

**XIX Congreso Español
sobre Tecnologías
y Lógica Fuzzy
(XIX ESTYLF)**

ESTYLF 9: SESIÓN ESPECIAL:
APLICACIONES DE
FUNCIONES DE AGREGACIÓN
Y PRE-AGREGACIÓN

Organizadores:

HUMBERTO BUSTINCE, TOMASA CALVO,
JAVIER FERNÁNDEZ, LAURA DE MIGUEL





Condiciones de monotonía direccional: interpretación geométrica

1st Laura De Miguel

Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, Spain
laura.demiguel@unavarra.es

2nd Mikel Sesma-Sara

Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, Spain
mikel.sesma@unavarra.es

3rd Antonio Roldán

Departamento de Didáctica de la Matemática
Universidad de Granada
Granada, Spain
aroldan@ugr.es

4th Julio Lafuente,

Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, Spain
lafuente@unavarra.es

5th Radko Mesiar

Department of Mathematics and Descriptive Geometry
Slovak University of Technology in Bratislava
Bratislava, Slovak Republic
mesiar@math.sk

6th Humberto Bustince

Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, Spain
bustince@unavarra.es

Abstract—En los últimos años ha habido un gran interés en estudiar condiciones de monotonía menos restrictivas para construir nuevas familias de funciones de fusión de datos. Varias de estas condiciones son la monotonía débil, la monotonía direccional, la monotonía direccional ordenada, y la monotonía direccional ordenada reforzada. En este trabajo presentamos algunos resultados sobre las monotonías anteriores desde el punto de vista de la geometría en dos dimensiones.

Index Terms—Funciones de agregación, condiciones de monotonía, geometría bidimensional

I. INTRODUCCIÓN

La cantidad de aparatos inteligentes y sensores que nos rodea ha crecido exponencialmente en los últimos años lo que conlleva que cada vez tenemos mayor volumen de datos de todo tipo y naturaleza. Debido al prácticamente inmanejable volumen de datos, la necesidad de encontrar operadores apropiados para fusionar los datos de acuerdo a la aplicación considerada es más importante cada día.

Desde sus inicios, las funciones de agregación han jugado un papel crucial en varias y diversas aplicaciones [1]–[4]. Además, dependiendo del tipo de propiedades que se impongan a las funciones de agregación (véase [5] para más información), distintas clases de funciones han sido definidas en la literatura: t-normas/t-conormas [6], uninormas [7] o funciones de solapamiento (*overlap functions*) [8], entre otras.

Este trabajo ha sido financiado por el Servicio de Investigación de la Universidad Pública de Navarra y por el proyecto TIN2016-77356-P del MINECO, AEI/FEDER, UE y por el proyecto APVV-14-0013.

La condición de monotonía impuesta a las funciones de agregación juega un papel muy importante en aplicaciones como la toma de decisión donde el aumento de un valor (que significa que un criterio o una alternativa es mejor) no debe nunca implicar un descenso de la puntuación final o utilidad. Sin embargo, en otro tipo de aplicaciones, en especial en áreas donde es frecuente la aparición de *outliers* o datos afectados por diversos ruidos, la monotonía impuesta en cada argumento es un inconveniente en vez de una ventaja. Esto demuestra, que para algunos casos particulares de aplicación, la condición de monotonía que satisfacen las funciones de agregación, es decir, el crecimiento en cada argumento de la función puede ser demasiado restrictiva y a veces no es necesario que las funciones cumplan una condición tan fuerte para dar unos buenos resultados.

Teniendo en cuenta todas las consideraciones anteriores sobre las funciones de agregación, en la literatura reciente, encontramos diversos esfuerzos de definir y estudiar diversas condiciones de monotonía menos restrictivas. La monotonía débil [9], la monotonía direccional [10], la monotonía direccional ordenada [11] y la monotonía direccional ordenada reforzada son algunos ejemplos de nuevas condiciones de monotonía que generan nuevas clases de operadores de fusión de datos. Estas nuevas nociones de monotonía han sido aplicados en diversas aplicaciones como la clasificación o procesamiento de imagen. Por ejemplo, funciones de pre-agregación [12], que solo satisfacen monotonías direccionales han demostrado mejorar los resultados en algunos problemas de clasificación [13], mientras que las monotonías or-

denadas direccionales han demostrado mejorar los resultados en detección de bordes [11], [14] y extracción de características [15].

Todos los trabajos en los que se introducen las distintas nociones de monotonía trabajan estos conceptos desde un punto de vista formal dentro del análisis matemático. En este trabajo, con el propósito de conseguir una mejor comprensión de las distintas condiciones, así como sus diferencias nos centramos en la interpretación geométrica de las distintas condiciones de monotonía introducidas en la literatura reciente [9]–[11], [16]. Para ello, nos centramos en el caso de las funciones 2-dimensionales, cuyo conjunto dominio es el $[0, 1]^2$ (representable en el plano). Las ventajas de esta aproximación son claras, ya que nos permite descubrir gráficamente el comportamiento de cada noción de monotonía descrita en los trabajos anteriores.

El artículo está organizado como sigue. En la Sección II, recordamos las diferentes nociones de monotonía (restringidas al caso 2-dimensional) a la vez que fijamos algunas notaciones. En la Sección III, estudiamos la interpretación geométrica de las distintas nociones centrándonos en la monotonía direccional ordenada y monotonía direccional ordenada reforzada, probando algunos resultados sobre el comportamiento de estas funciones en el caso 2-dimensional. En la Sección IV, acabamos con algunas conclusiones y trabajo a futuro.

II. CONDICIONES DE MONOTONÍA

En este trabajo nos centramos en la representación geométrica de algunas nociones de monotonía definidas en la literatura reciente. Por claridad y sencillez de los resultados nos restringimos al caso 2-dimensional, pero es importante resaltar que las definiciones originales, que podemos encontrar en los trabajos [9]–[11], [16] están introducidas para el caso n -dimensional.

Consideramos vectores $\vec{r} = (r_1, r_2) \in \mathbb{R}^2$ y puntos en el plano $x = (x_1, x_2) \in [0, 1]^2$ ordenados según el orden parcial heredado de \mathbb{R}^2 que viene dado por $x = (x_1, x_2) \leq y = (y_1, y_2)$ si y solo si $x_1 \leq x_2$ y $y_1 \leq y_2$. Sea $S = \{\sigma_1 = (1, 2), \sigma_2 = (2, 1)\}$ el conjunto de permutaciones en dos dimensiones. Dados los elementos $x \in [0, 1]^2$ y $\vec{r} \in \mathbb{R}^2$, el elemento x_{σ_i} representa el punto en el plano ordenado según la permutación $x_{\sigma_i} = (x_{\sigma_i(1)}, x_{\sigma_i(2)})$ mientras que \vec{r}_{σ_i} representa el vector ordenado $\vec{r}_{\sigma_i} = (r_{\sigma_i(1)}, r_{\sigma_i(2)})$.

Las funciones de agregación son una clase de funciones de fusión a las que se les impone la condición de monotonía con respecto a cada argumento así como condiciones frontera en el 0 y en el 1. La definición de las funciones de agregación 2-dimensionales es la siguiente.

Definición 2.1: ([17], [18]) Una función $A : [0, 1]^2 \rightarrow [0, 1]$ se llama una función de agregación si satisface la siguientes propiedades:

- $A(0, 0) = 0$
- $A(1, 1) = 1$
- $A(x_1, x_2) \leq A(y_1, y_2)$ para todo par de puntos $x = (x_1, x_2), y = (y_1, y_2) \in [0, 1]^2$ tales que $x_1 \leq y_1$ y $x_2 \leq y_2$.

Aunque las agregaciones han demostrado ser muy útiles en diversas aplicaciones, algunos operadores muy conocidos, como las medias Lehmer o la moda no cumplen las monotonía exigida a las funciones de agregación. Es por ello, que estudios recientes han intentado estudiar funciones de fusión con condiciones menos restrictivas para poder estudiar clases de funciones de fusión más generales. Algunas de estas condiciones menos restrictivas son la monotonía direccional, la monotonía direccional ordenada y la monotonía direccional ordenada reforzada. A continuación recordamos estas condiciones de monotonía para el caso de funciones 2-dimensionales.

Definición 2.2: ([9]): Una función $F : [0, 1]^2 \rightarrow [0, 1]$ se dice débilmente creciente si la siguiente desigualdad se satisface

$$F(x_1 + c, x_2 + c) \geq F(x_1, x_2)$$

para todo $(x_1, x_2) \in [0, 1]^2$ y $c \in [0, 1]$ tal que $(x_1, x_2) + (c, c) \in [0, 1]^2$.

Definición 2.3 ([10]): Sea $\vec{r} = (r_1, r_2)$ un vector real 2-dimensional distinto del nulo. Una función $F : [0, 1]^2 \rightarrow [0, 1]$ se dice \vec{r} -creciente si para todo punto $(x_1, x_2) \in [0, 1]^2$ y para todo $c > 0$ tales que $(x_1 + cr_1, x_2 + cr_2) \in [0, 1]^2$ se cumple la siguiente condición

$$F(x_1 + cr_1, x_2 + cr_2) \geq F(x_1, x_2).$$

Definición 2.4 ([11]): Sea $\vec{r} = (r_1, r_2)$ un vector real 2-dimensional distinto del nulo. Una función $F : [0, 1]^2 \rightarrow [0, 1]$ se dice \vec{r} -creciente Ordenada Direccional (OD) si para todo $(x_1, x_2) \in [0, 1]^2$, y para cada permutación $\sigma \in S$ tal que $x_{\sigma(1)} \geq x_{\sigma(2)}$ y para cualquier $c > 0$ que cumple

$$1 \geq x_{\sigma(1)} + cr_1 \geq x_{\sigma(2)} + cr_2 \geq 0, \quad (1)$$

se cumple que

$$F(x_1 + cr_{\sigma^{-1}(1)}, x_2 + cr_{\sigma^{-1}(2)}) \geq F(x_1, x_2).$$

Definición 2.5 ([16]): Sea $\vec{r} = (r_1, r_2)$ un vector real 2-dimensional distinto del nulo. Una función $F : [0, 1]^2 \rightarrow [0, 1]$ se dice \vec{r} -creciente ordenada direccional reforzada (SOD, del inglés *Strengthened Ordered Directional*) si para todo $(x_1, x_2) \in [0, 1]^2$, y para cada permutación $\sigma \in S$ tal que $x_{\sigma(1)} \geq x_{\sigma(2)}$ y para cualquier $c > 0$ que cumple $(x_{\sigma(1)} + cr_1, x_{\sigma(2)} + cr_2) \in [0, 1]^2$, se satisface la siguiente inecuación

$$F(x_1 + cr_{\sigma^{-1}(1)}, x_2 + cr_{\sigma^{-1}(2)}) \geq F(x_1, x_2),$$

Si una función F es \vec{r} -creciente SOD, entonces F es también \vec{r} -creciente OD. El contrario no se cumple en general. Sin embargo, si el vector \vec{r} satisface $r_1 \geq r_2$, entonces ambas definiciones son equivalentes.

Nótese que por sencillez solo nos hemos referido a las condiciones de crecimiento, pero condiciones similares pueden ser introducidas para el decrecimiento de las funciones.

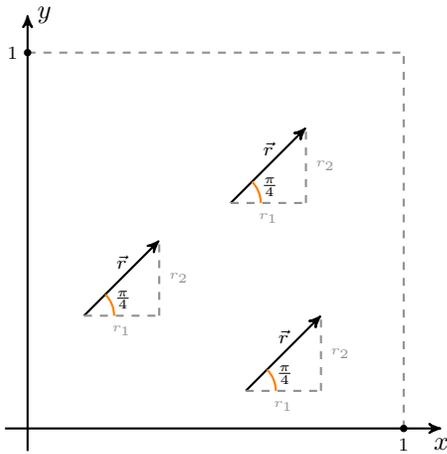


Fig. 1. Ejemplo de la dirección de una función débilmente creciente.

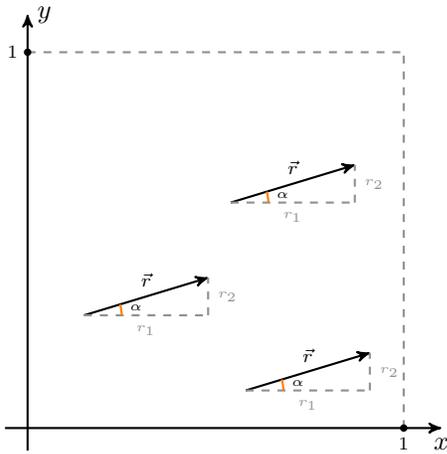


Fig. 2. Ejemplo de dirección de una función \vec{r} -creciente.

III. INTERPRETACIÓN GEOMÉTRICA DE LAS CONDICIONES DE MONOTONÍA EN DOS DIMENSIONES

Una función débilmente creciente puede entenderse como una función que en cada punto del plano crece en la dirección $(1,1)$, es decir, crece a través de la recta que pasa por este punto en la dirección del vector $(1,1)$ (véase Fig. 1). De manera similar, un función $\vec{r} = (r_1, r_2)$ -creciente puede entenderse como una función que en cada punto del plano crece en la dirección (r_1, r_2) , es decir, crece a través de la recta que pasa por ese punto en la dirección del vector (r_1, r_2) (véase Fig. 2).

Tanto en el caso de las funciones débilmente crecientes como en el de las funciones \vec{r} -crecientes para un vector fijo \vec{r} , la dirección de la recta sobre la que crece está prefijada y es igual para todos los puntos del dominio. Por el contrario, la mayor diferencia entre las funciones anteriores y las funciones \vec{r} -crecientes OD y SOD se debe a que la dirección varía de una región a otra. En particular, si $(x_1, x_2) \in [0, 1]^2$ y $\sigma \in \mathcal{S}_n$ es la permutación que cumple $x_{\sigma(1)} \geq x_{\sigma(2)}$, entonces la reordenación de las componentes del vector $\vec{r}_{\sigma^{-1}}$ en las

Definiciones 2.4 y 2.5 significa que la primera componente del vector, r_1 , afecta al mayor valor de x , mientras que r_2 afecta al menor de los valores de x . Esto significa que diferentes direcciones son consideradas dependiendo de la permutación que ordene las componentes del punto x . Formalicemos esta idea.

Dado que estamos trabajando únicamente en el caso 2-dimensional, dado un punto $x = (x_1, x_2) \in [0, 1]^2$ se satisface que $x_1 \leq x_2$ o $x_2 \leq x_1$. Esto significa que podemos diferenciar dos regiones de direcciones distintas.

Denotemos Ω_1 y Ω_2 a los subespacios

$$\Omega_1 = \{(x_1, x_2) \in [0, 1]^2 \mid x_1 \geq x_2\} \quad (2)$$

and

$$\Omega_2 = \{(x_1, x_2) \in [0, 1]^2 \mid x_1 \leq x_2\}. \quad (3)$$

Así mismo, denotamos por Δ a la diagonal del espacio, que está formada por todos los puntos del plano que pertenecen a los dos espacios, es decir

$$\Delta = \Omega_1 \cap \Omega_2 = \{(x_1, x_2) \in [0, 1]^2 \mid x_1 = x_2\}.$$

Proposición 3.1: Sea $\vec{r} = (r_1, r_2) \in \mathbb{R}^2$ un vector 2-dimensional distinto del nulo. La función $F: [0, 1]^2 \rightarrow [0, 1]$ es \vec{r} -creciente SOD si y solo si

- para todo $(x_1, x_2) \in \Omega_1$, F es (r_1, r_2) -creciente en (x_1, x_2) ;
- para todo $(x_1, x_2) \in \Omega_2$, F es (r_2, r_1) -creciente en (x_1, x_2) .

Además, también se conocen los ángulos (con respecto al eje horizontal positivo) de las direcciones de crecimiento.

Proposición 3.2: Sea $F: [0, 1]^2 \rightarrow [0, 1]$ una función \vec{r} -creciente SOD para un vector fijo $\vec{r} = (r_1, r_2) \in \mathbb{R}^2$ no nulo. Se cumple que

- para todo $(x_1, x_2) \in \Omega_1$, el ángulo de la recta (con respecto al eje horizontal positivo) a través de la cuál la función F crece en el punto (x_1, x_2) es el siguiente:
 - si $r_1 \geq 0$, entonces $\alpha = \arctan(\frac{r_2}{r_1})$ (véase nota a pie de página ¹).
 - si $r_1 < 0$, entonces $\alpha = \pi + \arctan(\frac{r_2}{r_1})$.
- para todo $(x_1, x_2) \in \Omega_2$, el ángulo de la recta (con respecto al eje horizontal positivo) a través de la cuál la función F crece en el punto (x_1, x_2) es el siguiente:
 - si $r_2 \geq 0$, entonces $\beta = \arctan(\frac{r_1}{r_2})$;
 - si $r_2 < 0$, entonces $\beta = \pi + \arctan(\frac{r_1}{r_2})$.

Además, es fácil comprobar que se cumple $\beta + \alpha = \frac{\pi}{2}$.

