. [1, 4, 6], ont été proposées pour apprendre (en des sens proches) la capacité d'une intégrale de Sugeno en fonction de N données d'entraînement, chaque donnée étant représentée par une paire (objet, évaluation globale ciblée).

Dans cet article, nous proposons une approche d'apprentissage d'une capacité sous-jacente à une intégrale de Sugeno en fonction de données en se basant sur des systèmes d'équations relationnelles floues. Notre but est le suivant : l'intégrale de Sugeno définie par une capacité apprise doit nous permettre d'obtenir pour chaque objet figurant dans les données d'entraînement, son évaluation globale ciblée, ou *une valeur approchée dans un sens précis*. Notre approche repose sur le fait que la définition de l'intégrale de Sugeno d'un objet définie par une capacité, est donnée par un produit $\max - \min$ entre une matrice ligne contenant le minimum des évaluations partielles de l'objet étudié relativement aux sous-ensembles de critères et un vecteur colonne dont les composantes sont

les mesures des sous-ensembles de critères définies par la capacité. L'intégrale de Sugeno s'obtient aussi par un produit min – max défini similairement. Ces deux expressions de l'intégrale de Sugeno nous permettent d'introduire deux systèmes d'équations relationnelles floues, un de type max – min et un second de type min – max pour notre problème d'apprentissage. La matrice de chaque système est construite à partir des objets des données d'entraînement, le second membre de chaque système est constitué par les évaluations globales ciblées. Sous certaines conditions naturelles, la résolution de chacun des deux systèmes, basée sur les travaux de Sanchez [13], produit un vecteur dont les composantes définissent les mesures des sous-ensembles des critères d'une capacité.

Les capacités q -maxitives/minitives ont été introduites [5, 11, 15] pour réduire la complexité du calcul de l'évaluation globale d'un objet par l'intégrale de Sugeno. Nous montrons que l'apprentissage de capacités q -maxitives (resp. q -minitives), peut être étudié par la réduction du système de type max – min (resp. de type min – max) aux sous-ensembles de critères de cardinalité inférieure ou égale à q (resp. de cardinalité supérieure ou égale à $n - q$ où n est le nombre critères). Pour établir une condition nécessaire et suffisante (Théorème 2, resp. 3) pour qu'une capacité q -maxitive (resp. q -minitive) représente les données d'entraînement, nous commençons par donner une construction d'une telle capacité (Proposition 1) (resp. (Proposition 2)). Dans le cas où le système réduit de type max – min (resp. de type min – max) est incompatible, sous une hypothèse naturelle portant sur les données, nous nous appuyons sur les résultats de [2, 3] pour calculer explicitement dans Théorème 4 (resp. Théorème 5), la plus grande (resp. la plus petite) capacité q -maxitive (resp. q -minitive) approchée en un sens précis (Corollaire 2) (resp. (Corollaire 3)).

L'article est structuré comme suit. La Section 2 est consacrée à quelques rappels sur les intégrales de Sugeno et les systèmes d'équa-

tions relationnelles floues de type max – min et min – max. Dans la Section 3, nous introduisons deux systèmes d'équations relationnelles floues qui permettent d'apprendre des capacités d'intégrales de Sugeno en fonction de données et nous donnons une construction de capacité q -maxitive (resp. q -minitive). Nous illustrons nos résultats sur un exemple. Dans la Section 4, nous énonçons nos principaux résultats sur l'apprentissage de capacités q -maxitives/minitives.

2 Rappels

Dans ce qui suit, nous rappelons la définition de l'intégrale de Sugeno et les résultats classiques portant sur la résolution des systèmes d'équations relationnelles floues de type max – min et min – max.

2.1 Intégrales de Sugeno

Le cadre de l'intégrale de Sugeno (voir [7] pour plus de détails) comprend :

- $\mathcal{C} = \{1, 2, \dots, n\}$: un ensemble de n critères,
- $L = \{\xi_1 = 0, \xi_2, \dots, \xi_l = 1\}$: une échelle d'évaluation (un ensemble fini totalement ordonné $\xi_1 = 0 < \xi_2 < \dots < \xi_l = 1$), ou $L = [0, 1]$.
- $\mu : 2^{\mathcal{C}} \rightarrow L$ une capacité i.e., une fonction d'ensembles telle que $\mu(\emptyset) = 0, \mu(\mathcal{C}) = 1$ et $A \subseteq B \implies \mu(A) \leq \mu(B)$. La capacité conjuguée μ^c est définie par $\mu^c(A) = \iota(\mu(\bar{A}))$ avec $\iota : L \rightarrow L : \xi_i \mapsto \xi_{l-i+1}$, ou $\iota : [0, 1] \rightarrow [0, 1] : t \mapsto 1 - t$.
- $x = [x_i] \in L^{n \times 1}$: un objet où chaque x_i est une évaluation partielle de l'objet x en fonction du critère i .