Demostración: Sea $(x_1, x_2) \in \Omega_1$. Esto significa que $x_1 \geq x_2$ y, por lo tanto, la permutación identidad σ_1 , dada por $\sigma_1(1) = 1$ y $\sigma_1(2) = 2$, satisface que $x_{\sigma_1(1)} \geq x_{\sigma_1(2)}$. Entonces, la dirección de crecimiento en el punto x es $\vec{r} = (r_1, r_2)$ y, por lo tanto, el ángulo con respecto al eje horizontal positivo es $\alpha = \arctan(\frac{r_2}{r_1})$ para todo vector \vec{r} tal que $r_1 \geq 0$ y $\alpha = \pi + \arctan(\frac{r_2}{r_1})$ para todo vector \vec{r} tal que $r_1 < 0$. De manera análoga, sea $(x_1, x_2) \in \Omega_2$. Esto

¹Por convenio consideramos $\arctan \frac{a}{0} = \text{sign}(a) \frac{\pi}{2}$.

significa que $x_2 \geq x_1$ y, por lo tanto, la permutación σ_2 , dada por $\sigma_2(1) = 2$ y $\sigma_2(2) = 1$, satisface que $x_{\sigma_2(1)} \geq x_{\sigma_2(2)}$. Entonces, la dirección de crecimiento en el punto x es $\vec{r} = (r_{\sigma_2^{-1}(1)}, r_{\sigma_2^{-1}(2)}) = (r_2, r_1)$ y, por lo tanto, el ángulo con respecto al eje horizontal positivo es $\beta = \arctan(\frac{r_1}{r_2})$ para todo vector \vec{r} tal que $r_2 \geq 0$ y $\beta = \pi + \arctan(\frac{r_1}{r_2})$ para todo vector \vec{r} tal que $r_2 < 0$.

La igualdad $\alpha + \beta = \frac{\pi}{2}$ se cumple por el hecho de que para $a \neq 0$, se cumple $\arctan(\frac{1}{a}) = \text{sign}(a)\frac{\pi}{2} - \arctan(a)$. Nótese que algunas veces, el valor de la suma real es $\frac{5\pi}{2}$, pero este es el mismo ángulo que $\frac{\pi}{2}$. ■

Nota 1: Nótese que debido a que los elementos de la diagonal Δ pertenecen tanto a Ω_1 como a Ω_2 ambas direcciones de crecimiento son impuestas.

Resultados similares pueden ser formulados para el caso de las funciones \vec{r} -crecientes OD en los que se impone el crecimiento solo al segmento formado de puntos de la recta que pertenecen al mismo conjunto Ω_i que el punto inicial x .

Corolario 3.3: Sea $\vec{0} \neq \vec{r} \in \mathbb{R}^2$. Una función $F: [0, 1]^2 \rightarrow [0, 1]$ es \vec{r} -creciente OD si y solo si

- para todo $(x_1, x_2) \in \Omega_1$, F es (r_1, r_2) -creciente en (x_1, x_2) restringido al subconjunto Ω_1 ;
- para todo $(x_1, x_2) \in \Omega_2$, F es (r_2, r_1) -creciente en (x_1, x_2) restringido al subconjunto Ω_2 .

Además, los ángulos de las rectas de crecimiento son los mismos que el caso de las funciones \vec{r} -crecientes SOD.

En el trabajo [16] puede encontrarse el siguiente resultado sobre la monotonía \vec{r} -creciente SOD.

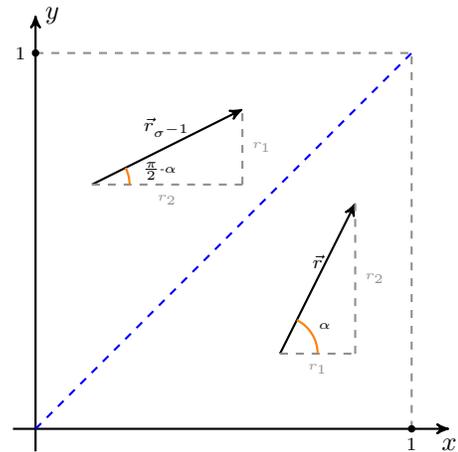
Teorema 3.4: (Teorema 6.5 de [16]) Sean $\vec{0} \neq \vec{r}, \vec{s} \in \mathbb{R}^n$ y $a, b \in \mathbb{R}^+$. Supongamos que $x = (x_1, \dots, x_n) \in [0, 1]^n$, $c > 0$ y σ es una permutación que ordena los elementos de x . Si $x_{\sigma(1)} \geq \dots \geq x_{\sigma(n)}$ y $x_{\sigma} + c(a\vec{r} + b\vec{s}) \in [0, 1]^n$, entonces o $x_{\sigma} + ca\vec{r} \in [0, 1]^n$ o $x_{\sigma} + cb\vec{s} \in [0, 1]^n$. Por lo tanto, si una función es \vec{r} - y \vec{s} -creciente SOD, entonces es $(a\vec{r} + b\vec{s})$ -creciente SOD.

Nótese que lo mismo puede probarse para funciones OD crecientes [11]. Estudiemos ahora el caso particular de los elementos de la diagonal. Recordemos, que como se ha mencionado anteriormente en la Nota 1, los elementos $\bar{x} = (x, x) \in \Delta$ satisfacen que $\bar{x} \in \Omega_1$ y $\bar{x} \in \Omega_2$, simultáneamente. Esto significa que los elementos de la diagonal en las funciones \vec{r} -crecientes SOD crecen en dos direcciones distintas.

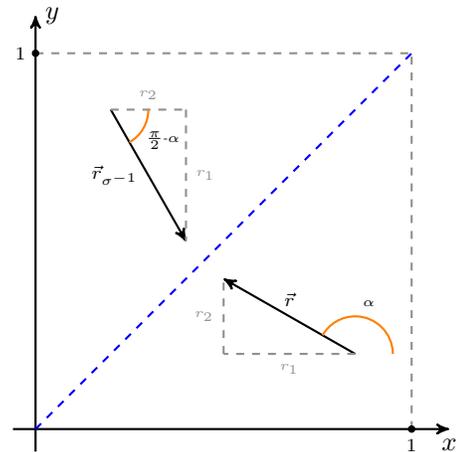
Corolario 3.5: Sea $\vec{0} \neq \vec{r} = (r_1, r_2) \in \mathbb{R}^2$ y sea F una función \vec{r} -creciente SOD. Para cada elemento de la diagonal $\bar{x} \in \Delta$, se tiene que F es (r_1, r_2) - y (r_2, r_1) -creciente SOD en \bar{x} .

Sea F una función $\vec{r} = (r_1, r_2)$ -creciente SOD para un vector \vec{r} satisfaciendo $r_1 \neq -r_2$. Por el Corolario 3.5, se tiene que F es una función (r_1, r_2) - y (r_2, r_1) -creciente en \bar{x} , para todo elemento $\bar{x} \in \Delta$. Entonces, debido al Teorema 3.4, F es creciente en \bar{x} a través de cualquier dirección entre (r_1, r_2) y (r_2, r_1) (véase Fig. 4(a)).

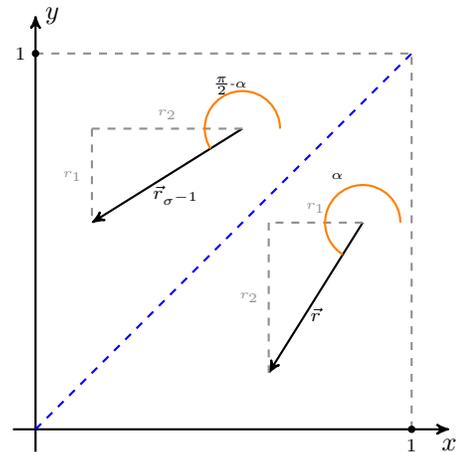
Dos casos extremos de direcciones son los vectores $\vec{r} = (-r, r)$ cuyo ángulo es $\alpha = \frac{3\pi}{4}$ y $\vec{r} = (r, -r)$ cuyo ángulo es $\alpha = \frac{7\pi}{8}$ (o $-\frac{\pi}{4}$).



a) Vector \vec{r} con $r_1 > 0$ y $r_2 > 0$.



b) Vector \vec{r} con $r_1 < 0$ y $r_2 > 0$.



c) Vector \vec{r} con $r_1 < 0$ y $r_2 < 0$.

Fig. 3. Ejemplos de algunas direcciones (r_1, r_2) para funciones crecientes OD y SOD: caso a) $r_1 > 0$ y $r_2 > 0$; b) $r_1 > 0$ y $r_2 < 0$ y c) $r_1 < 0$ y $r_2 < 0$.



Proposición 3.6: Sea $r > 0$ y F una función $(r, -r)$ -creciente SOD. Para todo $(x_1, x_2) \in [0, 1]^2$, se cumple que $F(x_m, x_m) \leq F(x_1, x_2)$, donde $x_m = \frac{x_1 + x_2}{2}$.

Demostración: Sea F una función $(r, -r)$ -creciente SOD, $(x_1, x_2) \in [0, 1]^2$ y $x_m = \frac{x_1 + x_2}{2}$. Como consecuencia del Corolario 3.5, F es $(r, -r)$ y $(-r, r)$ -creciente en $\bar{x}_m = (x_m, x_m) \in \Delta$. Distinguimos dos posibles casos dependiendo del punto $x = (x_1, x_2)$.

- Caso 1: Sea $(x_1, x_2) \in \Omega_1$. Dado $c = \frac{x_1 - x_2}{2r} \geq 0$, se cumple que $(x_m, x_m) + c(r, -r) = (x_1, x_2)$ y, por lo tanto, $F(x_m, x_m) \leq F(x_1, x_2)$.
- Caso 2: Sea $(x_1, x_2) \in \Omega_2$. Dado $c = \frac{x_2 - x_1}{2r} \geq 0$, se cumple que $(x_m, x_m) + c(-r, r) = (x_1, x_2)$ y, por lo tanto, $F(x_m, x_m) \leq F(x_1, x_2)$.

Como consecuencia de la Proposición anterior se puede concluir que si una función es $(r, -r)$ -increasing SOD, para algún $r > 0$, se cumple que $F(\bar{x}) \leq F(y_1, y_2)$ para todo punto satisfaciendo $y_1 + y_2 = 2x$, es decir, el valor en el punto de la diagonal es mínimo con respecto a todos los demás puntos de la recta que pasa por el punto con dirección $(-r, r)$ (véanse líneas verdes en la Fig. 4(b)).

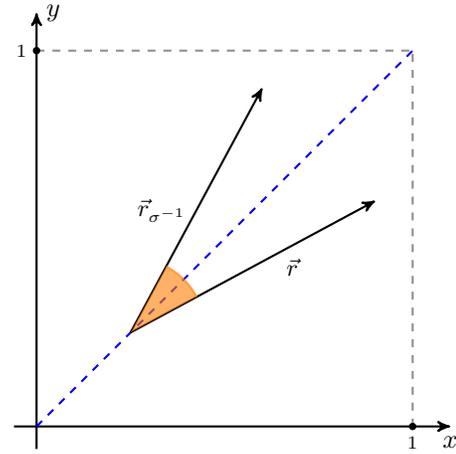
Proposición 3.7: Sea $r > 0$ y F una función $(-r, r)$ -creciente SOD. Para todo $(x_1, x_2) \in [0, 1]^2$, se cumple que $F(x_m, x_m) = F(x_1, x_2)$, donde $x_m = \frac{x_1 + x_2}{2}$.

Demostración: Sea F una función $(-r, r)$ -creciente SOD, $(x_1, x_2) \in [0, 1]^2$ y $x_m = \frac{x_1 + x_2}{2}$. Como consecuencia del Corolario 3.5, F es $(-r, r)$ y $(r, -r)$ -creciente en $\bar{x}_m = (x_m, x_m) \in \Delta$. Distinguimos dos posibles casos dependiendo del punto $x = (x_1, x_2)$.

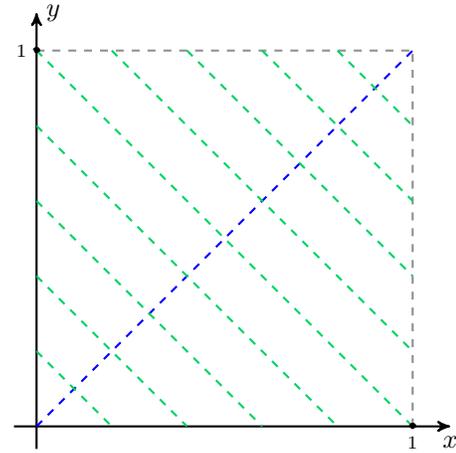
- Caso 1: Sea $(x_1, x_2) \in \Omega_1$. Dado $c = \frac{x_1 - x_2}{2r} \geq 0$, se cumple $(x_1, x_2) + c(-r, r) = (x_m, x_m)$. Como F es $(-r, r)$ -creciente en (x_1, x_2) , se tiene que $F(x_1, x_2) \leq F(x_m, x_m)$. De manera similar, se tiene que $(x_m, x_m) + c(r, -r) = (x_1, x_2)$. Como F es también $(r, -r)$ -creciente en (x_m, x_m) , se cumple que $F(x_m, x_m) \leq F(x_1, x_2)$. Por lo tanto, $F(x_1, x_2) = F(x_m, x_m)$.
- Caso 2: Sea $(x_1, x_2) \in \Omega_2$. Dado $c = \frac{x_2 - x_1}{2r} \geq 0$, se cumple que $(x_1, x_2) + c(r, -r) = (x_m, x_m)$. Como F es $(r, -r)$ -creciente en (x_1, x_2) (nótese que en Ω_2 , $\sigma_2(1) = 2$ y $\sigma_2(2) = 1$), y se tiene que $F(x_1, x_2) \leq F(x_m, x_m)$. De manera similar, se tiene que $(x_m, x_m) + c(-r, r) = (x_1, x_2)$. Como F es también $(-r, r)$ -creciente en (x_m, x_m) , se cumple que $F(x_m, x_m) \leq F(x_1, x_2)$. Por lo tanto, $F(x_1, x_2) = F(x_m, x_m)$.

IV. CONCLUSIONES

En la literatura reciente encontramos diversas nociones de monotonía menos restrictivas que la monotonía en cada variable que se impone a la funciones de agregación. En este trabajo, hemos analizado la interpretación geométrica de estas monotonías. Se ha estudiado la diferencia entre monotonía direccional ordenada y monotonía direccional ordenada reforzada, viendo que la principal diferencia es que



a) Direcciones de crecimiento en los puntos de la diagonal de una función \vec{r} -creciente SOD.



b) Líneas paralelas con dirección $(-r, r)$.

Fig. 4. Representación gráfica del comportamiento de las funciones $\vec{r} = (r_1, r_2)$ -creciente SOD con $r_1 = -r_2$.

la monotonía direccional ordenada restringe sus condiciones de monotonía al subespacio Ω en el que se encuentra. El análisis geométrico llevado a cabo para el caso 2-dimensional, también nos ha permitido deducir algunas consecuencias del comportamiento de las funciones direccionalmente ordenadas reforzadas cuando el vector de las direcciones $\vec{r} = (r_1, r_2)$ satisface que $r_1 = -r_2$.

Como trabajo a futuro, destaca la posibilidad de realizar un estudio más general de la interpretación geométrica para el caso n -dimensional. En este caso, se perdería la posibilidad de dibujar el dominio en el plano, pero existe una mayor cantidad de permutaciones posibles, que darían un mayor número de regiones donde las direcciones varían.

REFERENCIAS

[1] G. Beliakov, S. James, and D. G. Nimmo, "Using aggregation functions to model human judgements of species diversity," *Information Sciences*, vol. 306, pp. 21–33, 2015.

[2] U. Bentkowska, "Aggregation of diverse types of fuzzy orders for decision making problems," *Information Sciences*, vol. 424, pp. 317–336, 2018.

- [3] D. Paternain, H. Bustince, M. Pagola, P. Sussner, A. Kolesárová, and R. Mesiar, “Capacities and overlap indexes with an application in fuzzy rule-based classification systems,” *Fuzzy Sets and Systems*, vol. 305, pp. 70–94, 2016.
- [4] J. L. García-Lapresta and R. A. M. Pereira, “The dual decomposition of aggregation functions and its application in welfare economics,” *Fuzzy Sets and Systems*, vol. 281, pp. 188–197, 2015.
- [5] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, “Aggregation functions: Construction methods, conjunctive, disjunctive and mixed classes,” *Information Sciences*, vol. 181, no. 1, pp. 23–43, 2011.
- [6] E. P. Klement, R. Mesiar, and E. Pap, *Triangular Norms*. Springer, Dordrecht, 2000.
- [7] J. Fodor and B. D. Baets, *Studies in Fuzziness and Soft Computing, vol 215*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, ch. Uninorm Basics, pp. 49–64.
- [8] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, and R. Orduna, “Overlap functions,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 72, no. 3-4, pp. 1488–1499, 2010.
- [9] T. Wilkin and G. Beliakov, “Weakly monotonic averaging functions,” *International Journal of Intelligent Systems*, vol. 30, no. 2, pp. 144–169, 2015.
- [10] H. Bustince, J. Fernandez, A. Kolesárová, and R. Mesiar, “Directional monotonicity of fusion functions,” *European Journal of Operational Research*, vol. 244, no. 1, pp. 300–308, 2015.
- [11] H. Bustince, E. Barrenechea, M. Sesma-Sara, J. Lafuente, G. Dimuro, R. Mesiar, and A. Kolesárová, “Ordered directionally monotone functions. Justification and application,” *IEEE Transactions on Fuzzy Systems (In Press)*.
- [12] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, and H. Bustince, “Preaggregation functions: Construction and an application,” *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 2, pp. 260–272, 2016.
- [13] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, H. Bustince, and R. Mesiar, “CF-integrals: a new family of pre-aggregation functions with application to fuzzy rule-based classification systems,” *Information Sciences*, vol. 435, pp. 94–110, 2017.
- [14] M. Sesma-Sara, H. Bustince, E. Barrenechea, J. Lafuente, A. Kolesárová, and R. Mesiar, “Edge detection based on ordered directionally monotone functions,” in *Advances in Fuzzy Logic and Technology 2017*. Springer, 2017, pp. 301–307.
- [15] C. Marco-Detchart, C. Lopez-Molina, J. Fernandez, M. Pagola, and H. Bustince, *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations. IPMU 2018*. Springer, 2018, ch. Image Feature Extraction Using OD-Monotone Functions.
- [16] M. Sesma-Sara, J. Lafuente, A. F. Roldán-López-de-Hierro, R. Mesiar, and H. Bustince, “Strengthened ordered directionally monotone functions. Links between the different notions of monotonicity,” *Submitted to Fuzzy Sets and Systems*.
- [17] M. Grabisch, J. Marichal, R. Mesiar, and E. Pap, *Aggregation functions*. Cambridge University Press, 2009.
- [18] G. Beliakov, H. Bustince, and T. Calvo, *A Practical Guide to Averaging Functions*, ser. Studies in Fuzziness and Soft Computing. Springer International Publishing, 2016.



Operadores de agregación y comparación para diferenciación de señales discretas

C. Lopez-Molina, S. Iglesias-Rey, H. Bustince
Dpto. Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, Spain
Email: carlos.lopez@unavarra.es

N. Madrid
Escuela de Ingenierías Industriales
Universidad de Málaga
Málaga, España
Email: nicolas.madrid@uma.es

Abstract—La tecnología digital ha hecho que la mayor parte de los datos se registren en forma de señales discretas. A pesar de las ventajas que esto tiene, la naturaleza de esta información representa desafíos novedosos. En este trabajo analizamos la diferenciación de señales discretas, una tarea básica para el procesamiento automático de las mismas. En concreto, proponemos el uso de operadores de comparación y agregación para la generación de familias de operadores de diferenciación con alta interpretabilidad. Nuestra propuesta es validada en el contexto de procesamiento de imagen, con la finalidad de caracterizar derivadas parciales en señales bidimensionales.

Index Terms—Operadores de agregación; operadores de comparación; diferenciación de señales discretas; filtros de convolución.

I. INTRODUCCIÓN

La era digital ha producido un incremento exponencial en la cantidad de datos recogidos en casi cualquier ámbito de conocimiento. Esto ha provocado cambios en la manera de tratar información, siendo uno de los más importantes el papel del procesamiento de señal [1], [2]. Nótese que el procesamiento de señal trata con datos nativamente representados en forma de señal (audio, imagen, video), y también de datos escalares recogidos en forma de señal (como las series temporales o los espacios-escala). El procesamiento de señal, entendido de una manera amplia, es una rama aplicada del análisis matemático. Sin embargo, la portabilidad de algoritmos entre ambos campos está en ocasiones muy dificultada por la naturaleza discreta de las señales digitales.

En este trabajo nos centramos en el problema de la diferenciación de señales discretas [3]. La importancia de la diferenciación (e integración) en el procesamiento de señal es capital [4]. Ambas operaciones, aunque en especial la primera de ellas, se usan para localizar tanto puntos críticos (p.e. máximos o mínimos locales) como para localizar o *clusterizar* regiones de interés en las señales. Un ejemplo de fácil interpretación es el de las imágenes en escala de grises. En estas imágenes, entendidas como señales discretas bidimensionales, el uso de derivadas parciales es central a casi cualquier procesamiento de medio o alto nivel. Las diferencias parciales permiten analizar la variación local de las imágenes, y suponen la información primaria en aplicaciones que incluyen la segmentación, detección de bordes, la búsqueda de puntos de interés y la regularización adaptativa.

Muchas de las propuestas para diferenciación de señales discretas están basadas en filtros de convolución para primera o segunda derivada. Al inicio de la era digital aparecieron filtros de convolución discretos, generalmente asociados al cómputo de gradientes en imágenes digitales [5], [6], [7]. En los años 80, nuevos trabajos se inspiraron en el análisis matemático para diseñar filtros continuos, que debían ser discretizados para su aplicación. Esta naturaleza continua les permitía, antes de la discretización, presentarse como óptimos ante ciertos criterios de calidad [8], [9]. Estos trabajos incluyen las propuestas de Marr y Hildreth [10] y Canny [11], [12], así como ciertas expansiones posteriores [9], [13].

El estudio de filtros de convolución para el cálculo de diferencias parciales se ha visto frenado en los últimos 20 años. A pesar de que se han presentado trabajos de impacto, estos generalmente se limitan a reescribir los criterios de optimalidad y derivar nuevas propuestas, sin que esto repercuta claramente en una mejora en la aplicabilidad [14], [15], [16]. Si bien es cierto que la aparición de los espacios-escala para el análisis multidimensional de señales ha aportado cierta novedad en la diferenciación de señales discretas [17], [18], apenas ha sido aprovechado para producir mejores análisis diferenciales [19].

En este trabajo analizamos los filtros de convolución para el cálculo de diferencias parciales, con especial atención a la diferenciación parcial de imágenes. En concreto, reformulamos el proceso de convolución para considerar, de manera independiente, (a) la cuantificación de diferencias locales y (b) el modelado del vecindario de interés en torno a cada punto de la señal. Esta reformulación viene inspirada por los filtros bilaterales de Tomasi y Manduchi [20], así como por otras iniciativas de aplicación de operadores de agregación en la caracterización de diferencias parciales en imágenes [21]. Proponemos esta generalización para poder incorporar a este campo diferentes operadores de la Teoría de Conjuntos Difusos, en particular las funciones de disimilitud restringidas y las funciones de agregación.

En la Sección II repasamos algunos conocimientos básicos de aplicación en este trabajo. Las Secciones III y III desgranar la propuesta, que es validada de manera experimental en la Sección V. Para finalizar, la Sección VI lista algunas conclusiones y futuras líneas de trabajo.

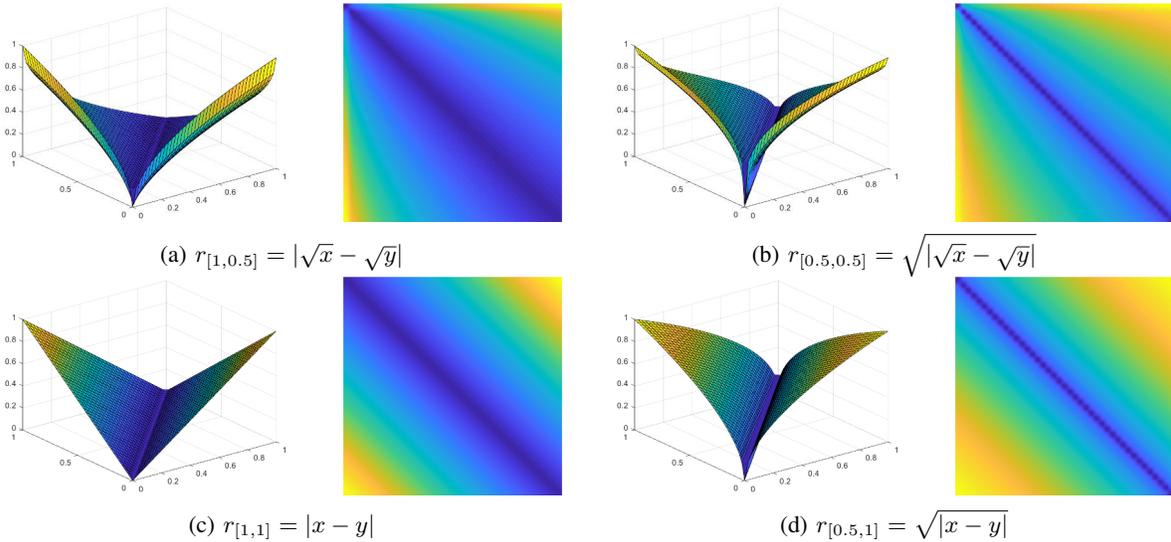


Fig. 1. Funciones de disimilitud restringidas construidas de acuerdo a la Proposición 1.

II. PRELIMINARIES

La teoría de funciones de comparación ha tenido un impacto significativo en la Teoría de Conjuntos Difusos. Esto se debe, probablemente, a la naturaleza subjetiva de las comparaciones humanas, en la cuales pueden aparecer ambigüedades e inconsistencias. En este contexto, la Teoría de Conjuntos Difusos ha sido prolífica, bien fuera centrada en términos de métricas [22] o en términos de medidas de similitud [23]. En este trabajo usamos las Funciones de Disimilitud Restringidas (o, RDFs, por sus siglas inglesas [24]). Estas funciones modelan la diferencia perceptual entre dos elementos en el intervalo $[0, 1]$, y han sido aplicadas en diferentes contextos [25], [26].

Definition 1: [24] Una *función de disimilitud restringida* es una aplicación $r : [0, 1]^2 \rightarrow [0, 1]$ que satisfaga

- (R1) $r(x, y) = r(y, x)$ para todo $x, y \in [0, 1]$;
- (R2) $r(x, y) = 1$ si y sólo si $x = 0$ e $y = 1$ o $x = 1$ e $y = 0$;
- (R3) $r(x, y) = 0$ si y sólo si $x = y$;
- (R4) para todo $x, y, z \in [0, 1]$, si $x \leq y \leq z$, entonces $r(x, y) \leq r(x, z)$ y $r(y, z) \leq r(x, z)$.

Existen diferentes maneras de construir RDFs, si bien la más sencilla es la combinación de automorfismos.

Proposition 1: Sean φ_1 y φ_2 dos automorfismos de $[0, 1]$. La aplicación $r : [0, 1]^2 \rightarrow [0, 1]$ construída tal que

$$r(x, y) = \varphi_1(|\varphi_2(x) - \varphi_2(y)|) \quad (1)$$

es una función de disimilitud restringida (o RDF).

De cara a reducir el alcance de este trabajo, consideraremos únicamente las RDFs construídas de acuerdo a la Proposición 1. Específicamente, nos referiremos como $r_{[\alpha, \beta]}$, con $\alpha, \beta \in \mathbb{R}$, a la RDF tal que $\varphi_1(x) = x^\alpha$ y $\varphi_2(x) = x^\beta$. Es decir, a

$$r_{[\alpha, \beta]}(x, y) = (|x^\beta - y^\beta|)^\alpha. \quad (2)$$

La Fig. 1 muestra algunas RDFs construídas de acuerdo a la Ec. (2). Nótese que $r_{[1,1]}$ es equivalente a la diferencia absoluta entre los argumentos de la función.

III. UN MODELO GENERAL PARA DIFERENCIACIÓN

La noción de *derivada* está ligada a la idea de variación de una función. Por ello, al trabajar con funciones sobre dominios discretos, la diferencia local puede tomarse como un reemplazo fidedigno de la derivada. Sean dos funciones $f : \mathbb{R}^2 \mapsto \mathbb{R}$ y $g : \mathbb{N}^2 \mapsto \mathbb{R}$ tal que la segunda (mediciones) se construya a partir del muestreo de la primera (realidad subyacente). Podría aproximarse la derivada de f tal que:

$$\begin{aligned} \frac{\partial f}{\partial x}(i, j) &= g(i+1, j) - f(i, j) \text{ y} \\ \frac{\partial f}{\partial y}(i, j) &= g(i, j+1) - f(i, j). \end{aligned} \quad (3)$$

Sin embargo, la aplicación de esta estrategia suele generar resultados decepcionantes, en buena medida debido a la contaminación de los datos. En entornos reales, el empleo de Ec. (3) es poco fiable, y sus resultados tienen una interpretabilidad muy baja. Un ejemplo de este hecho se ve en la Fig. 2. En esta figura vemos, en la fila superior, dos señales sintéticas, tal que la situada en la columna derecha es una versión contaminada de la que ocupa la columna izquierda. En esta señal coexisten dos *objetos* cada uno de los cuales genera dos puntos de interés (críticos), uno a cada lado del propio objeto. Estos puntos deberían representarse como dos máximos (subidas, al lado izquierdo de cada objeto) y dos mínimos (bajadas, al lado derecho). En la segunda fila de la Fig. 2 vemos la aproximación a $\frac{\partial f}{\partial x}$ realizada por el filtro de Lalignant [15], que corresponde en su versión 1D a la formulación de la Ec. (3). Se observa que el resultado es muy preciso en la señal sintética original, pero también muy poco interpretable en la señal contaminada. Esto se achaca a la alta frecuencia del muestreo para computar la diferencia (dos posiciones contiguas), lo que generalmente se soluciona usando operadores de mayor soporte.

Al usar un filtro de diferenciación la aproximación a las derivadas parciales de f se hace mediante la convolución de

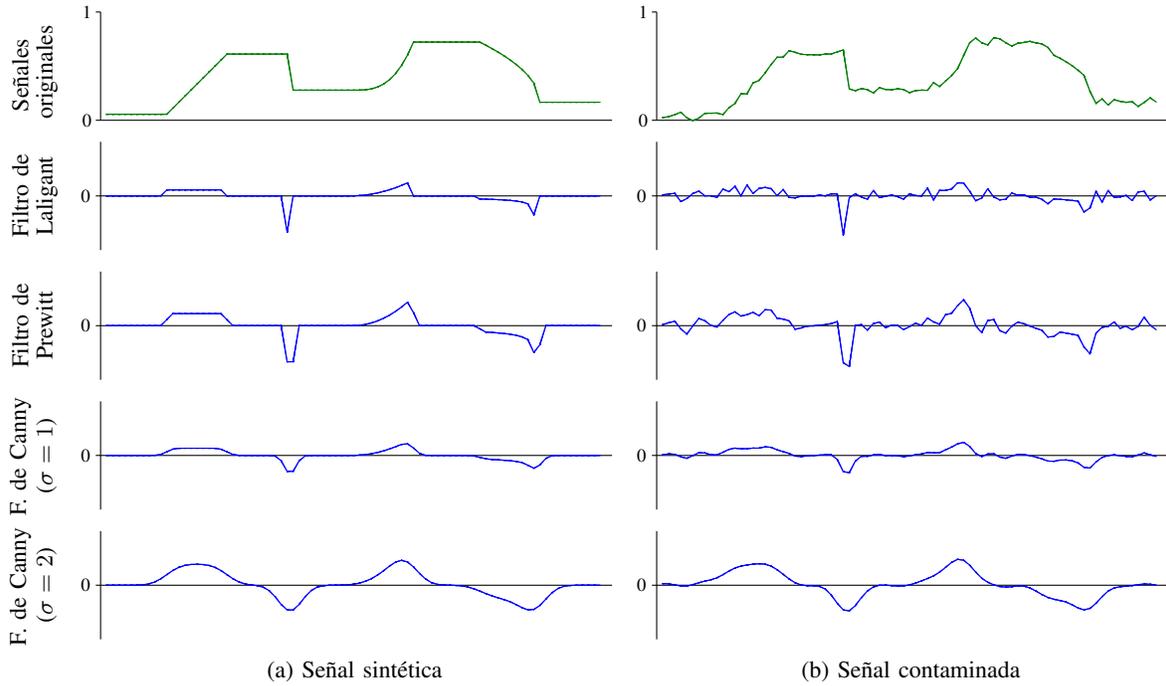


Fig. 2. Estimaciones de la primera derivada de una señal discreta usando diferentes filtros de convolución 1D. La señal contaminada se obtiene añadiendo ruido blanco Gaussiano con desviación estándar 0.02. Los operadores usados son el de Laligant [15], Prewitt [6] and Canny [12], y pueden observarse en la Fig. 3.

la señal g con un filtro κ . Algunos ejemplos son los operadores discretos de Sobel [5] y Prewitt [6], o los operadores basados en señales continuas de Canny [11] (y sus evoluciones [16], [27]) o Shen-Castan [28]. Algunos de estos filtros, en su representación 1D, se incluyen en la Fig. 3.

Una derivada parcial en x en el punto \mathbf{p} se aproxima tal que¹:

$$\frac{\partial f}{\partial x}(\mathbf{p}) = g * \kappa(\mathbf{p}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \kappa(\xi) \cdot f(\mathbf{p} + \xi) \cdot d\xi, \quad (4)$$

donde ξ representa el desplazamiento.

La elección de κ , así como los parámetros que éste pudiera tener, ha sido ampliamente discutido en la comunidad científica. En esta discusión toman relevancia las características de las funciones a derivar, las contaminaciones que pueda sufrir, o la finalidad del proceso. De hecho, muchas propuestas son óptimas para el conjunto de restricciones y objetivos propuestos por sus autores, si bien estos difieren unos de otros. Una de las propiedades más aceptada es la antisimetría o imparidad, es decir, que para cualquier ξ en el dominio se cumple que $\kappa(\xi) = -\kappa(-\xi)$. Esta propiedad no siempre se cumple en los filtros de diferenciación, véase el caso del filtro de Laligant [15] en la Fig. 3, pero sí es sostenida en la mayor parte de los casos. Nótese que, en algunos filtros (como el ISEF [28], ver Fig. 3) el filtro no está definido en el origen, suponiéndose un valor 0.