L'intégrale de Sugeno est donnée par deux formules équivalentes [10, 7] (d'autres formules équivalentes existent) :

$$S_\mu(x) = \max_{A \in 2^{\mathcal{C}}} \min_{i \in A} (\min x_i, \mu(A)) \quad (1)$$

$$= \min_{A \in 2^{\mathcal{C}}} \max_{i \in \bar{A}} (\max x_i, \mu(A)). \quad (2)$$

où \bar{A} désigne le complémentaire du sous-ensemble A de \mathcal{C} . Les deux notions suivantes ont été introduites dans [5, 11, 15] :

Definition 1 Soient $\mu : 2^{\mathcal{C}} \rightarrow L$ une capacité et $q \in \{1, 2, \dots, n\}$ où $n = |\mathcal{C}|$.

1. μ est dite q -maxitive ssi pour tout sous-ensemble $X \subseteq \mathcal{C}$ vérifiant $|X| > q$, nous avons $\mu(X) = \max_{Y \subset X, |Y| \leq q} \mu(Y)$.
2. μ est dite q -minitive ssi pour tout sous-ensemble $X \subseteq \mathcal{C}$ vérifiant $|X| < (n - q)$, nous avons $\mu(X) = \min_{Y \supset X, |Y| \geq (n - q)} \mu(Y)$.

On vérifie facilement qu'une capacité μ est q -maxitive ssi sa conjuguée μ^c est q -minitive. Toute capacité $\mu : 2^{\mathcal{C}} \rightarrow L$ est (trivialement) n -maxitive et n -minitive. De plus, les deux formules (1) et (2) deviennent :

Lemme 1 Soient $\mu : 2^{\mathcal{C}} \rightarrow L$ une capacité, $q \in \{1, 2, \dots, n\}$ où $n = |\mathcal{C}|$ et $x = [x_i] \in L^{n \times 1}$.

1. Si μ est q -maxitive, alors :

$$S_{\mu}(x) = \max_{A \in 2^{\mathcal{C}} \setminus \{\emptyset\}, |A| \leq q} \min_{i \in A} (\min x_i, \mu(A)).$$

2. Si μ est q -minitive, alors nous avons :

$$S_{\mu}(x) = \min_{A \in 2^{\mathcal{C}} \setminus \{\emptyset\}, |A| \geq n - q} \max_{i \in \bar{A}} (\max x_i, \mu(A)).$$

Les capacités q -maxitives et q -minitives permettent donc de calculer l'évaluation globale d'un objet x en considérant moins de sous-ensembles de critères que dans le cas général.

2.2 Résoudre les systèmes d'équations relationnelles floues de type max – min

Un système d'équations relationnelles floues de type max – min basé sur une matrice $A = [a_{ij}] \in L^{n \times m}$ de taille (n, m) et un vecteur $b = [b_i] \in L^{n \times 1}$ de n composantes, est de la forme suivante :

$$(S) : A \square_{\min}^{\max} x = b, \quad (3)$$

où $x = [x_j]_{1 \leq j \leq m} \in L^{m \times 1}$ est un vecteur inconnu de m composantes et l'opérateur \square_{\min}^{\max} est le produit matriciel qui utilise la t-norme min pour le produit et max pour l'addition.

À partir du vecteur

$$e = A^t \square_{\rightarrow G}^{\min} b, \quad (4)$$

où A^t est la transposée de A et le produit matriciel $\square_{\rightarrow G}^{\min}$ utilise l'implication de Gödel \rightarrow_G définie par :

$$x \rightarrow_G y = \begin{cases} 1 & x \leq y \\ y & \text{sinon.} \end{cases} \quad (5)$$

pour le produit et min pour l'addition, Sanchez [13] a montré l'équivalence suivante :

$$\text{Le système } (S) \text{ est compatible} \iff A \square_{\min}^{\max} e = b.$$

Si le système (S) est compatible, le vecteur e est sa plus grande solution [13].

2.3 Résoudre les systèmes d'équations relationnelles floues de type min – max

Un système d'équations relationnelles floues de type min – max se définit similairement. À partir d'une matrice $\Gamma = [\gamma_{ij}] \in L^{n \times m}$ de taille (n, m) et un vecteur $\beta = [\beta_i] \in L^{n \times 1}$ de n composantes, on définit un système d'équations relationnelles floues de type min – max par :

$$(\Sigma) : \Gamma \square_{\max}^{\min} x = \beta, \quad (6)$$

où $x = [x_j]_{1 \leq j \leq m} \in L^{m \times 1}$ est un vecteur inconnu de m composantes et l'opérateur \square_{\max}^{\min} est le produit matriciel qui utilise la fonction max pour le produit et min pour l'addition.

À partir du vecteur

$$f = \Gamma^t \square_{\epsilon}^{\max} \beta, \quad (7)$$

où Γ^t est la transposée de Γ et le produit matriciel $\square_{\epsilon}^{\max}$ utilise le produit epsilon (noté ϵ) défini par :

$$x \epsilon y = \begin{cases} y & x < y \\ 0 & x \geq y. \end{cases} \quad (8)$$

pour le produit et la fonction max pour l'addition, nous avons l'équivalence suivante :

$$\text{Le système } (\Sigma) \text{ est compatible} \iff \Gamma \square_{\max}^{\min} f = \beta.$$

Si le système (Σ) est compatible, le vecteur f est sa plus petite solution [13].