¹En adelante, se considera únicamente la derivada $\frac{\partial f}{\partial x}$, ya que la derivada $\frac{\partial f}{\partial y}$ se computa de manera análoga.

Partiendo de la Ec. (4), y dando por hecha la antisimetría del filtro κ , tenemos que:

$$\begin{aligned} g * \kappa(\mathbf{p}) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \kappa(\xi) \cdot g(\mathbf{p} + \xi) \cdot d\xi = \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} \kappa(\xi) g(\mathbf{p} + \xi) d\xi + \int_{-\infty}^{\infty} \int_{-\infty}^0 \kappa(\xi) g(\mathbf{p} + \xi) d\xi = \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} \kappa(\xi) g(\mathbf{p} + \xi) d\xi + \int_{-\infty}^{\infty} \int_0^{\infty} \kappa(-\xi) g(\mathbf{p} - \xi) d\xi = \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} \kappa(\xi) g(\mathbf{p} + \xi) d\xi - \int_{-\infty}^{\infty} \int_0^{\infty} \kappa(\xi) g(\mathbf{p} - \xi) d\xi = \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} \kappa(\xi) (g(\mathbf{p} + \xi) - g(\mathbf{p} - \xi)) d\xi. \end{aligned} \quad (5)$$

Los dominios de las señales se definen, en el plano teórico, con dominio infinito. Por ejemplo, un filtro Gaussiano es, evidentemente, positivo en $[-\infty, \infty]$, que pasa a ser el rango de las integrales en Ec. (5). Sin embargo, en las aplicaciones prácticas esto no tiene demasiada relevancia, y los filtros se restringen a un cierto soporte discreto a partir del cual se considera que $K(\xi) = 0$. Por tanto, se puede reformular el resultado de la Ec. (5) tal que:

$$g * \kappa(i) = \int_{-t}^t \int_0^t \kappa(\xi) \cdot (g(\mathbf{p} + \xi) - g(\mathbf{p} - \xi)) \cdot d\xi. \quad (6)$$

donde $[-t, t]^2$ es el soporte del filtro κ , generalmente establecido a partir del descenso del valor κ por debajo de un cierto umbral cercano a 0.

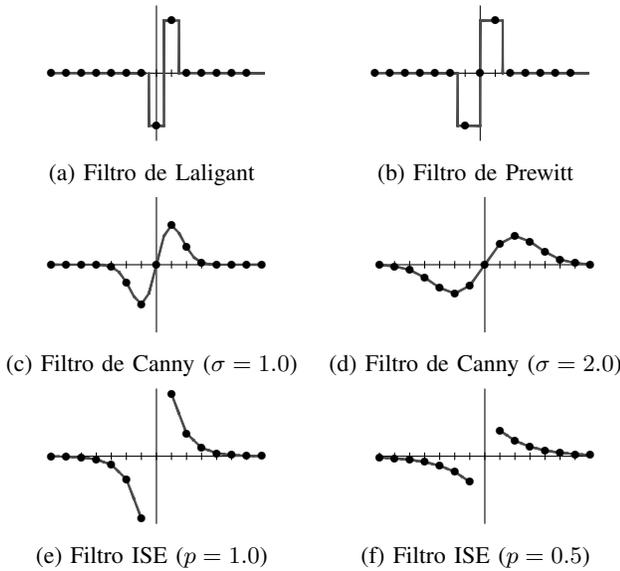


Fig. 3. Filtros de convolución 1D usados para diferenciación. Los puntos representan la versión discreta, mientras que la línea señala la función continua subyacente.

La estimación de las derivadas parciales de una señal g involucra, pues, dos componentes diferentes:

- la cuantificación de las diferencias entre los valores a ambos lados de la posición, representado en la Ec. (6) por el término $(g(x + \xi) - g(x - \xi))$. Nos referimos a este término como *término de intensidad*;
- la cuantificación de la relevancia de cada posición en el entorno, representado en la Ec. (6) por el término $\kappa(\xi)$. Nos referimos a este término como *término espacial*.

Al tener ambos términos identificados, podemos modelarlos de manera independiente. Por ejemplo, en lugar de considerar la diferencia absoluta entre los valores de la señal g , podemos modelar el término de intensidad usando otras medidas. Dado que las señales discretas suelen tener límites bien definidos, podemos considerar que f y g tienen el rango $[0, 1]$. Para el *término de intensidad* de la Ec. (6) consideramos una función de diferencia tal que $\epsilon: [0, 1]^2 \rightarrow [-1, 1]$. Esta función debe encargarse de cuantificar la diferencia percibida entre dos valores, sean $g(\mathbf{p} - \xi)$ y $g(\mathbf{p} + \xi)$. La Ec. (6) resulta en:

$$g * \kappa(\mathbf{p}) = \int_{-k}^k \int_0^k \kappa(\xi) \cdot \epsilon((g(\mathbf{p} + \xi), g(\mathbf{p} - \xi))) \cdot d\xi \quad (7)$$

Así, las derivadas parciales dependen de dos piezas de información: el filtro que modela la relevancia espacial (κ) y la función que modula la disimilitud (ϵ). Cada uno de estos debería poder estudiarse y entrenarse de manera independiente. Además, ambos se combinan de manera local usando un operador conjuntivo (producto) y de manera semi-local usando la suma (media aritmética, en caso de normalización).

En este trabajo nos centramos en el término de intensidad, relacionándolo con la teoría de comparación. En concreto, estudiamos el papel que pueden tener las funciones de disimilitud restringidas para el modelado de las diferencias.

IV. DIFERENCIACIÓN USANDO FUNCIONES DE DISIMILITUD RESTRINGIDAS

Dada la Ec. (7), una pregunta evidente se refiere al conjunto mínimo de propiedades que debe cumplir la función ϵ . Por ejemplo, debería producir valores *altos* cuando sus argumentos sean diferentes, mientras que deberían ser bajos o cercanos a cero cuando los argumentos sean similares. Sería esperable también la antisimetría, ya que las derivadas de una función no deberían verse afectadas por la dirección del análisis (más allá del signo de las mismas). Creemos que una función ϵ debe cumplir las siguientes propiedades:

- (E1) $\epsilon_r(y, z) = -\epsilon_r(z, y)$;
- (E2) $\epsilon_r(y, z) = 0$ si y sólo si $y = z$;
- (E3a) $\epsilon_r(y, z) = 1$ si y sólo si $y = 0$ y $z = 1$;
- (E3b) $\epsilon_r(y, z) = -1$ si y sólo si $y = 1$ y $z = 0$;
- (E4) Para cada $t, y, z \in [0, 1]$, $t \leq y \leq z$, tenemos que $\epsilon_r(t, y) \leq \epsilon_r(t, z)$ y $\epsilon_r(y, z) \leq \epsilon_r(z, t)$.

Estas propiedades garantizan el comportamiento esperado en el término de intensidad. Por ejemplo, (E2) garantiza que la derivada de una función será 0 si y sólo si $g(k + \xi) = g(k - \xi)$ para cada posición $\xi \in [-t, t]^2$.

Tras un análisis detallado, creemos que las RDF suponen una herramienta útil para generar funciones ϵ . En concreto, proponemos usar funciones ϵ construidas tal que:

$$\epsilon_r(y, z) = \frac{z - y}{|z - y|} \cdot r(y, z), \quad (8)$$

donde r es una RDF. Una función ϵ_r , construida como en la Ec. (8), satisface las propiedades (E1)-(E4).

Trabajos anteriores han estudiado las técnicas clásicas de filtrado para incorporar dos términos independientes (y explícitos), especialmente en el campo de procesamiento de imagen. Las técnicas de regularización o diferenciación adaptadas al contenido local de la imagen (*content-aware filtering*), por ejemplo, discriminan el análisis de tonos (el término de intensidad) con el estudio de la posición de cada pixel (el término espacial). La manera en que los términos se aplican y combinan puede, sin embargo, ser diferente. La contribución más relevante en este sentido, así como la más parecida a nuestro objetivo, es el *filtrado bilateral* de Tomasi y Manduchi (*bilateral filtering* [20]). En este trabajo, los autores proponen la combinación de filtros en las dimensiones tonal y espacial para la regularización adaptativa de imágenes. Estos filtros consideran la cercanía tonal y espacial de cada par de píxeles a la hora de regularizar una zona de la imagen. Específicamente, proponen generar una imagen regularizada g^* a partir de una imagen monocal g tal que:

$$g^*(\mathbf{p}) = \frac{1}{k_d} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\xi) \cdot g(\mathbf{p} + \xi) \cdot h(|g(\mathbf{p}) - g(\mathbf{p} + \xi)|) d\xi \quad (9)$$

donde k_d es un factor de normalización, s es un filtro de paso bajo y h es una función unaria decreciente. En esta formulación, s modula el peso de los elementos en función de su distancia a \mathbf{p} , mientras que h modula el peso de los elementos en función de su diferencia de valor respecto a $g(\mathbf{p})$.



Los filtros bilaterales se han aplicado en diferentes escenarios, si bien se han mantenido casi siempre en el ámbito de la regularización adaptativa. Los objetivos, sin embargo, han variado, desde la regularización sin contexto (en el trabajo original [20]) a la eliminación de ruido o texturas [29]. Algunos autores han relacionado los filtros bilaterales con otras estrategias de regularización adaptativas y análisis de datos, como pueden ser la difusión anisotrópica o el *mean-shift* [30], [31], [32]. Finalmente, en 2017, Wilkin y Beliaikov presentaron una reformulación de los filtros bilaterales [21] usando funciones débilmente monótonas (*weakly monotone functions* [33]).

Nuestra propuesta, aunque es conceptualmente similar a la de Tomasi y Manduchi, tiene una formulación diferente. El cambio más significativo es que el término de intensidad incluye dos posiciones simétricas respecto a \mathbf{p} , mientras que en un filtro bilateral se evalúa la similitud tonal entre los valores respecto de la posición central (ver Ec. (9)). Otra diferencia significativa es que el dominio espacial de los filtros bilaterales es simétrico desde el origen (\mathbf{p}), mientras que en la Ec. (7) κ sólo necesita definirse en la mitad del mismo.

V. VALIDACIÓN EXPERIMENTAL

La propuesta de la Sección IV se basa en la asunción de que la diferencia aritmética no es necesariamente el mejor descriptor de la disimilitud entre dos valores. O, al menos, de que es bueno tener la capacidad de modelar esta disimilitud por medio de una función dedicada. Esta sección trata de comprobar si el aislar el *término de intensidad* puede conllevar mejoras en los resultados prácticos de una aplicación de diferenciación. Para ello, hemos aplicado nuestras ideas al contexto de detección de bordes.

Proponemos un esquema de detección de bordes relativamente estándar, compuesto por cuatro pasos acordes a la estructura de descomposición de Bezdek [34], [35]:

1. Regularizar la imagen con un filtro Gaussiano con $\sigma = 1$;
2. Calcular los gradientes como la combinación de las derivadas parciales, aproximados de acuerdo a la Ec. (7) con los ajustes que se detallan a continuación.
3. Estimar la intensidad de borde como la magnitud del gradiente en cada píxel.
4. Binarizar la imagen usando *non-maxima suppresion* [36] e *hysteresis* [11], para la cuál los umbrales se establecen usando la técnica de Medina-Carnicer *et al.* [37].

Este esquema es una aplicación directa de la diferenciación de señales discretas (imágenes). Los ajustes de nuestros filtros son los que siguen:

- ϵ se construye como se ve en la Ec. (8), donde las RDFs son del tipo $r_{[\alpha,\beta]}$ con valores $\alpha, \beta \in \{0.5, 1, 1.5, \dots, 3\}$;
- κ_h y κ_v son equivalentes al lado positivo de un filtro de Canny, i.e

$$\kappa_h(\xi) = \frac{-\xi_x}{\sigma^2} e^{-\frac{\xi_x^2 + \xi_y^2}{2\sigma^2}} \quad \text{and} \quad \kappa_v(\xi) = \frac{-\xi_y}{\sigma^2} e^{-\frac{\xi_x^2 + \xi_y^2}{2\sigma^2}}, \quad (10)$$

con $\sigma = 1$.

		β					
		0.5	1.0	1.5	2.0	2.5	3.0
α	0.5	.56	.63	.65	.64	.63	.61
	1.0	.56	.63	.64	.63	.61	.58
	1.5	.56	.61	.61	.59	.57	.55
	2.0	.56	.59	.58	.55	.53	.50
	2.5	.55	.57	.55	.52	.49	.47
	3.0	.54	.54	.51	.48	.46	.43

(a) Rendimiento medio (medida F) para cada RDF $r_{[\alpha,\beta]}$.

		β					
		0.5	1.0	1.5	2.0	2.5	3.0
α	0.5	4/13	26/-	24/-	16/-	16/-	16/3
	1.0	2/3	11/-	14/-	6/-	8/-	2/2
	1.5	4/3	7/1	2/-	1/1	1/1	3/-
	2.0	5/3	3/-	3/-	1/-	-/-	1/2
	2.5	2/9	4/-	2/2	-/-	2/2	6/12
	3.0	2/19	1/5	3/2	1/9	-/13	1/95

(b) Número de imágenes para las cuales cada RDF $r_{[\alpha,\beta]}$ obtiene el mejor/peor resultados de entre los contendientes.

TABLE I

RESULTADOS CUANTITATIVOS OBTENIDOS EN EL EXPERIMENTO PARA CADA RDF $r_{[\alpha,\beta]}$ EN EL BSDS500 *test set*.

Nótese que, cuando usamos $r_{[1,1]}$, el algoritmo es computacionalmente equivalente al método de Canny. De esta manera, al incluir $(\alpha, \beta) = (1, 1)$ en la comparación, podemos contrastar la importancia de tener un término independiente y adaptable para la cuantificación de las disimilitudes.

En este experimento utilizamos un esquema estándar de cuantificación de resultados basado en la medida F [38], [39] sobre el *dataset de test* del BSDS [40].

En la Tabla I mostramos los resultados obtenidos en los experimentos. Primero, en la Tabla I(a) listamos el rendimiento medio de cada combinación de (α, β) . Luego, en la Tabla I(b) mostramos el número de imágenes para los cuales cada combinación de (α, β) produce el mejor y/o peor resultado (en términos de la medida F).

Los resultados indican que la combinación $(\alpha, \beta) = (1, 1)$ (que recupera el método de Canny) no es necesariamente la mejor opción para el término de intensidad. Otras configuraciones producen resultados que son, al menos, competitivos, especialmente con valores bajos de α y β . Esta conclusión se refiere a los valores medios de rendimiento, pero también a los números de mejores y peores resultados. Son llamativos los casos de $r_{[0.5,1.0]}$ y, por ejemplo, $r_{[0.5,2.5]}$, ya que ambos tienen mejores resultados medios, así como mejores ratios de mejor/peor resultado que $r_{[1,1]}$. Si bien experimentos exhaustivos podrían llevarse a cabo para confirmar nuestras hipótesis, creemos que nuestros resultados, dado el limitado espacio, refuerzan nuestra tesis de que la diferencia absoluta podría no ser siempre la mejor medida de disimilitud entre (en este caso) tonos en una imagen y (de manera general) valores en una función discreta.

VI. CONCLUSIONES

En este trabajo hemos presentado un modelo para generalizar filtros antisimétricos (impares) para la aproximación de derivadas mediante diferencias discretas. Nuestro modelo propone el uso de términos explícitos para las diferencias de valor (término de intensidad) y posición (término espacial). En el caso de las primeras, hemos presentado una clase de funciones de comparación basadas en RDFs que garantizan una serie de propiedades, añadidas a la alta flexibilidad. La validez del modelo se ha testeado en la diferenciación de imágenes en escala de grises para detección de bordes, encontrando que los resultados de algoritmos clásicos pueden mejorarse al usar las RDFs dentro de nuestra reformulación del modelo. Los resultados son prometedores, aunque deberían ser respaldados por trabajos más exhaustivos. Además, cabe señalar que nuestro modelo propuesto abre la puerta al uso de técnicas como el entrenamiento de funciones en el espacio de intensidades (tonos, en el caso de procesamiento de imagen), que en una formulación clásica del problema no eran elegibles.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia del Gobierno de España (Proyectos TIN2016-77356-P y TIN15-70266-C2-P-1).