3 Apprentissage de capacités pour les intégrales de Sugeno en fonction de données avec des systèmes d'équations relationnelles floues

Dans cette section, nous introduisons notre méthode d'apprentissage de capacités pour les intégrales de Sugeno en fonction de données d'entraînement basée sur les systèmes d'équations relationnelles floues. Nous considérons un ensemble de N données d'apprentissage $\{(x^{(1)}, \alpha^{(1)}), (x^{(2)}, \alpha^{(2)}), \dots, (x^{(N)}, \alpha^{(N)})\}$. Pour $k = 1, 2, \dots, N$, chaque paire $(x^{(k)}, \alpha^{(k)})$ est composée d'un objet $x^{(k)} = [x_i^{(k)}] \in L^{n \times 1}$ qui contient l'évaluation partielle $x_i^{(k)}$ de l'objet $x^{(k)}$ selon le critère i et l'évaluation globale ciblée de l'objet qui est désignée par $\alpha^{(k)} \in L$.

3.1 Construction de deux systèmes d'équations à partir de données

À partir des données $(x^{(k)}, \alpha^{(k)})_{1 \leq k \leq N}$, on construit deux matrices et un vecteur :

- $M = [m_{k,A}]_{1 \leq k \leq N, A \in 2^C \setminus \{\emptyset\}}$ où $m_{k,A} = \min_{i \in A} x_i^{(k)}$,
- $\Gamma = [\gamma_{k,A}]_{1 \leq k \leq N, A \in 2^C \setminus \{C\}}$ où $\gamma_{k,A} = \max_{i \in \bar{A}} x_i^{(k)}$,
- $\alpha = [\alpha^{(k)}] \in L^{N \times 1}$: un vecteur contenant les évaluations globales ciblées des objets.

Nous formons le système d'équations de type max – min suivant à partir de M et α :

$$(S) : M \square_{\min}^{\max} X = \alpha \quad (9)$$

où $X = [\xi_A]_{A \in 2^C \setminus \{\emptyset\}}$ est un vecteur inconnu et $e = M^t \square_{\rightarrow G}^{\min} \alpha$ est la plus grande solution potentielle du système (S) , voir (4). Nous formons le système d'équations de type min – max suivant à partir de Γ et α :

$$(\Sigma) : \Gamma \square_{\max}^{\min} X = \alpha \quad (10)$$

où $X = [\xi_A]_{A \in 2^C \setminus \{C\}}$ est un vecteur inconnu et $f = \Gamma^t \square_{\leftarrow G}^{\max} \alpha$ est la plus petite solution potentielle du système (Σ) , voir (7).

3.2 Résultats préliminaires

Les solutions extrémales potentielles des deux systèmes (S) et (Σ) vues comme des fonctions d'ensembles $e : 2^C \setminus \{\emptyset\} \rightarrow L$ et $f : 2^C \setminus \{C\} \rightarrow L$ sont croissantes :

Lemme 2

1. $\emptyset \subset A \subseteq B \implies e(A) \leq e(B)$.
2. $A \subseteq B \subset C \implies f(A) \leq f(B)$.

Ces deux implications se déduisent des propriétés de l'implication de Gödel \rightarrow_G , voir (5) et de celles du produit ϵ , voir (8) : $\forall x, x' \in L$ tels que $x \leq x'$, on a $\forall y \in L$, $x' \rightarrow_G y \leq x \rightarrow_G y$ et $x' \epsilon y \leq x \epsilon y$.

Pour tout $\emptyset \subset A \subseteq C$ et $B \subset C$, posons :

$$J_A = \{k \in \{1, 2, \dots, N\} \mid \min_{i \in A} x_i^{(k)} > \alpha^{(k)}\} \text{ et}$$

$$K_B = \{k \in \{1, 2, \dots, N\} \mid \max_{i \in \bar{B}} x_i^{(k)} < \alpha^{(k)}\}.$$

Avec ces notations, nous avons :

Lemme 3

1. $e(A) = \min_{k \in J_A} \alpha^{(k)}$, $\min_{\emptyset} = 1$.
2. $f(B) = \max_{k \in K_B} \alpha^{(k)}$, $\max_{\emptyset} = 0$.

En particulier, on a :

$$e(C) = 1 \iff \forall k \in \{1, 2, \dots, N\}, \min_{i \in C} x_i^{(k)} \leq \alpha^{(k)} \text{ et}$$

$$f(\emptyset) = 0 \iff \forall k \in \{1, 2, \dots, N\}, \alpha^{(k)} \leq \max_{i \in C} x_i^{(k)}.$$

3.3 Apprentissage de capacités

Nous construisons deux fonctions d'ensembles à partir des solutions extrémales potentielles e et f des systèmes (S) et (Σ) définies à partir des données d'entraînement, voir (9) et (10), comme suit :

$$\mu_e(A) = \begin{cases} e(A) & \text{si } A \neq \emptyset \\ 0 & \text{si } A = \emptyset \end{cases}, \quad (11)$$

$$\mu_f(A) = \begin{cases} f(A) & \text{si } A \subset C \\ 1 & \text{si } A = C \end{cases}. \quad (12)$$

Pour garantir que μ_e, μ_f sont des capacités compatibles avec les données d'apprentissage, nous voulons que les systèmes (S) et (Σ) aient des solutions et nous voulons que $\mu_e(C) = e(C) = 1$ et $\mu_f(\emptyset) = f(\emptyset) = 0$. Nous avons établi le résultat suivant :

Théorème 1 Les trois conditions suivantes sont équivalentes :

1. Il existe une capacité $\mu : 2^{\mathcal{C}} \rightarrow L$ qui représente les données i.e., une capacité telle que $\forall k \in \{1, 2, \dots, N\}$, on a $S_\mu(x^{(k)}) = \alpha^{(k)}$,
2. Le système $(S) : M \square_{\min}^{\max} X = \alpha$ est compatible et $e(\mathcal{C}) = 1$,
3. Le système $(\Sigma) : \Gamma \square_{\max}^{\min} X = \alpha$ est compatible et $f(\emptyset) = 0$.

Preuve abrégée : pour toute capacité $\mu : 2^{\mathcal{C}} \rightarrow L$, posons : $v : 2^{\mathcal{C}} \setminus \{\emptyset\} \rightarrow L : A \mapsto v(A) = \mu(A)$ et $w : 2^{\mathcal{C}} \setminus \{\mathcal{C}\} \rightarrow L : A \mapsto w(A) = \mu(A)$. Alors il résulte de (1) et (2) qu'on a :

$$\forall k \in \{1, 2, \dots, N\}, S_\mu(x^{(k)}) = \alpha^{(k)} \iff M \square_{\min}^{\max} v = \alpha \\ \iff \Gamma \square_{\max}^{\min} w = \alpha.$$

Corollaire 1 Si les conditions équivalentes du Théorème 1 sont vérifiées, on a :

- μ_e est une capacité qui représente les données d'entraînement et c'est la plus grande.
- μ_f est une capacité qui représente les données d'entraînement et c'est la plus petite.
- Pour toute capacité $\mu : 2^{\mathcal{C}} \rightarrow L$ représentant les données d'entraînement, nous avons : $\mu_f \leq \mu \leq \mu_e$.

Les méthodes présentées dans cet article pour obtenir les capacités extrémales μ_e et μ_f qui représentent les données d'entraînement, basées sur les résultats de Sanchez [13], sont plus simples que celles données dans [12].

3.4 Constructions de capacités q -maxitives et q -minitives

Pour une donnée $(x^{(k)}, \alpha^{(k)})$, une condition nécessaire pour qu'il soit possible de relier l'objet $x^{(k)}$ à son évaluation globale ciblée $\alpha^{(k)}$ par une intégrale de Sugeno, est $\min_{i \in \mathcal{C}} x_i^{(k)} \leq \alpha^{(k)} \leq \max_{i \in \mathcal{C}} x_i^{(k)}$. Pour notre problème d'apprentissage, si pour tout $k \in \{1, 2, \dots, N\}$, chaque donnée d'entraînement $(x^{(k)}, \alpha^{(k)})$ respecte cette contrainte, alors nous pouvons donner des conditions suffisantes pour l'existence de capacités q -maxitives et q -minitives (où $1 \leq q \leq n$), que l'on déduit des solutions extrémales des systèmes (S) et (Σ) .

Proposition 1 Soit $q \in \{1, 2, \dots, n\}$. Supposons qu'il existe $A \subseteq \mathcal{C}$ tel que $|A| = q$ et $e(A) = 1$. Posons :

$$\mu_{\vee} : 2^{\mathcal{C}} \rightarrow L \\ : X \mapsto \mu_{\vee}(X) \\ = \begin{cases} 0 & \text{si } X = \emptyset \\ e(X) & \text{si } 0 < |X| \leq q \\ \max_{\emptyset \subset Y \subset X, |Y| \leq q} e(Y) & \text{si } |X| > q \end{cases}$$

Alors μ_{\vee} est une capacité q -maxitive sur \mathcal{C} .

Proposition 2 Soit $q \in \{1, 2, \dots, n\}$. Supposons qu'il existe $A' \subseteq \mathcal{C}$ tel que $|A'| = n - q$ et $f(A') = 0$. Posons :

$$\mu_{\wedge} : 2^{\mathcal{C}} \rightarrow L \\ : X \mapsto \mu_{\wedge}(X) \\ = \begin{cases} 1 & \text{si } X = \mathcal{C} \\ f(X) & \text{si } n - q \leq |X| < n \\ \min_{Y \supset X, n - q \leq |Y| < n} f(Y) & \text{si } |X| < n - q \end{cases}$$

Alors μ_{\wedge} est une capacité q -minitive sur \mathcal{C} .