REFERENCES

- [1] J. Astola, E. Dougherty, I. Shmulevich, and I. Tabus, "Genomic signal processing," *Signal Processing*, vol. 83, no. 4, pp. 691–694, 2003.
- [2] T.-S. Chua, X. He, W. Liu, M. Piccardi, Y. Wen, and D. Tao, "Big data meets multimedia analytics," *Signal Processing*, vol. 124, pp. 1–4, 2016.
- [3] L. Florack, *Image structure*. Springer Science & Business Media, 1997, vol. 10.
- [4] D. Marr, *Vision*. MIT Press, 1982.
- [5] I. Sobel and G. Feldman, "A 3x3 isotropic gradient operator for image processing," 1968, presented at a talk at the Stanford Artificial Intelligence Project.
- [6] J. M. S. Prewitt, *Object enhancement and extraction*, ser. Picture Processing and Psychopictorics. Academic Press, 1970, pp. 75–149.
- [7] L. G. Roberts, "Machine perception of three-dimensional solids," Ph.D. dissertation, Massachusetts Institute of Technology, 1963.
- [8] R. M. Haralick, "Digital step edges from zero crossing of second directional derivatives," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, no. 1, pp. 58–68, 1984.
- [9] M. Basu, "Gaussian-based edge-detection methods- A survey," *IEEE Trans. on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 32, no. 3, pp. 252–260, 2002.
- [10] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London*, vol. 207, no. 1167, pp. 187–217, 1980.
- [11] J. Canny, "Finding edges and lines in images," Massachusetts Institute of Technology, Tech. Rep., 1983.
- [12] —, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
- [13] G. Wang and B. De Baets, "Edge detection based on the fusion of multiscale anisotropic edge strength measurements," in *Advances in Fuzzy Logic and Technology 2017*. Springer, 2017, pp. 530–536.
- [14] S. Mahmoodi, "Edge detection filter based on Mumford–Shah Green function," *SIAM Journal on Imaging Sciences*, vol. 5, no. 1, pp. 343–365, 2012.
- [15] O. Lalignant and F. Truchetet, "A nonlinear derivative scheme applied to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 2, pp. 242–257, 2010.
- [16] W. McIlhagga, "The Canny edge detector revisited," *International Journal of Computer Vision*, vol. 91, pp. 251–261, 2011.
- [17] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 117–156, 1998.
- [18] J. Weickert, *Anisotropic Diffusion in Image Processing*, ser. ECMI Series. Teubner-Verlag, 1998.
- [19] C. Lopez-Molina, J. Montero, H. Bustince, and B. De Baets, "Self-adapting weighted operators for multiscale gradient fusion," *Information Fusion*, vol. 44, pp. 136–146, 2018.
- [20] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. of the IEEE International Conf. on Computer Vision*, 1998, pp. 838–846.
- [21] T. Wilkin and G. Beliakov, "Robust image denoising and smoothing with generalised spatial-tonal averages," in *IEEE International Conf. on Fuzzy Systems*. IEEE, 2017, pp. 1–7.
- [22] V. Gregori, S. Morillas, and A. Sapena, "Examples of fuzzy metrics and applications," *Fuzzy Sets and Systems*, vol. 170, no. 1, pp. 95–111, 2011.
- [23] S. Santini and R. Jain, "Similarity measures," *IEEE Trans. on Pattern Analysis Machine Intelligence*, vol. 21, no. 9, pp. 871–883, 1999.
- [24] H. Bustince, E. Barrenechea, and M. Pagola, "Relationship between restricted dissimilarity functions, restricted equivalence functions and normal EN-functions: Image thresholding invariant," *Pattern Recognition Letters*, vol. 29, no. 4, pp. 525–536, 2008.
- [25] C. Lopez-Molina, B. De Baets, M. Galar, and H. Bustince, "A generalization of the Perona-Malik anisotropic diffusion method using restricted dissimilarity functions," *International Journal of Computational Intelligent Systems*, vol. 6, no. 1, pp. 14–28, 2013.
- [26] C. Marco-Detchart, J. Cerron, L. De Miguel, C. Lopez-Molina, H. Bustince, and M. Galar, "A framework for radial data comparison and its application to fingerprint analysis," *Applied Soft Computing*, vol. 46, pp. 246–259, 2016.
- [27] P.-L. Shui and W.-C. Zhang, "Noise-robust edge detector combining isotropic and anisotropic Gaussian kernels," *Pattern Recognition*, vol. 45, no. 2, pp. 806–820, 2012.
- [28] J. Shen and S. Castan, "An optimal linear operator for step edge detection," *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 2, pp. 112–133, 1992.
- [29] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, "Bilateral filtering: Theory and applications," *Foundations and Trends in Computer Graphics and Vision*, vol. 4, no. 1, pp. 1–73, 2008.
- [30] M. Elad, "On the origin of the bilateral filter and ways to improve it," *IEEE Trans. on Image Processing*, vol. 11, no. 10, pp. 1141–1151, 2002.
- [31] D. Barash, "A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 844–847, 2002.
- [32] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [33] T. Wilkin and G. Beliakov, "Weakly monotonic averaging functions," *International Journal of Intelligent Systems*, vol. 30, no. 2, pp. 144–169, 2015.
- [34] J. Bezdek, R. Chandrasekhar, and Y. Attikouzel, "A geometric approach to edge detection," *IEEE Trans. on Fuzzy Systems*, vol. 6, no. 1, pp. 52–75, 1998.
- [35] C. Lopez-Molina, "The breakdown structure of edge detection- analysis of individual components and revisit of the overall structure," Ph.D. dissertation, Universidad Publica de Navarra, 2012.
- [36] A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene analysis," *IEEE Trans. on Computers*, vol. 20, no. 5, pp. 562–569, 1971.
- [37] R. Medina-Carnicer, F. Madrid-Cuevas, A. Carmona-Poyato, and R. Muñoz-Salinas, "On candidates selection for hysteresis thresholds in edge detection," *Pattern Recognition*, vol. 42, no. 7, pp. 1284–1296, 2009.
- [38] D. Martin, C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 530–549, 2004.
- [39] C. Lopez-Molina, B. De Baets, and H. Bustince, "Quantitative error measures for edge detection," *Pattern Recognition*, vol. 46, no. 4, pp. 1125–1139, 2013.
- [40] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 898–916, 2011.



Una nueva aproximación al concepto de similitud intervalar teniendo en cuenta la longitud de los intervalos

Humberto Bustince, Javier Fernandez, Laura de Miguel, Jose Antonio Sanz, Mikel Sesma-Sara

Depto. de Estadística, Informática y Matemáticas

Universidad Pública de Navarra

Pamplona, Spain

{bustince, fcojavier.fernandez,laura.demiguel, joseantonio.sanz, mikel.sesma}@unavarra.es

Abstract—En este trabajo proponemos una definición de función de equivalencia restringida intervalo-valorada que determina el grado de similitud de dos intervalos teniendo en cuenta su amplitud y utilizando órdenes admisibles. Utilizamos estas funciones para definir medidas de similitud intervalo-valoradas y discutimos algunas posibles aplicaciones.

Index Terms—Función de equivalencia restringida intervalo-valorada, Función de similitud intervalo-valorada, Orden admisible, Medida de similitud

I. INTRODUCCIÓN

Los conjuntos intervalo-valorados (IVFSs) son cada vez más utilizados debido a su excelente rendimiento en múltiples aplicaciones [1]–[3], [10]. Dado que muchas de estas aplicaciones, en su versión difusa, hacen uso del concepto de medida de similitud [4], [12], las medidas de similitud intervalo-valoradas están atrayendo también un gran interés [11].

El objetivo de este trabajo es construir medidas de similitud intervalo-valoradas que, por una parte, utilicen órdenes totales (y no solo parciales) a la hora de comparar los intervalos involucrados y que, además, tengan en cuenta la amplitud de los intervalos, interpretada como una medida del grado de incertidumbre asociado a los datos. Para lograr este objetivo, consideramos nuevas definiciones de funciones de agregación y funciones de equivalencia restringida intervalo-valoradas que tienen en cuenta tanto órdenes totales como la amplitud de los intervalos.

La estructura del trabajo es la siguiente. En la Sección II presentamos algunos resultados y definiciones preliminares. En la Sección III presentamos la definición de función de equivalencia restringida intervalo-valorada y en la Sección IV, discutimos la construcción de funciones de agregación intervalo-valoradas que preservan la amplitud. La Sección V se centra en el nuevo concepto de medida de similitud intervalo-valorada basada en la amplitud. Terminamos con algunas conclusiones y referencias.

Este trabajo ha sido parcialmente financiado por el proyecto de investigación TIN2016-77356-P del Gobierno de España.

II. PRELIMINARES

Vamos a trabajar con subintervalos cerrados del intervalo unidad. Por ello, definimos el siguiente conjunto:

$$L([0, 1]) = \{[\underline{X}, \overline{X}] \mid 0 \leq \underline{X} \leq \overline{X} \leq 1\}.$$

La amplitud de un intervalo $X \in L([0, 1])$ se denota por $w(X)$. Una función intervalar $f : (L([0, 1]))^n \rightarrow L([0, 1])$ se dice que preserva la amplitud si para todo $X_1, \dots, X_n \in L([0, 1])$ tales que $w(X_1) = \dots = w(X_n)$, se tiene que $w(f(X_1, \dots, X_n)) = w(X_i)$ para cualquier $i \in \{1, \dots, n\}$.

Denotamos por \leq_L una relación de orden arbitraria en $L([0, 1])$ con $0_L = [0, 0]$ como elemento mínimo y $1_L = [1, 1]$ como elemento máximo. Esta relación de orden puede ser total o parcial. Si queremos hablar específicamente de un orden total, lo denotaremos por \leq_{TL} .

Ejemplo 1: Ejemplos de relaciones de orden en $L([0, 1])$.

- a) La relación de orden parcial en $L([0, 1])$ inducida por el orden (parcial) usual en \mathbb{R}^2 es:

$$[\underline{X}, \overline{X}] \lesssim_L [\underline{Y}, \overline{Y}] \text{ si } \underline{X} \leq \underline{Y} \text{ y } \overline{X} \leq \overline{Y}. \quad (1)$$

- b) Como ejemplo de orden total en $L([0, 1])$ tenemos el de Xu y Yager (véase [14]):

$$[\underline{X}, \overline{X}] \leq_{XY} [\underline{Y}, \overline{Y}] \text{ si } \begin{cases} \underline{X} + \overline{X} < \underline{Y} + \overline{Y} \text{ o} \\ \underline{X} + \overline{X} = \underline{Y} + \overline{Y} \text{ y} \\ \overline{X} - \underline{X} \leq \overline{Y} - \underline{Y}. \end{cases} \quad (2)$$

Definición 1: Un orden admisible en $L([0, 1])$ es un orden lineal \leq_{TL} que extiende el orden parcial \lesssim_L .

En este trabajo, cuando hablamos de un orden lineal entre intervalos, asumimos que es admisible. El siguiente resultado proporciona un método de construcción de órdenes admisibles.

Proposición 1: ([9]) Sean $M_1, M_2 : [0, 1]^2 \rightarrow [0, 1]$ dos funciones de agregación (es decir, funciones crecientes tales que, si todas las entradas son cero, el resultado es cero y, si todas las entradas son uno, el resultado es uno) tales que, para cualesquiera $X, Y \in L([0, 1])$, las identidades $M_1(\underline{X}, \overline{X}) = M_1(\underline{Y}, \overline{Y})$ y $M_2(\underline{X}, \overline{X}) = M_2(\underline{Y}, \overline{Y})$ se

satisfacen simultáneamente si y solo si $X = Y$. Entonces, el orden \leq_{M_1, M_2} en $L([0, 1])$ dado por

$$X \leq_{M_1, M_2} Y \quad \text{si} \quad \begin{cases} M_1(\underline{X}, \bar{X}) < M_1(\underline{Y}, \bar{Y}) \text{ o} \\ M_1(\underline{X}, \bar{X}) = M_1(\underline{Y}, \bar{Y}) \text{ y} \\ M_2(\underline{X}, \bar{X}) \leq M_2(\underline{Y}, \bar{Y}) \end{cases}$$

es un orden admisible en $L([0, 1])$.

Ejemplo 2:

- (i) El orden de Xu y Yager es un ejemplo de orden admisible con $M_1(x, y) = \frac{x+y}{2}$ y $M_2(x, y) = y$.
- (ii) Los órdenes lexicográficos \leq_{lex1} (\leq_{lex2}) también son ejemplos de órdenes admisibles con $M_1(x, y) = x$ ($M_1(x, y) = y$) y $M_2(x, y) = y$ ($M_2(x, y) = x$).
- (iii) En general, si, para $\alpha \in [0, 1]$ definimos la función de agregación

$$K_\alpha(x, y) = (1 - \alpha)x + \alpha y$$

entonces, para $\alpha, \beta \in [0, 1]$ con $\alpha \neq \beta$, podemos obtener el orden admisible $\leq_{\alpha, \beta}$ tomando $M_1(x, y) = K_\alpha(x, y)$ y $M_2(x, y) = K_\beta(x, y)$. Véase [9] para más detalles.

Recordemos también la definición de función de equivalencia restringida (REF) [7]

Definición 2: Una función $R : [0, 1]^2 \rightarrow [0, 1]$ es una función de equivalencia restringida (REF) si:

- 1) $R(x, y) = 0$ si y solo si $\{x, y\} = \{0, 1\}$;
- 2) $R(x, y) = 1$ si y solo si $x = y$;
- 3) $R(x, y) = R(y, x)$ para todo $x, y \in [0, 1]$;
- 4) Si $x \leq y \leq z$, entonces $R(x, z) \leq R(x, y)$ y $R(x, z) \leq R(y, z)$ para todo $x, y, z \in [0, 1]$.

III. FUNCIONES DE EQUIVALENCIA RESTRINGIDA INTERVALO-VALORADAS QUE PRESERVAN LA AMPLITUD

En esta sección proponemos una nueva definición de REF para datos intervalo-valorados que tiene en cuenta la amplitud de los mismos.

Definición 3: Sea \leq_L un orden en $L([0, 1])$. Una función de equivalencia restringida intervalo-valorada (IV REF) con respecto al orden \leq_L es una función $R_{IV} : L([0, 1])^2 \rightarrow L([0, 1])$ tal que:

- 1) $R_{IV}(X, Y) = 0_L$ si y solo si $\{X, Y\} = \{0_L, 1_L\}$;
- 2) $R_{IV}(X, X) = [1 - w(X), 1]$ para todo $X \in L([0, 1])$;
- 3) $R_{IV}(X, Y) = R_{IV}(Y, X)$ para todo $X, Y \in L([0, 1])$;
- 4) Si $X, Y, Z \in L([0, 1])$ son tales que $X \leq_L Y \leq_L Z$ y $w(X) = w(Y) = w(Z)$, entonces $R_{IV}(X, Z) \leq_L R_{IV}(X, Y)$ y $R_{IV}(X, Z) \leq_L R_{IV}(Y, Z)$.

La principal diferencia de esta definición respecto a su contrapartida difusa radica en el axioma 2. Dado que consideramos que la amplitud del intervalo de pertenencia de un elemento es una medida de la incertidumbre asociada al valor preciso de pertenencia, y que dicho valor preciso es un número dentro del intervalo de pertenencia, de esta forma se evita que el resultado sea menos impreciso que los datos considerados.

Ejemplo 3:

La función $R_{IV} : L([0, 1])^2 \rightarrow L([0, 1])$ dada por:

$$R_{IV}(X, Y) = \left[\max(0, 1 - |K_\alpha(X) - K_\alpha(Y)| - \frac{1}{2}(w(X) + w(Y))), \max(1 - |K_\alpha(X) - K_\alpha(Y)|, \frac{1}{2}(w(X) + w(Y))) \right]$$

es, para todo $\alpha \in]0, 1]$, un ejemplo de IV REF con respecto a cualquier orden admisible.

A continuación discutimos un método de construcción de estas funciones. Para ello, necesitamos el siguiente lema previo.

Lema 1: Sean $X, Y \in L([0, 1])$ intervalos tales que $w(X) = w(Y)$. Entonces

$$X \lesssim_L Y \quad \Leftrightarrow \quad X \leq_{TL} Y$$

para cualquier orden admisible \leq_{TL} .

Prueba. La demostración se sigue al tener en cuenta que dos intervalos cualesquiera de la misma amplitud son siempre comparables por medio del orden parcial \lesssim_L . Dado que todo orden admisible refina este orden parcial, se tiene el resultado. ■

Teorema 1: Sea $\alpha \in]0, 1[$, $M : [0, 1]^2 \rightarrow [0, 1]$ una función de agregación simétrica e idempotente y $R : [0, 1]^2 \rightarrow [0, 1]$ una REF. Entonces, la función $R_{IV} : L([0, 1])^2 \rightarrow L([0, 1])$ dada por

$$R_{IV}(X, Y) = \left[\max(0, R(K_\alpha(X), K_\alpha(Y)) - M(w(X), w(Y))), \max(R(K_\alpha(X), K_\alpha(Y)), M(w(X), w(Y))) \right] \quad (3)$$

es una IV REF con respecto a cualquier orden admisible \leq_{TL} que preserva la amplitud

Prueba. Por comodidad, escribimos \mathcal{R} por $R(K_\alpha(X), K_\alpha(Y))$, y \mathcal{M} por $M(w(X), w(Y))$. Entonces, la ecuación (3) se simplifica como:

$$R_{IV}(X, Y) = \left[\max(0, \mathcal{R} - \mathcal{M}), \max(\mathcal{R}, \mathcal{M}) \right] \quad (4)$$

$$= \begin{cases} [\mathcal{R} - \mathcal{M}, \mathcal{R}], & \text{si } \mathcal{R} \geq \mathcal{M}, \\ [0, \mathcal{M}], & \text{en otro caso.} \end{cases}$$

De (4), R_{IV} está bien definido. Además, $R_{IV}(X, Y) = 0_L$ si y solo si $\mathcal{R} = 0$ y $\mathcal{M} = 0$. Pero esto ocurre si y solo si $\{K_\alpha(X), K_\alpha(Y)\} = \{0, 1\}$, lo que sucede si y solo si $\{X, Y\} = \{0_L, 1_L\}$. Luego $w(X) = w(Y) = 0$ y tenemos la primera condición de la Definición 3.