3.5 Exemple

Soit $n = 3$ critères i.e., $\mathcal{C} = \{1, 2, 3\}$. L'échelle est $L = \{0.05 \cdot l \mid 1 \leq l \leq 20\}$. On considère les $N = 3$ données d'entraînement suivantes :

- $x^{(1)} = (0.15, 0.2, 0.3)$, $\alpha^{(1)} = 0.2$,
- $x^{(2)} = (0.5, 0.25, 0.3)$, $\alpha^{(2)} = 0.3$,
- $x^{(3)} = (0.4, 0.7, 0.35)$, $\alpha^{(3)} = 0.4$.

On construit le système $(S) : M \square_{\min}^{\max} X = \alpha$ où

$$M = \begin{array}{c|c|c|c|c|c|c} \{1\} & \{2\} & \{3\} & \{1,2\} & \{1,3\} & \{2,3\} & \{1,2,3\} \\ \hline 0.15 & 0.2 & 0.3 & 0.15 & 0.15 & 0.2 & 0.15 \\ 0.5 & 0.25 & 0.3 & 0.25 & 0.3 & 0.25 & 0.25 \\ 0.4 & 0.7 & 0.35 & 0.4 & 0.35 & 0.35 & 0.35 \end{array}$$

et $\alpha = [0.2 \ 0.3 \ 0.4]^t$. Le système (S) est compatible et sa plus grande solution (4) est :

$$e = \begin{array}{c|c|c|c|c|c|c} \{1\} & \{2\} & \{3\} & \{1,2\} & \{1,3\} & \{2,3\} & \{1,2,3\} \\ \hline 0.3 & 0.4 & 0.2 & 1 & 1 & 1 & 1 \end{array}^t$$

et on a donc $e(\mathcal{C}) = 1$. La fonction d'ensembles notée μ_e , voir (11), est une capacité 2-maxitive qui représente les données. C'est la plus grande, et elle n'est pas une mesure de possibilité, car nous avons

$\mu_e(\{1, 2\}) = 1 > \max(\mu_e(\{1\}), \mu_e(\{2\}))$. On peut construire similairement le système (Σ) et obtenir la capacité μ_f qui représente les données et qui est la plus petite.

4 Apprentissage de capacités q -maxitives/minitives en pratique

Dans la pratique, on se restreint à apprendre les mesures des ensembles de critères de cardinalité inférieure ou égale à q (avec $1 \leq q < n$), d'une capacité en considérant le système suivant :

$$(S_q) : M_q \square_{\min}^{\max} X = \alpha,$$

où $M_q = [m_{k,A}]_{1 \leq k \leq N, A \in 2^{\mathcal{C}} \setminus \{\emptyset\}, |A| \leq q}$ et on note e_q la plus grande solution potentielle de (S_q) , voir (4). Pour $q = n$, il est clair que le système (S_n) coïncide avec le système (S) , voir (9).

Propriété 1 *On dit que le système (S_q) vérifie la propriété de q -maxitivité si $\exists A \subseteq C$ tel que $0 < |A| = q$ et $e_q(A) = 1$.*

Pour tout sous-ensemble non vide $A \subseteq C$ tel que $|A| \leq q$, on a $e_q(A) = e(A)$. La condition $e_q(A) = 1$ signifie que $\forall k \in \{1, 2, \dots, N\}$, $\min_{i \in A} x_i^{(k)} \leq \alpha^{(k)}$ voir Lemme 3. On a le résultat suivant :

Théorème 2 *Les deux assertions suivantes sont équivalentes :*

1. *Il existe une capacité q -maxitive $\mu : 2^{\mathcal{C}} \rightarrow L$ qui représente les données d'entraînement : $\forall k \in \{1, 2, \dots, N\}$, on a $S_\mu(x^{(k)}) = \alpha^{(k)}$,*
2. *Le système $(S_q) : M_q \square_{\min}^{\max} X = \alpha$ est compatible et vérifie Propriété 1.*

Preuve abrégée : Soit $\mu : 2^{\mathcal{C}} \rightarrow L$ une capacité q -maxitive. Pour tout $A \in 2^{\mathcal{C}} \setminus \{\emptyset\}$ vérifiant $|A| \leq q$, posons $v(A) = \mu(A)$. Alors il résulte de la formule (1) qu'on a :

$$\forall k \in \{1, 2, \dots, N\}, S_\mu(x^{(k)}) = \alpha^{(k)} \iff M_q \square_{\min}^{\max} v = \alpha.$$

Pour $q \in \{1, 2, \dots, n\}$, l'étude de l'apprentissage en pratique de capacité q -minitive est analogue. On construit le système réduit $(\Sigma_q) : \Gamma_q \square_{\min}^{\max} X = \alpha$ avec $\Gamma_q = [\gamma_{k,A}]_{1 \leq k \leq N, A \in 2^{\mathcal{C}} \setminus \{C\}, |A| \geq n-q}$ et on note f_q sa plus petite solution potentielle, voir (7).

Propriété 2 *On dit que le système (Σ_q) vérifie la propriété de q -minitivité si $\exists A \subset C$ tel que $|A| = n - q$ et $f_q(A) = 0$.*

Pour tout sous-ensemble $A \subset C$ tel que $|A| \geq n - q$, on a $f_q(A) = f(A)$. La condition $f_q(A) = 0$ signifie que $\forall k \in \{1, 2, \dots, N\}$, $\alpha^{(k)} \leq \max_{i \in \bar{A}} x_i^{(k)}$ (Lemme 3).

Théorème 3 *Les deux assertions suivantes sont équivalentes :*

1. *Il existe une capacité q -minitive $\mu : 2^{\mathcal{C}} \rightarrow L$ qui représente les données d'entraînement : $\forall k \in \{1, 2, \dots, N\}$, on a $S_\mu(x^{(k)}) = \alpha^{(k)}$,*
2. *Le système $(\Sigma_q) : \Gamma_q \square_{\min}^{\max} X = \alpha$ est compatible et vérifie Propriété 2.*

Preuve abrégée : Soit $\mu : 2^{\mathcal{C}} \rightarrow L$ une capacité q -minitive. Pour tout $A \in 2^{\mathcal{C}} \setminus \{C\}$ vérifiant $|A| \geq n - q$, posons $w(A) = \mu(A)$. Alors il résulte de la formule (2) qu'on a :

$$\forall k \in \{1, 2, \dots, N\}, S_\mu(x^{(k)}) = \alpha^{(k)} \iff \Gamma_q \square_{\min}^{\max} w = \alpha.$$

4.1 Apprentissage de capacités approchées

Si l'on utilise les systèmes d'équations relationnelles floues (S_q) et (Σ_q) avec des données réelles, qui peuvent être sujettes au bruit ou contenir des valeurs aberrantes, les systèmes (S_q) et (Σ_q) peuvent être incompatibles. Dans ce cas, si $L = [0, 1]$, nous pouvons nous appuyer sur les travaux de [2, 3] sur l'incompatibilité des systèmes d'équations relationnelles floues de type max - min et min - max pour obtenir des capacités qui représentent approximativement les données d'entraînement. Pour un système d'équations relationnelles floues incompatible de type max-min $A \square_{\min}^{\max} x = b$, l'auteur de [2, 3] donne une formule analytique explicite pour calculer la distance de Tchebychev notée $\Delta = \Delta(A, b) = \min_{c \in \mathcal{C}} \|b - c\|_\infty$ (exprimée avec la norme L_∞) entre le second membre b et l'ensemble

$$\mathcal{C} = \{c \in [0, 1]^{n \times 1} \mid \text{le système } A \square_{\min}^{\max} x = c \text{ est compatible}\}$$

formé des seconds membres des systèmes compatibles définis avec la même matrice : la matrice A du système incompatible considéré. Le

même résultat est donné dans [3] pour les systèmes de type min – max, où la distance de Tchebychev est notée ∇ . La formule de $\Delta_q = \Delta(M_q, \alpha)$ pour le système (S_q) est la suivante (nous utilisons la notation : $x^+ = \max(x, 0)$) :

$$\Delta_q = \max_{1 \leq i \leq N} \delta_i \text{ avec } \delta_i = \min_{A \in 2^{\mathcal{C}} \setminus \{\emptyset\}, |A| \leq q} \delta_{i,A} \quad (13)$$

où $\delta_{i,A} = \max(\alpha^{(i)} - m_{i,A})^+, \max_{1 \leq l \leq N} \sigma_G(\alpha^{(i)}, m_{i,A}, \alpha^{(l)})$ et $\sigma_G(\alpha^{(i)}, m_{i,A}, \alpha^{(l)}) = \min(\frac{(\alpha^{(i)} - \alpha^{(l)})^+}{2}, (m_{i,A} - \alpha^{(l)})^+)$. La valeur de Δ_q donne une estimation de la qualité des données d'entraînement utilisées pour construire le système (S_q) i.e., les données sont de bonne qualité si $\Delta_q \approx 0$ et non dans le cas contraire. Notons que pour $q \leq q'$, on a $\Delta_{q'} \leq \Delta_q$.

Les outils donnés dans [3] permettent ensuite d'obtenir des solutions approchées du système (S_q) via Δ_q dont on peut déduire des capacités q -maxitives. Nous nous appuyons sur la plus grande solution approchée du système (S_q) , qui est le vecteur $\eta_q = M_q^t \square_{\rightarrow G}^{\min} \bar{\alpha}(\Delta_q)$ où $\bar{\alpha}(\Delta_q) = [\min(\alpha^{(i)} + \Delta_q, 1)]_{1 \leq i \leq N}$, voir [3], pour donner des conditions suffisantes pour l'existence de la plus grande capacité q -maxitive approchée et nous montrons comment la construire :

Proposition 3 Si $\exists A \subseteq \mathcal{C}$ tel que $|A| = q$ et $\eta_q(A) = 1$, posons :

$$\begin{aligned} \mu^* : 2^{\mathcal{C}} &\rightarrow L \\ &: X \mapsto \mu^*(X) \\ &= \begin{cases} 0 & \text{si } X = \emptyset \\ \eta_q(X) & \text{si } 0 < |X| \leq q \\ \max_{\emptyset \subset Y \subset X, |Y| \leq q} \eta_q(Y) & \text{si } |X| > q \end{cases} \end{aligned}$$

Alors μ^* est une capacité q -maxitive qui vérifie $\max_{1 \leq k \leq N} |S_{\mu^*}(x^{(k)}) - \alpha^{(k)}| = \Delta_q$.