La segunda condición de la Definición 3 se sigue al observar que $R(K_\alpha(X), K_\alpha(X)) = 1$ y $M(w(X), w(X)) = w(X)$.

La simetría de R_{IV} es una consecuencia inmediata de la simetría de R y M .

La monotonía con respecto a cualquier orden admisible es clara a partir de la monotonía de R , el lema 1 y el hecho de que, si $X \leq_{TL} Y \leq_{TL} Z$ y $w(X) = w(Y) = w(Z)$, entonces $K_\alpha(X) \leq K_\alpha(Y) \leq K_\alpha(Z)$.



Finalmente, la preservación de la amplitud R_{IV} se sigue de la Ecuación (4) y de la idempotencia de M ■

Podemos utilizar cualquier REF R y cualquier función de agregación idempotente y simétrica M en la Ecuación (3) para construir funciones de equivalencia restringida intervalo-valoradas que preserven la amplitud. De hecho, es posible simplificar la Ecuación (3) imponiendo restricciones adicionales sobre R y M .

Corolario 1: Sean $\alpha \in]0, 1[$ y $M : [0, 1]^2 \rightarrow [0, 1]$ una función de agregación simétrica idempotente tal que

$$M(x, y) \leq \min((1 - \alpha)x + \alpha y, \alpha x + (1 - \alpha)y)$$

para cualesquiera $x, y \in [0, 1]$. Sea $R : [0, 1]^2 \rightarrow [0, 1]$ una función de equivalencia restringida tal que

$$R(x, y) \geq 1 - |x - y|$$

para todo $x, y \in [0, 1]$. Entonces, la función $R_{IV} : L([0, 1])^2 \rightarrow L([0, 1])$ dada por

$$\begin{aligned} R_{IV}(X, Y) &= [R(K_\alpha(X), K_\alpha(Y)) \\ &- M(w(X), w(Y)), R(K_\alpha(X), K_\alpha(Y))] \end{aligned} \quad (5)$$

es una función de equivalencia restringida intervalo-valorada con respecto a cualquier orden admisible \leq_{TL} y que preserva la amplitud.

Prueba Es necesario ver que

$$R(K_\alpha(X), K_\alpha(Y)) \geq M(w(X), w(Y))$$

para todo $X, Y \in L([0, 1])$, ya que, de este modo, la Ecuación (5) es un caso especial de la Ecuación (3).

Por nuestras hipótesis sobre M y R , es suficiente mostrar que

$$\begin{aligned} 1 - |K_\alpha(X) - K_\alpha(Y)| \\ \geq \min((1 - \alpha)w(X) + \alpha w(Y), \alpha w(X) + (1 - \alpha)w(Y)). \end{aligned} \quad (6)$$

Supongamos que $K_\alpha(X) \geq K_\alpha(Y)$. Entonces

$$\begin{aligned} 1 - |K_\alpha(X) - K_\alpha(Y)| \\ = 1 - (1 - \alpha)\underline{X} - \alpha\overline{X} + (1 - \alpha)\underline{Y} + \alpha\overline{Y} \end{aligned}$$

y como

$$\begin{aligned} 1 &\geq \overline{X} - \underline{Y} \\ &= (1 - \alpha)(\overline{X} - \underline{Y}) + \alpha(\overline{X} - \underline{Y}) \\ &= (1 - \alpha)(\overline{X} - \underline{X} + \underline{X} - \underline{Y}) \\ &+ \alpha(\overline{X} - \overline{Y} + \overline{Y} - \underline{Y}), \end{aligned}$$

tenemos que

$$\begin{aligned} 1 - (1 - \alpha)\underline{X} - \alpha\overline{X} + (1 - \alpha)\underline{Y} + \alpha\overline{Y} \\ \geq (1 - \alpha)(\overline{X} - \underline{X}) + \alpha(\overline{Y} - \underline{Y}), \end{aligned}$$

luego (6) se satisface.

Supongamos ahora que $K_\alpha(X) < K_\alpha(Y)$. Tenemos que

$$\begin{aligned} 1 - |K_\alpha(X) - K_\alpha(Y)| \\ = 1 + (1 - \alpha)\underline{X} + \alpha\overline{X} - (1 - \alpha)\underline{Y} - \alpha\overline{Y} \end{aligned}$$

y como

$$\begin{aligned} 1 &\geq \overline{Y} - \underline{X} \\ &= (1 - \alpha)(\overline{Y} - \underline{X}) + \alpha(\overline{Y} - \underline{X}) \\ &= (1 - \alpha)(\overline{Y} - \underline{Y} + \underline{Y} - \underline{X}) \\ &+ \alpha(\overline{Y} - \overline{X} + \overline{X} - \underline{X}), \end{aligned}$$

vemos que

$$\begin{aligned} 1 + (1 - \alpha)\underline{X} + \alpha\overline{X} - (1 - \alpha)\underline{Y} - \alpha\overline{Y} \\ \geq (1 - \alpha)(\overline{Y} - \underline{Y}) + \alpha(\overline{X} - \underline{X}), \end{aligned}$$

de donde se cumple (6) y tenemos el resultado. ■

Corolario 2: Consideremos la función de equivalencia restringida intervalo-valorada R_{IV} propuesta en el Corolario 1. Entonces, para todo $X, Y \in L([0, 1])$ se tiene que

$$\begin{aligned} \min(w(X), w(Y)) &\leq w(R_{IV}(X, Y)) = M(w(X), w(Y)) \\ &\leq \min((1 - \alpha)w(X) + \alpha w(Y), \alpha w(X) + (1 - \alpha)w(Y)). \end{aligned}$$

Prueba. La primera desigualdad se verifica porque toda función de agregación idempotente es siempre mayor o igual que el mínimo. En cuanto a la segunda desigualdad, es una consecuencia de las hipótesis sobre M asumidas en el Corolario 1. ■

El siguiente resultado es directo.

Lema 2: Si $M_1, M_2 : [0, 1]^2 \rightarrow [0, 1]$ son funciones de agregación simétricas idempotentes, entonces la función $M : [0, 1]^2 \rightarrow [0, 1]$ dada por

$$M(x, y) = \min(M_1(x, y), M_2(x, y)),$$

para todo $x, y \in [0, 1]$, es una función de agregación simétrica idempotente.

Ejemplo 4: Vamos a construir un ejemplo específico de función de equivalencia restringida intervalo-valorada construida de acuerdo con el Corolario 1. Sea $R^p(x, y) = 1 - |x - y|^p$, con $p > 0$, que es una REF. Es claro que $R^p(x, y) \geq R^1 = 1 - |x - y|$ para todo $x, y \in [0, 1]$ si y solo si $p \in [1, \infty[$.

(i) Si tomamos $\alpha = 1/2$ y $M(x, y) = \frac{x+y}{2}$ obtenemos una clase de funciones de equivalencia restringida intervalo-valoradas respecto a cualquier orden admisible:

$$\begin{aligned} R_{IV}^p(X, Y) &= \left[R^p\left(\frac{\underline{X} + \overline{X}}{2}, \frac{\underline{Y} + \overline{Y}}{2}\right) \right. \\ &- \left. \frac{w(X) + w(Y)}{2}, R^p\left(\frac{\underline{X} + \overline{X}}{2}, \frac{\underline{Y} + \overline{Y}}{2}\right) \right] \end{aligned}$$

para cualquier $p \in [1, \infty[$.

(ii) Si tomamos $M(x, y) = \min(x, y)$, obtenemos:

$$\begin{aligned} R_{IV}^{p,\alpha}(X, Y) &= [R^p(K_\alpha(X), K_\alpha(Y)) - \min(w(X), w(Y)), \\ &R^p(K_\alpha(X), K_\alpha(Y))] \end{aligned}$$

para $p \in [1, \infty[$ y $\alpha \in]0, 1[$.

(iii) Sea $\alpha \in [0, 1]$. Es fácil observar que

$$\begin{aligned} \min((1 - \beta)x + \beta y, \beta x + (1 - \beta)y) \\ \leq \min((1 - \alpha)x + \alpha y, \alpha x + (1 - \alpha)y) \end{aligned}$$

para todo $\beta \in [0, 1]$ tal que

$$\max(\beta, 1 - \beta) \geq \max(\alpha, 1 - \alpha).$$

Por tanto, obtenemos una clase más general que la del ítem

(ii) si tomamos $\alpha \in]0, 1[$ y

$$M(x, y) = \min((1 - \beta)x + \beta y, \beta x + (1 - \beta)y)$$

para $\beta \in [\max(\alpha, 1 - \alpha), 1]$ (o equivalentemente para $\beta \in [0, \min(\alpha, 1 - \alpha)]$):

$$\begin{aligned} & R_{IV}^{p, \alpha, \beta}(X, Y) \\ &= [R^p(K_\alpha(X), K_\alpha(Y)) - \min((1 - \beta)w(X) + \beta w(Y), \\ & \quad \beta w(X) + (1 - \beta)w(Y)), R^p(K_\alpha(X), K_\alpha(Y))]. \end{aligned}$$

En particular, para $\beta = 1$ (o equivalentemente para $\beta = 0$) obtenemos la clase descrita en el ítem (ii).

IV. FUNCIONES DE AGREGACIÓN INTERVALO-VALORADAS QUE PRESERVAN LA AMPLITUD

De cara a posibles aplicaciones, introducimos en esta sección un análisis de la posible definición de funciones de agregación intervalo-valoradas que también tengan en cuenta la amplitud de los datos a agregar. Recordamos primero la definición de función de agregación intervalo-valorada con respecto a un orden admisible arbitrario.

Definición 4: Sea $n \geq 2$. Una función de agregación intervalo-valorada (n -dimensional) en $L([0, 1])$ con respecto a un orden \leq_L es una aplicación $M_{IV} : (L([0, 1]))^n \rightarrow L([0, 1])$ tal que:

- (i) $M_{IV}(0_L, \dots, 0_L) = 0_L$.
- (ii) $M_{IV}(1_L, \dots, 1_L) = 1_L$.
- (iii) M_{IV} es no decreciente respecto a \leq_L .

A continuación proponemos un método de construcción de funciones de agregación intervalo-valoradas que preservan la amplitud de los intervalos a agregar.

Para ello, dada una función de agregación $M : [0, 1]^n \rightarrow [0, 1]$, consideramos las dos propiedades siguientes:

- (P1) $M(cx_1, \dots, cx_n) \geq cM(x_1, \dots, x_n)$ para todo $c \in [0, 1]$, $x_1, \dots, x_n \in [0, 1]$.
- (P2) $M(x_1, \dots, x_n) \leq 1 - M(1 - x_1, \dots, 1 - x_n)$ para todo $x_1, \dots, x_n \in [0, 1]$.

Teorema 2: Sean $\alpha, \beta \in [0, 1]$ con $\beta \neq \alpha$. Sean $M_1, M_2 : [0, 1]^n \rightarrow [0, 1]$ dos funciones de agregación tales que M_1 es estrictamente creciente, $M_1(x_1, \dots, x_n) \geq M_2(x_1, \dots, x_n)$ para todo $x_1, \dots, x_n \in [0, 1]$, M_1 o M_2 satisfacen la propiedad (P1) y M_1 o M_2 satisfacen la propiedad (P2). Entonces, $M_{IV} : (L([0, 1]))^n \rightarrow L([0, 1])$ dada por:

$$M_{IV}(X_1, \dots, X_n) = Y, \quad \text{donde}$$

$$\begin{cases} K_\alpha(Y) = M_1(K_\alpha(X_1), \dots, K_\alpha(X_n)), \\ w(Y) = M_2(w(X_1), \dots, w(X_n)), \end{cases}$$

para todo $X_1, \dots, X_n \in L([0, 1])$, es una función de agregación intervalo-valorada con respecto a $\leq_{\alpha, \beta}$.

Además, si M_2 es idempotente, entonces M_{IV} preserva la amplitud.

Prueba.

Primero veamos que M_{IV} está bien definida. Tenemos

$$Y = [\underline{Y}, \overline{Y}] = [K_\alpha(Y) - \alpha w(Y), K_\alpha(Y) + (1 - \alpha)w(Y)].$$

Como $\underline{Y} \leq \overline{Y}$, solo debemos probar que

- 1) $\underline{Y} \geq 0$: Para $\alpha = 0$ tenemos que $\underline{Y} = M_1(\underline{X}_1, \dots, \underline{X}_n) \geq 0$ y para $\alpha \in]0, 1[$ tenemos que

$$\begin{aligned} K_\alpha(Y) &= M_1(K_\alpha(X_1), \dots, K_\alpha(X_n)) \\ &\geq \alpha M_2\left(\frac{K_\alpha(X_1)}{\alpha}, \dots, \frac{K_\alpha(X_n)}{\alpha}\right) \\ &\geq \alpha M_2(w(X_1), \dots, w(X_n)) = \alpha w(Y) \end{aligned}$$

donde la primera desigualdad se sigue de que M_2 satisface la propiedad (P1) y la segunda, de que $K_\alpha(X) = (1 - \alpha)\underline{X} + \alpha\overline{X} \geq \alpha(\overline{X} - \underline{X}) = \alpha w(X)$ for all $X \in L([0, 1])$.

- 2) $\overline{Y} \leq 1$: Para $\alpha = 1$ tenemos que $\overline{Y} = M_1(\overline{X}_1, \dots, \overline{X}_n) \leq 1$ y para $\alpha \in [0, 1[$ vemos que

$$\begin{aligned} & K_\alpha(Y) + (1 - \alpha)w(Y) \\ &= M_1(K_\alpha(X_1), \dots, K_\alpha(X_n)) \\ &+ (1 - \alpha)M_2(w(X_1), \dots, w(X_n)) \\ &\leq M_1(K_\alpha(X_1), \dots, K_\alpha(X_n)) \\ &+ (1 - \alpha)M_2\left(\frac{1 - K_\alpha(X_1)}{1 - \alpha}, \dots, \frac{1 - K_\alpha(X_n)}{1 - \alpha}\right) \\ &\leq M_1(K_\alpha(X_1), \dots, K_\alpha(X_n)) \\ &+ M_2(1 - K_\alpha(X_1), \dots, 1 - K_\alpha(X_n)) \\ &\leq M_1(K_\alpha(X_1), \dots, K_\alpha(X_n)) \\ &+ 1 - M_2(K_\alpha(X_1), \dots, K_\alpha(X_n)) = 1 \end{aligned}$$

donde la primera desigualdad se sigue de que $1 - K_\alpha(X) = 1 - (1 - \alpha)\underline{X} - \alpha\overline{X} \geq (1 - \alpha)(\overline{X} - \underline{X}) = (1 - \alpha)w(X)$ for all $X \in L([0, 1])$, y la segunda y la tercera, de las hipótesis del Teorema.

Veamos ahora que M_{IV} es una función de agregación intervalo-valorada.

(i) $M_{IV}(0_L, \dots, 0_L) = Y$ donde $K_\alpha(Y) = M_1(0, \dots, 0) = 0$ y $w(Y) = M_2(0, \dots, 0) = 0$, luego $Y = 0_L$.

(ii) $M_{IV}(1_L, \dots, 1_L) = Y$ donde $K_\alpha(Y) = M_1(1, \dots, 1) = 1$ y $w(Y) = M_2(0, \dots, 0) = 0$, luego $Y = 1_L$.

(iii) Sea $X_i \leq_{\alpha, \beta} Y_i$ para todo $i = 1, \dots, n$. Entonces $K_\alpha(X_i) \leq K_\alpha(Y_i)$ para todo $i = 1, \dots, n$ y hay dos posibilidades:

- 1) Existe $j \in \{1, \dots, n\}$ tal que $K_\alpha(X_j) < K_\alpha(Y_j)$. Entonces

$$M_1(K_\alpha(X_1), \dots, K_\alpha(X_n)) < M_1(K_\alpha(Y_1), \dots, K_\alpha(Y_n)),$$

ya que M_1 es estrictamente creciente, luego $M_{IV}(X_1, \dots, X_n) <_{\alpha, \beta} M_{IV}(Y_1, \dots, Y_n)$.

- 2) $K_\alpha(X_i) = K_\alpha(Y_i)$ para todo $i = 1, \dots, n$. Si $\beta > \alpha$, entonces $w(X_i) \leq w(Y_i)$ para todo $i = 1, \dots, n$, de donde $M_2(w(X_1), \dots, w(X_n)) \leq$



$M_2(w(Y_1), \dots, w(Y_n))$, y por tanto $M_{IV}(X_1, \dots, X_n) \leq_{\alpha, \beta} M_{IV}(Y_1, \dots, Y_n)$. Si $\beta < \alpha$, entonces $w(X_i) \geq w(Y_i)$ para todo $i = 1, \dots, n$, luego $M_2(w(X_1), \dots, w(X_n)) \geq M_2(w(Y_1), \dots, w(Y_n))$, y $M_{IV}(X_1, \dots, X_n) \leq_{\alpha, \beta} M_{IV}(Y_1, \dots, Y_n)$.

Finalmente, la conservación de la amplitud se sigue fácilmente de la idempotencia de M_2 . ■

Ejemplo 5: Una función $M_{IV} : (L([0, 1]))^n \rightarrow L([0, 1])$ definida como en el Teorema 2, es una función de agregación intervalo-valorada que preserva la amplitud (con respecto a $\leq_{\alpha, \beta}$), si, por ejemplo:

- 1) (i) $M_1(x_1, \dots, x_n) = M_2(x_1, \dots, x_n) = \frac{x_1 + \dots + x_n}{n}$ para todo $x_1, \dots, x_n \in [0, 1]$, o
- 2) (ii) $M_1(x_1, \dots, x_n) = \frac{x_1 + \dots + x_n}{n}$, $M_2(x_1, \dots, x_n) = \min\{x_1, \dots, x_n\}$ para todo $x_1, \dots, x_n \in [0, 1]$.

V. MEDIDAS DE SIMILITUD INTERVALO-VALORADAS QUE PRESERVAN LA AMPLITUD

A continuación proponemos una nueva definición de medida de similitud intervalo-valorada que tiene en cuenta la amplitud de los intervalos.

Definición 5: Sea \leq_L un orden en $L([0, 1])$ y $M : [0, 1]^n \rightarrow [0, 1]$ una función de agregación. Una medida de similitud intervalo-valorada asociada a M basada en la amplitud los conjuntos intervalo-valorados definidos en un referencia finito U , $IVFS(U)$, y con respecto al orden \leq_L es una función $S_M : IVFS(U) \times IVFS(U) \rightarrow L([0, 1])$ tal que, para todo $A, B, A', B' \in IVFS(U)$,

$$[(SM1)] S_M(A, B) = S(B, A);$$

$$[(SM2)]$$

$$S_M(A, A) = \left[1 - M(w(A(u_1)), \dots, w(A(u_n))), 1\right];$$

[(SM3)] $S_M(A, B) = 0_L$ si y solo si $\{A(u_i), B(u_i)\} = \{0_L, 1_L\}$ para todo $i \in \{1, \dots, n\}$;

[(SM4)] Si $A \subseteq A' \subseteq B' \subseteq B$ w.r.t. \leq_L y $w(A(u_i)) = w(A'(u_i)) = w(B'(u_i)) = w(B(u_i))$ para todo $i \in \{1, \dots, n\}$, entonces $S_M(A, B) \leq_L S_M(A', B')$.

Preentamos ahora un método de construcción de estas similitudes. Recordemos que una función de agregación $M : [0, 1]^n \rightarrow [0, 1]$ es auto-dual con respecto a la negación estándar si

$$M(x_1, \dots, x_n) = 1 - M(1 - x_1, \dots, 1 - x_n)$$

si $x_1, \dots, x_n \in [0, 1]$.

Teorema 3:

Sea $M_{IV} : (L([0, 1]))^n \rightarrow L([0, 1])$ una función de agregación intervalo-valorada (respecto a un orden \leq_L) tal que $M_{IV}(X_1, \dots, X_n) = [M_L(\underline{X}_1, \dots, \underline{X}_n), M_U(\overline{X}_1, \overline{X}_n)]$ con M_L, M_U funciones de agregación y M_L auto-dual, y sea $M_{IV}(X_1, \dots, X_n) = 0_L$ si y solo si $X_1 = \dots = X_n = 0_L$. Sea $R_{IV} : L([0, 1])^2 \rightarrow L([0, 1])$ una IV REF con respecto al orden \leq_L . Entonces, la función $S_{M_L} : IVFS(U) \times IVFS(U) \rightarrow L([0, 1])$ dada por:

$$S_{M_L}(A, B) = M_{IV}(R_{IV}(A(u_1), B(u_1)), \dots, R_{IV}(A(u_n), B(u_n)))$$

para todo $A, B \in IVFS(U)$ es una medida de similitud intervalo-valorada asociada a M basada en la amplitud sobre $IVFS(U)$ con respecto al orden \leq_L .

Prueba. (SM2) Tenemos

$$\begin{aligned} S_{M_L}(A, A) &= M_{IV}([1 - w(A(u_1)), 1], \dots, [1 - w(A(u_n)), 1]) \\ &= [M_L(1 - w(A(u_1)), \dots, 1 - w(A(u_n))), 1] \\ &= [1 - M_L(w(A(u_1)), \dots, w(A(u_n))), 1] \end{aligned}$$

(SM1), (SM3) y (SM4) son directos. ■

Ejemplo 6: El siguiente es un ejemplo de IV similitud basada en la amplitud con respecto a la media aritmética y el orden definido por las funciones $K_{0.5}$ y K_1 (de acuerdo con la proposición 1).

$$\begin{aligned} S_{M_1}(A, B) &= \left[1 - \frac{\sum_{i=1}^n |A(u_i) + \overline{A(u_i)} - B(u_i) - \overline{B(u_i)}|}{2n}\right. \\ &\quad - \frac{\sum_{i=1}^n \min(w(A(u_i)), w(B(u_i)))}{n}, \\ &\quad \left.1 - \frac{\sum_{i=1}^n |A(u_i) + \overline{A(u_i)} - B(u_i) - \overline{B(u_i)}|}{2n}\right]. \end{aligned} \quad (7)$$

VI. CONCLUSIONES

En este trabajo hemos presentado nuevas aproximaciones a los conceptos de función de equivalencia restringida y función de similitud en el marco intervalar teniendo en cuenta órdenes admisibles y la amplitud de los intervalos.

Estas nociones son aplicables directamente en todos aquellos problemas donde, o bien se han utilizado conjuntos intervalo-valorados, o bien se han utilizado conjuntos difusos pero existe una gran incertidumbre asociada a la construcción de los valores de pertenencia. En particular, en el futuro tenemos intención de aplicar estas nociones en problemas de de visión en estéreo.

REFERENCES

- [1] E. Barrenechea, H. Bustince, B. De Baets, C. Lopez-Molina, Construction of interval-valued fuzzy relations with application to the generation of fuzzy edge images IEEE Transactions on Fuzzy Systems 19 (5) (2011) 819–830.
- [2] E. Barrenechea, J. Fernandez, M. Pagola, F. Chiclana, H. Bustince, Construction of interval-valued fuzzy preference relations from ignorance functions and fuzzy preference relations. Application to decision making Knowledge-Based Systems 58 (2014) 33–44.
- [3] P. Burillo, H. Bustince, Construction theorems for intuitionistic fuzzy sets, Fuzzy Sets and Systems, 84 (1996) 271–281.
- [4] H. Bustince, E. Barrenechea, M. Pagola, Relationship between restricted dissimilarity functions, restricted equivalence functions and normal E_N -functions: Image thresholding invariant, Pattern Recognition Letters 29(4) (2008) 525–536.
- [5] P. Burillo, H. Bustince, Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets, Fuzzy Sets and Systems 78 (1996) 305–316.
- [6] H. Bustince, E. Barrenechea, M. Pagola, Image thresholding using restricted equivalence functions and maximizing the measure of similarity, Fuzzy Sets and Systems 128(5) (2007) 496–516.
- [7] H. Bustince, E. Barrenechea, M. Pagola, Restricted equivalence functions, Fuzzy Sets and Systems 157(17) (2006) 2333–2346.



- [8] H. Bustince, E. Barrenechea, M. Pagola, J. Fernández, Z. Xu, B. Bedregal, J. Montero, H. Hagra, F. Herrera, B. De Baets, A historical account of types of fuzzy sets and their relationship, *IEEE Transactions on Fuzzy Systems*, 24 (1) (2016) 179–194.
- [9] H. Bustince, J. Fernandez, , A. Kolesárová, R. Mesiar, Generation of linear orders for intervals by means of aggregation functions, *Fuzzy Sets and Systems* 220 (2013) 69–77.
- [10] H.M. Choi, G.S. Mun, J.Y. Ahn, A medical diagnosis based on interval-valued fuzzy sets, *Biomedical Engineering-Applications Basis Communications*, 24(4) (2012) 349–354.
- [11] A. Heidarzade, A new similarity measure for interval type-2 fuzzy sets: Application in fuzzy risk analysis, *International Journal of Applied Decision Sciences*, 9 (4) (2016) 400–412.
- [12] X. Liu, Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets Systems* 52 (1992) 305–318.
- [13] Z. Lu, and J. Ye, Logarithmic similarity measure between interval-valued fuzzy sets and its fault diagnosis method, *Information (Switzerland)* 9 (2018) art. no. 36.
- [14] Z.S. Xu, R.R. Yager, Some geometric aggregation operators based on intuitionistic fuzzy sets, *International Journal of General Systems* 35 (2006) 417–433.



Una generalización del algoritmo gravitacional de clústering utilizando funciones de overlap

I. Rodríguez, J. Ruiz-Aranguren, J. Fernández, H. Bustince
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra, Pamplona
bustince@unavarra.es

Javier Armentia
Planetario de Pamplona
Pamplona, España

Abstract—En este trabajo consideramos una modificación del algoritmo de clúster gravitacional propuesto por Wright y utilizamos el concepto de función de overlap para mejorar su eficacia. Nuestra generalización recupera el algoritmo original. Comparamos los resultados de nuestra propuesta con los obtenidos con algoritmos usuales como el K-means o el FCM, con el objetivo de descubrir los puntos fuertes y las carencias de la nueva versión.

Keywords—Clúster; overlap; fuerza gravitacional

I. INTRODUCCIÓN

La variedad en los problemas a tratar por las técnicas de clúster ha fomentado la aparición de decenas de algoritmos distintos, capaces de adaptarse a distintas situaciones y ofrecer resultados óptimos para cada caso [1]. Uno de estos algoritmos es el propuesto por W. E. Wright en 1977 [2]. Este algoritmo trata cada uno de los ejemplos sobre los que trabaja como una partícula que afecta al resto del sistema, según una adaptación de la ley de gravitación universal de Newton. Cada partícula compone inicialmente un clúster, que se van combinando a lo largo del tiempo hasta que solo queda uno.

El objetivo de este trabajo es generalizar el algoritmo de Wright utilizando funciones de overlap [3]. Nuestra generalización modifica la función fundamental que simula el movimiento de las partículas sustituyendo un término concreto por una función de overlap. Las funciones de overlap se introdujeron inicialmente en el campo difuso para tratar con el problema de determinar a cuál de dos conjuntos difusos pertenece en mayor medida una entrada dada. Sin embargo, desde sus orígenes se han aplicado con éxito en campos como los del procesamiento de imagen [4], la clasificación [5, 6] o la optimización [7].

En particular, en [7], una generalización muy competitiva del GSA, un algoritmo de optimización que también hace uso de la fuerza gravitacional, ver [8], fue propuesto, reemplazando el producto en la expresión de la fuerza por funciones de overlap (y otro tipo de funciones).

La estructura del trabajo es la siguiente. En la Sección II presentamos algunos conceptos preliminares. En la Sección III, presentamos el algoritmo de clúster gravitacional original y nuestra propuesta de generalización. La Sección IV se dedica a un estudio experimental de nuestra propuesta. Terminamos con algunas conclusiones y referencias.

II. PRELIMINARES

A. Funciones de Agregación

Definición. [9] Una función de agregación n-dimensional es una función $M:[0,1]^n \rightarrow [0,1]$ que verifica las dos condiciones siguientes:

- M es no decreciente en cada una de sus variables;
- $M(0, \dots, 0) = 0$ y $M(1, \dots, 1) = 1$.

Definición. Sea $M:[0,1]^n \rightarrow [0,1]$ una función de agregación n-dimensional.

- $a \in [0,1]$ es un aniquilador de M si $M(x_1, \dots, x_n) = a$ siempre que $a \in \{x_1, \dots, x_n\}$.
- Si M no tiene aniquilador, M se dice estrictamente creciente si lo es en el dominio $[0,1]^n$ como función real de n variables. Si a es un aniquilador de M, si es estrictamente creciente en el dominio $([0, 1] \setminus \{a\})^n$.
- M tiene divisores de cero si existen $x_1, \dots, x_n \in [0, 1]$ tales que $M(x_1, \dots, x_n) = 0$.
- M es idempotente si $M(x_1, \dots, x_n) = x$ para todo $x \in [0, 1]$.

Nos vamos a centrar en el caso de funciones de agregación en dimensión 2. Por ello, revisamos algunas propiedades que nos van a resultar de interés.

Definición. Sea M una función de agregación de dos variables.

- M es simétrica si $M(x, y) = M(y, x)$ para cualesquiera $x, y \in [0, 1]$.
- M es asociativa si $M(M(x, y), z) = M(x, M(y, z))$ para cualesquiera $x, y, z \in [0, 1]$.

La asociatividad permite extender las funciones de agregación bidimensionales a dimensiones mayores de una manera “razonable”. En particular, es una de las propiedades fundamentales que se demandan en la construcción de normas triangulares (t-normas).

Definición . Una norma triangular o t-norma es una función de agregación $T:[0, 1]^2 \rightarrow [0, 1]$ asociativa y simétrica tal que $T(1, x) = x$ para todo $x \in [0, 1]$.

Las t-normas permiten modelar intersecciones entre conjuntos difusos. Ahora bien, si estamos trabajando con solo dos elementos, la asociatividad no es en principio una propiedad que deba ser requerida. Esta idea, la de poder representar la intersección sin necesariamente requerir la asociatividad nos lleva al concepto de función de solapamiento (función de overlap) que estudiamos a continuación.

B. Funciones de Overlap

En esta subsección recordamos algunos conceptos y resultados presentados en [3].

Definición Una función $O:[0, 1]^2 \rightarrow [0, 1]$ es una función de solapamiento si

- O es simétrica;
- $O(x, y) = 0$ si y solo si $xy = 0$;
- $O(x, y) = 1$ si y solo si $xy = 1$;
- O es no decreciente;
- O es continua.

Podemos presentar el siguiente resultado sobre la relación entre funciones de solapamientos y t-normas.

Teorema. Sea O una función de solapamiento asociativa. Entonces O es una t-norma.

En cuanto a la construcción de funciones de overlap, es posible hacerlo de forma general tal y como muestra el siguiente resultado.

Teorema. La función $O:[0, 1]^2 \rightarrow [0, 1]$ es una función de solapamiento si y solo si

$$O(x, y) = f(x, y) / (f(x, y) + h(x, y))$$

Con f, h: $[0, 1]^2 \rightarrow [0, 1]$ tales que

- f y h son simétricas;
- f es no decreciente y h es no creciente;
- $f(x, y) = 0$ si y solo si $xy = 0$;
- $h(x, y) = 0$ si y solo si $xy = 1$;
- f y h son continuas;

III. EL ALGORITMO DE CLÚSTER GRAVITACIONAL

A. El Algoritmo Original

El algoritmo de clúster gravitacional pretende hacer uso de la Ley de Gravitación Universal de Newton para llevar a cabo un proceso de clasificación no supervisado [2].

El esquema del algoritmo, tal y como figura en [2], es el siguiente:

Supongamos que tenemos n partículas p_1, \dots, p_n , para las que conocemos sus posiciones $s_1, \dots, s_n \in \mathbb{R}^n$. Asignamos a cada partícula p_i una masa $m_i = 1$ para comenzar el algoritmo.

1. Fijamos dos parámetros reales positivos δ y ϵ .
 - Utilizamos δ para determinar la longitud de cada paso temporal dt. En concreto, en el intervalo $[t, t + dt]$ la partícula que se mueva más deprisa debe desplazarse una distancia δ .
 - Si dos partículas están en algún momento a distancia menor que ϵ , las unimos en una partícula con masa igual a la suma de las masas original y posición dada por el centro de masas de las partículas originales.
2. Inicializamos el tiempo con $t = 0$.
3. Repetimos los pasos (i)-(iv) siguientes hasta que solo quede una partícula.

- (i). En cada intervalo de tiempo $[t, t + dt]$, para cada partícula i calculamos el efecto gravitacional de las demás partículas sobre ella. En concreto:

$$g(i, t, dt) = \frac{1}{2} G \sum_{j \neq i} \frac{m_i(t)m_j(t)}{m_i(t)} \frac{s_j(t) - s_i(t)}{|s_j(t) - s_i(t)|} \frac{1}{|s_j(t) - s_i(t)|^2} dt^2 \tag{1}$$

donde G es una constante positiva.

- (ii). Para cada partícula i, su nueva posición es: $s_i(t + dt) = s_i(t) + g(i, t, dt)$.
- (iii). Incrementamos el tiempo de t a t + dt.
- (iv). Si dos partículas i y j están a distancia (euclídea) menor que ϵ , se funden en una partícula como se ha explicado anteriormente.

Al final tenemos una sola partícula y, si el proceso ha durado un tiempo T, la siguiente evolución:

- De $t_n = 0$ a t_{n-1} había n partículas.
- De t_{n-1} a t_{n-2} había n-1 partículas.
- ...
- De t_3 a t_2 había 3 partículas.
- De t_2 a t_1 había 2 partículas.

Definimos la vida relativa de la configuración de k clusters como

$$R_k = (t_k - t_{k-1}) / T$$

Tomamos como solución aquella configuración R_{k_0} con la mayor vida relativa.

Nótese que en una iteración dada puede haber varias fusiones de partículas. El caso en que en t_k se tienen k partículas y en t_{k-1} pasa a haber k-2 partículas, se puede ver como que de t_k a t_{k-1} había k partículas, de t_{k-1} a t_{k-2} pasa a haber k-1 partículas y $t_{k-1} = t_{k-2}$, luego $R_{k-1} = 0$.