La preuve de cette proposition repose sur le fait que le vecteur $[S_{\mu^*}(x^{(k)})]_{1 \leq k \leq N}$ (respectivement η_q), est la plus grande approximation de Tchebychev (respectivement la plus grande solution approchée) du système (S_q) : $M_q \square_{\min}^{\max} X = \alpha$, voir [3].

En utilisant cette proposition et les résultats de [3], nous déduisons les deux résultats suivants :

Théorème 4 Si $\exists A \subseteq \mathcal{C}$ tel que $|A| = q$ et $\eta_q(A) = 1$ alors :

$$\min_{\mu \text{ capacité } q\text{-maxitive}} \max_{1 \leq k \leq N} |S_{\mu}(x^{(k)}) - \alpha^{(k)}| = \Delta_q.$$

L'erreur d'apprentissage minimale des capacités q -maxitives est égale à Δ_q .

L'idée de la preuve repose sur Proposition 3 et le fait que pour toute capacité q -maxitive μ , le système $M_q \square_{\min}^{\max} X = c$, où $c = [S_{\mu}(x^{(k)})]_{1 \leq k \leq N}$, est un système compatible. Remarquons que l'algorithme de [4] donne une valeur approchée de $\min_{\mu \text{ capacité}} \max_{1 \leq k \leq N} |S_{\mu}(x^{(k)}) - \alpha^{(k)}|$ alors que la formule (13) nous permet d'obtenir la valeur exacte i.e., Δ_n . En conséquence de ce théorème, nous avons le résultat important suivant :

Corollaire 2 Si $\exists A \subseteq \mathcal{C}$ tel que $|A| = q$ et $\eta_q(A) = 1$, alors la capacité μ^* (Proposition 3), est la plus grande capacité q -maxitive μ qui vérifie $\max_{1 \leq k \leq N} |S_{\mu}(x^{(k)}) - \alpha^{(k)}| = \Delta_q$.

Pour apprendre des capacités q -minitives approchées, on utilise le système (Σ_q) : $\Gamma_q \square_{\max}^{\min} X = \alpha$ dont la distance de Tchebychev est :

$$\nabla_q = \max_{1 \leq i \leq N} \nabla_i \text{ avec } \nabla_i = \min_{A \in 2^{\mathcal{C}} \setminus \{\emptyset\}, |A| \geq n-q} \nabla_{i,A}$$

où $\nabla_{i,A} = \max((\gamma_{i,A} - \alpha^{(i)})^+, \max_{1 \leq l \leq N} \sigma_{\epsilon}(\alpha^{(i)}, \gamma_{i,A}, \alpha^{(l)}))$ et $\sigma_{\epsilon}(\alpha^{(i)}, \gamma_{i,A}, \alpha^{(l)}) = \min(\frac{(\alpha^{(l)} - \alpha^{(i)})^+}{2}, (\alpha^{(l)} - \gamma_{i,A})^+)$. Nous nous appuyons sur la plus petite solution approchée du système (Σ_q) , qui est le vecteur $\nu_q = \Gamma_q^t \square_{\epsilon}^{\max} \underline{\alpha}(\nabla_q)$ où $\underline{\alpha}(\nabla_q) = [\max(\alpha^{(i)} - \nabla_q, 0)]_{1 \leq i \leq N}$, voir [3]. Comme dans le cas des capacités q -maxitives, nous avons :

Proposition 4 Si $\exists A \subseteq \mathcal{C}$ tel que $|A| = n - q$ et $\nu_q(A) = 0$, posons :

$$\begin{aligned} \mu_* : 2^{\mathcal{C}} &\rightarrow L \\ &: X \mapsto \mu_*(X) \\ &= \begin{cases} 1 & \text{si } X = \mathcal{C} \\ \nu_q(X) & \text{si } n - q \leq |X| < n \\ \min_{Y \supset X, n-q \leq |Y| < n} \nu_q(Y) & \text{si } |X| < n - q \end{cases} \end{aligned}$$

Alors μ_* est une capacité q -minitive qui vérifie $\max_{1 \leq k \leq N} |S_{\mu_*}(x^{(k)}) - \alpha^{(k)}| = \nabla_q$.

La preuve de cette proposition repose sur le fait que le vecteur $[S_{\mu^*}(x^{(k)})]_{1 \leq k \leq N}$ (respectivement ν_q), est la plus petite approximation de Tchebychev (respectivement la plus petite solution approchée) du système $(\Sigma_q) : \Gamma_q \square_{\min}^{\max} X = \alpha$, voir [3]. En utilisant cette proposition et [3], nous déduisons les deux résultats suivants :

Théorème 5 *Si $\exists A \subseteq \mathcal{C}$ tel que $|A| = n - q$ et $\nu_q(A) = 0$ alors :*

$$\min_{\mu \text{ capacité } q\text{-minitive}} \max_{1 \leq k \leq N} |S_{\mu}(x^{(k)}) - \alpha^{(k)}| = \nabla_q.$$

L'erreur d'apprentissage minimale des capacités q -minitives est égale à ∇_q .

L'idée de la preuve repose sur Proposition 4 et le fait que pour toute capacité q -minitive μ , le système $\Gamma_q \square_{\min}^{\max} X = c$, où $c = [S_{\mu}(x^{(k)})]_{1 \leq k \leq N}$, est un système compatible.

Corollaire 3 *Si $\exists A \subseteq \mathcal{C}$ tel que $|A| = n - q$ et $\nu_q(A) = 0$, alors la capacité μ_* (Proposition 4), est la plus petite capacité q -minitive μ qui vérifie $\max_{1 \leq k \leq N} |S_{\mu}(x^{(k)}) - \alpha^{(k)}| = \nabla_q$.*

5 Conclusion

Dans cet article, nous avons introduit une méthode d'apprentissage des capacités pour les intégrales de Sugeno en fonction des données d'entraînement basée sur les systèmes d'équations relationnelles floues. Nous avons montré en particulier que cette approche permet d'apprendre des capacités approchées q -maxitives et q -minitives en un sens précis.

En perspective, nous travaillons actuellement sur les liens entre les systèmes (S) et (Σ) lorsqu'ils sont incompatibles, afin de relier leurs ensembles de solutions approximatives respectifs via les outils de [3]. Enfin, nous voulons tester notre méthode sur des données réelles.

Remerciements

L'auteur remercie Sébastien Destercke pour d'utiles discussions. Ce travail a été partiellement soutenu par TAILOR, un projet financé par le programme de recherche et d'innovation Horizon 2020 de l'UE sous le numéro GA 952215.

Références

- [1] Abbaszadeh, S. et E. Hüllermeier. 2020, «Machine learning with the sugeno integral : The case of binary classification», *IEEE Trans. Fuzzy Syst.*, vol. 29, n° 12, p. 3723–3733.
- [2] Baaj, I. 2023, «Approximations de tchebyshev d'un système incompatible d'équations relationnelles floues de type max- t», dans *Logique Floue et Applications*, Cepadues, p. 101–108.
- [3] Baaj, I. 2024, «On the handling of inconsistent systems of max-min fuzzy relational equations», *Fuzzy Sets and Systems*, p. 108912.
- [4] Beliakov, G., M. Gagolewski et S. James. 2020, «Robust fitting for the sugeno integral with respect to general fuzzy measures», *Inf. Sci.*, vol. 514, p. 449–461.
- [5] Calvo, T. et B. De Baets. 1998, «Aggregation operators defined by k-order additive/maxitive fuzzy measures», *Int. J. Uncertain. Fuzziness Knowl. Based Syst.*, vol. 6, n° 06, p. 533–550.
- [6] Dubois, D., C. Durrieu, H. Prade, A. Rico et Y. Ferro. 2015, «Extracting decision rules from qualitative data using sugeno integral : a case-study», dans *Eur. Conf. on Symb. and Quant. Approaches to Reasoning and Uncertainty*, Springer, p. 14–24.
- [7] Dubois, D., J.-L. Marichal, H. Prade, M. Roubens et R. Sabbadin. 2001, «The use of the discrete sugeno integral in decision-making : A survey», *Int. J. Uncertain. Fuzziness Knowl. Based Syst.*, vol. 9, n° 05, p. 539–561.
- [8] Grabisch, M. et C. Labreuche. 2010, «A decade of application of the choquet and sugeno integrals in multi-criteria decision aid», *Ann. Oper. Res.*, vol. 175, p. 247–286.
- [9] Grabisch, M. et collab.. 2016, *Set functions, games and capacities in decision making*, vol. 46, Springer.
- [10] Marichal, J.-L. 2000, «On sugeno integral as an aggregation function», *Fuzzy sets and systems*, vol. 114, n° 3, p. 347–365.
- [11] Mesiar, R. 1999, «K-order additive fuzzy measures», *Int. J. Uncertain. Fuzziness Knowl. Based Syst.*, vol. 7, n° 06, p. 561–568.
- [12] Rico, A., M. Grabisch, C. Labreuche et A. Chateauf. 2005, «Preference modeling on totally ordered sets by the sugeno integral», *Discrete Applied Mathematics*, vol. 147, n° 1, p. 113–124.
- [13] Sanchez, E. 1976, «Resolution of composite fuzzy relation equations», *Inf. Control.*, vol. 30, p. 38–48.
- [14] Sugeno, M. 1974, *Theory of fuzzy integrals and its applications*, thèse de doctorat, Tokyo Institute of technology.
- [15] Wu, J.-Z. et G. Beliakov. 2019, «k-minitive capacities and k-minitive aggregation functions», *Journal of Intelligent & Fuzzy Systems*, vol. 37, n° 2, p. 2797–2808.