Este modelo puede generalizarse reemplazando (1) por



$$g(i, t, dt) = \sum_{j \neq i} \frac{m_i(t)^p m_j(t)^q}{m_i(t)} \frac{s_j(t) - s_i(t)}{|s_j(t) - s_i(t)| |s_j(t) - s_i(t)|^2} dt^2 \quad (2)$$

con $p, q \geq 0$. En particular, los mejores resultados se obtienen cuando $p = q = 0$, obteniéndose el denominado modelo unitario de Markov

$$g(i, t, dt) = \frac{1}{m_i(t)} \sum_{j \neq i} \frac{s_j(t) - s_i(t)}{|s_j(t) - s_i(t)| |s_j(t) - s_i(t)|^2} dt^2 \quad (3)$$

Nótese que la constante $G/2$ se ha eliminado por tratarse únicamente de un valor heredado a partir de la ley de gravitación universal de Newton que no tiene un impacto útil en nuestro algoritmo, más allá del de escalar los valores obtenidos por la función g .

B. Generalización Usando Overlaps

En esta sección presentamos el contenido principal del trabajo: un nuevo algoritmo de clúster gravitacional que generaliza el mostrado en la sección anterior, reemplazando el factor $1/m_i(t)$ en la ecuación (3) por un término de la forma $O(1/m_i(t), 1)$ donde $O: [0,1]^2 \rightarrow [0,1]$ es una función de overlap general. Ahora bien, dado que la función de overlap está definida sobre el cuadrado unidad $[0,1]^2$, en principio es necesaria una normalización. Esto puede hacerse considerando la función $O(m_i^{p-1}/n^{p-1}, m_j^q/n^q)$, con $p, q > 0$.

Sin embargo, dado que estamos interesados en el caso p, q próximos a cero (que es el que mejor funciona en el trabajo original), se tiene que esta función siempre es mayor o igual que $O(1/m_i(t), 1)$. La ecuación final a utilizar será la siguiente:

$$g(i, t, dt) = O\left(\frac{1}{m_i(t)}, 1\right) \sum_{j \neq i} \frac{s_j(t) - s_i(t)}{|s_j(t) - s_i(t)| |s_j(t) - s_i(t)|^2} dt^2 \quad (4)$$

Los overlaps que ofrecen resultados más prometedores, y que por tanto se utilizarán en la fase de experimentación son los siguientes:

$$O(x, y) = (x * y)^p \rightarrow O\left(\frac{1}{m_i}, 1\right) = \frac{1}{m_i^p} \quad (5)$$

$$O(x, y) = \frac{\sqrt{x}}{\sqrt{x} + \max(1-x, 1-y)} \rightarrow O\left(\frac{1}{m_i}, 1\right) = \frac{\sqrt{\frac{1}{m_i}}}{\sqrt{\frac{1}{m_i}} + \max\left(1 - \frac{1}{m_i}, 0\right)} \quad (6)$$

IV. EXPERIMENTACIÓN

A. Pruebas

Para las pruebas se utilizaron una serie de datasets sintéticos bidimensionales con el fin de comprobar en qué casos se obtienen los mejores resultados para cada uno de los overlaps. Los datasets utilizados se representan visualmente en las Figs. 1-6.

Las pruebas se realizaron utilizando tanto métricas de evaluación interna (Dunn-Index [10]) como de evaluación externa (Purity, Rand-Index [11]). Se compararon los resultados del algoritmo empleando cada uno de los overlaps

con el modelo unitario de Markov, el algoritmo K-means [12] y el Fuzzy c-Means (FCM) [13].

Para estos dos últimos algoritmos se realizaron 100 ejecuciones y se tomaron los mejores valores obtenidos, buscando evitar el problema de caer en óptimos locales. Para el algoritmo gravitacional se asignaron valores tomados de 0.01 en 0.01 en $[0.1, 1]$ para δ , y de $\epsilon = 2\delta$. Como parámetro p para (5) se utilizó el valor 20, con el objetivo de que las masas grandes se vieran muy penalizadas. En las tablas se reflejan tanto los valores medios obtenidos tras todas las ejecuciones, como los valores máximos. La columna izquierda hace referencia al dataset para el cuál se han obtenido los resultados (Figs. 1-6).

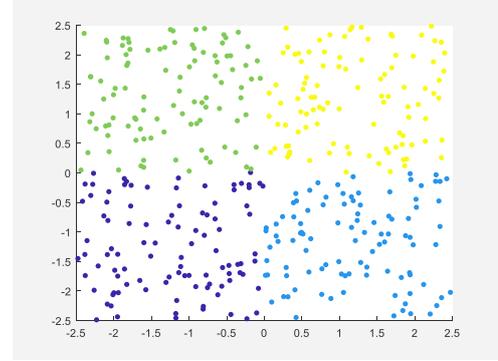


Fig. 1. Dataset "4 squares"

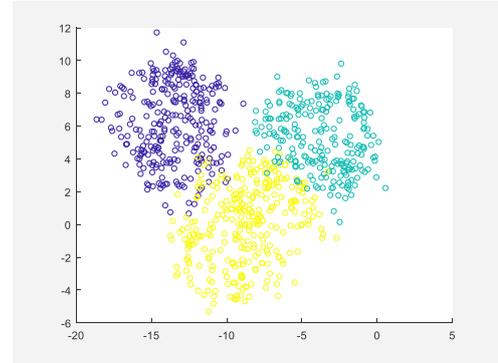


Fig. 2. Dataset "3 close circles"

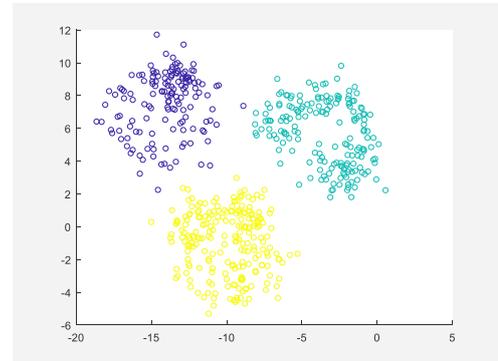


Fig. 3. Dataset "3 circles"

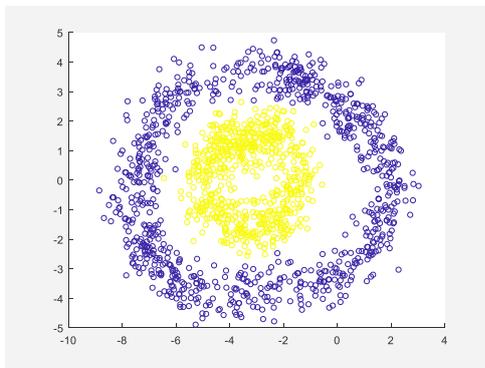


Fig. 4. Dataset "2 concentric rings"

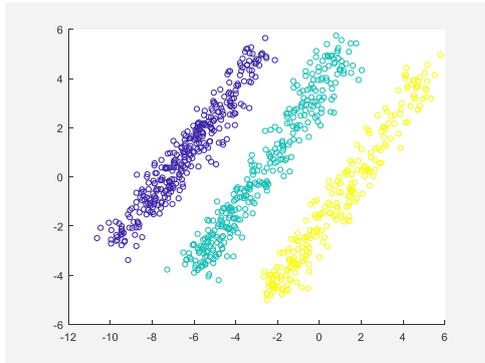


Fig. 5. Dataset "3 stripes"

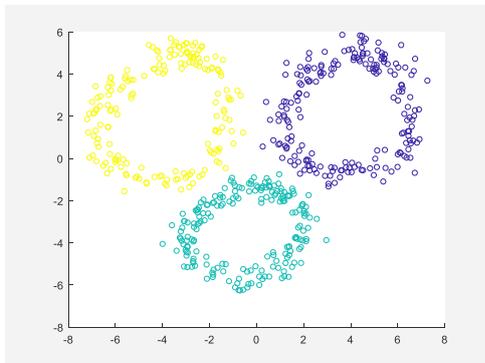


Fig. 6. Dataset "3 rings"

	Dunn-Index							
	(5)		(6)		Modelo unitario		K-means	FCM
	Mean	Max	Mean	Max	Mean	Max		
1	0,508	0,682	0,577	0,682	0,578	0,682	0,743	0,736
2	0,633	0,690	0,621	0,691	0,641	0,691	0,688	0,685
3	0,848	0,861	0,859	0,860	0,858	0,860	0,932	0,932
4	0,386	0,464	0,468	0,497	0,464	0,497	0,464	0,438
5	0,378	0,579	0,442	0,596	0,432	0,586	0,467	0,485
6	0,519	0,831	0,631	0,758	0,621	0,715	0,874	0,848

^a. Tabla 1 – Comparación de índice Dunn entre los algoritmos para los datasets 1-6

	Purity							
	(5)		(6)		Modelo unitario		K-means	FCM
	Mean	Max	Mean	Max	Mean	Max		
1	0,784	0,893	0,807	0,910	0,810	0,900	0,988	0,983
2	0,859	0,946	0,856	0,948	0,852	0,920	0,901	0,901
3	0,991	0,998	0,996	0,998	0,996	0,998	0,998	0,998
4	0,577	0,628	0,622	0,713	0,624	0,705	0,570	0,570
5	0,670	0,928	0,564	0,642	0,569	0,690	0,697	0,690
6	0,806	0,998	0,788	0,916	0,784	0,916	0,960	0,960

^b. Tabla 2 – Comparación de Purity entre los algoritmos para los datasets 1-6

	Rand-Index							
	(5)		(6)		Modelo unitario		K-means	FCM
	Mean	Max	Mean	Max	Mean	Max		
1	0,822	0,901	0,834	0,917	0,838	0,907	0,988	0,983
2	0,843	0,931	0,840	0,931	0,836	0,902	0,884	0,884
3	0,989	0,998	0,995	0,998	0,994	0,998	0,998	0,998
4	0,507	0,532	0,533	0,591	0,533	0,584	0,501	0,501
5	0,718	0,905	0,626	0,702	0,633	0,719	0,762	0,754
6	0,812	0,998	0,741	0,900	0,735	0,900	0,949	0,949

^c. Tabla 3 – Comparación de índice Rand entre los algoritmos para los datasets 1-6

B. Resultados

Lo primero que se puede observar es que no siempre existe una relación directa entre los mejores resultados para la evaluación interna y los de la externa. Esto supone un problema, dado que dificulta la búsqueda de los parámetros óptimos para el algoritmo en los casos en que carecemos de información sobre las clasificaciones a priori, que serán la mayoría.

Si nos fijamos en las medidas de evaluación externa, sin embargo, vemos que el modelo usando (6) tiende a mejorar el mejor resultado del modelo clásico, mientras que suele ofrecer valores similares en media. Por tanto, parece una alternativa razonable al mismo. En concreto, obtiene los mejores resultados de cualquier método cuando opera con datasets circulares, del mismo modo que cuando lo hace con el dataset de la Fig. 4, aunque en este caso los resultados son mediocres.

En el caso del overlap indicado en (5) tenemos que se obtienen valores con mucha varianza para la mayoría de datasets. No obstante, se obtienen resultados más que aceptables al trabajar con los dos últimos, casos en los que ninguna de las alternativas del algoritmo ofrece buenos resultados. Esto parece deberse a la gran penalización de las partículas con masas mayores. El modo de trabajo del algoritmo asegura que en todas las iteraciones, una partícula se desplazará la distancia δ . En este caso también sucede así, pero el resto de partículas de masa mayor se desplazan distancias ínfimas, de modo que cuando se tiene un número elevado de partículas, en cada iteración tiende a moverse una sola, obteniéndose un funcionamiento similar al del algoritmo "single-linkage" [15], que opera bien con conjuntos de datos de esta forma.



V. CONCLUSIONES Y FUTUROS TRABAJOS

Las pruebas realizadas parecen apuntar a una mejoría del funcionamiento del algoritmo cuando se utilizan los overlaps propuestos en ciertos casos concretos. Además, lo dotan de una mayor capacidad de adaptación a conjuntos de datos de forma no circular, problema con el que no pueden tratar otro tipo de técnicas.

En cualquier caso, resulta importante recalcar que nuestro caso sigue siendo una generalización del algoritmo original, por lo que siempre se puede recuperar el funcionamiento clásico.

Esta generalización abre la puerta a un estudio en mayor profundidad que permita extender la generalización empleando funciones de overlap a funciones de otro tipo, consiguiendo posiblemente comportamientos mejores. Del mismo modo, puede haber alternativas a la distancia euclidiana a tener en cuenta a la hora de calcular la distancia entre partículas. Ambas posibilidades se tendrán en cuenta para futuros trabajos.

Este trabajo ha sido parcialmente financiado por el proyecto de investigación TIN2016-77356-P (MINECO/UE/FEDER) del Gobierno de España.

REFERENCIAS

- [1] A. Saxena, M. Prasad, A. Gupta, N. Bharill, O. Prakash Patel, A. Tiwari, E. M. Joo, D. Weiping, L. Chin-Teng, "A review of clustering techniques and developments," in *Neurocomputing*, vol. 267, pp.664-681, 2017
- [2] W. E. Wright, "Gravitational clustering", in *Pattern Recognition*, Pergamon Press 9, 1977, pp.151-166
- [3] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, R. Orduna, "Overlap functions," in *Nonlinear Analysis*, vol. 72, pp.1488-1499, 2010
- [4] A. Jurio, H. Bustince, M. Pagola, A. Pradera, R. R. Yager, "Some properties of overlap and grouping functions and their application to image thresholding," in *Fuzzy Sets and Systems*, vol. 229, pp.69-90, 2013
- [5] D. Gómez, J. T. Rodríguez, J. Montero, H. Bustince, E. Barrenechea, "n-Dimensional overlap functions," in *Fuzzy Sets and Systems*, vol. 287, pp.57-75, 2016
- [6] D. Paternain, H. Bustince, M. Pagola, P. Sussner, A. Kolesrov, R. Mesiar, "Capacities and overlap indexes with an application in fuzzy rule-based classification systems," in *Fuzzy Sets and Systems*, vol. 305, pp.70-94, 2016
- [7] M. Minárová, D. Paternain, A. Jurio, J. Ruiz-Aranguren, Z. tak, H. Bustince, "Modifying the gravitational search algorithm: A functional study," in *Information Sciences*, vol. 430, pp.87-103, 2018
- [8] E. Rashedi, H. Neamabadi-Pour, S. Sariazdi, "GSA: a gravitational search algorithm," in *Information Sciences*, vol 179, pp.2232-2248, 2009
- [9] M. Grabisch, J. Marichal, R. Mesiar, E. Pap. *Aggregation Functions*, Encyclopedia of Mathematics and its Applications, vol. 127, Cambridge University Press, Cambridge, 2009.
- [10] J. C. Dunn, "Well-separated clusters and optimal fuzzy partitions," in *Journal of Cybernetics*, vol. 4, pp.95-104, 1973
- [11] D. L. Davies, D. W. Bouldin, "A cluster separation measure," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, pp.224-227, 1979
- [12] D. Arthur, S. Vassilvitskii, "K-means++: the advantages of careful seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, New Orleans, 2007
- [13] J. C. Bezdek, R. Ehrlich, W. Full, "FCM: the fuzzy c-means algorithm," in *Computer & Geosciences*, vol. 10, pp.191-203, 1984
- [14] R. Sibson, "SLINK: an optimally efficient algorithm for the single-link cluster method," in *The Computer Journal*, vol. 16, pp.30-34, 1973

Una definición de la Integral de Choquet intervalo-valorada basada en permutaciones admisibles

Daniel Paternain, Laura De Miguel, Gustavo Ochoa, Inmaculada Lizasoain, Humberto Bustince
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra, Pamplona, España
{daniel.paternain,laura.demiguel,ochoa,ilizasoain,bustince}@unavarra.es

Radko Mesiar
Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering,
Slovak University of Technology, Bratislava, Eslovaquia,
mesiar@math.sk

Abstract—La agregación o fusión de datos intervalo-valorados no es una tarea trivial. Esto es debido a que para aplicar determinadas funciones de agregación es necesario ordenar totalmente los datos a agregar. Para solucionar el problema del orden, en la literatura podemos encontrar soluciones basadas en el concepto de orden admisible. Tras analizar las ventajas e inconvenientes de esta propuesta, en este trabajo proponemos el concepto de permutación admisible y de integral de Choquet intervalo-valorada basada en permutaciones admisibles. Este nuevo operador nos permite agregar de una manera efectiva un conjunto de intervalos solucionando los problemas encontrados en las propuestas basadas en órdenes admisibles.

Index Terms—Agregación, Fusión, Integral de Choquet, intervalos, permutación admisible.

I. INTRODUCCIÓN

La fusión de información es un paso clave en multitud de aplicaciones del mundo real, dado que en casi todas ellas es necesario combinar o fusionar en algún momento varias fuentes de información en un único valor representativo [2], [3], [11]. Ejemplos de dichas aplicaciones son la toma de decisiones multi-experto y multi-criterio, la fusión de sensores, el procesamiento digital de imágenes, la minería de datos o el aprendizaje automático, entre otros.

Cuando el proceso de fusión de información tiene un alto grado de incertidumbre asociada a los datos, las extensiones de los conjuntos difusos pueden ser muy útiles para manejar dicha incertidumbre [1], [4], [8]. Algunas extensiones que han sido muy utilizadas en la literatura son los conjuntos intervalo-valorados difusos, los conjuntos intuicionistas de Atanassov o los tipo-2. En este trabajo nos centramos en la fusión de información intervalo-valorada, ya que con una complejidad reducida, los intervalos permiten tratar la incertidumbre de manera efectiva. Ejemplos de la utilización de información intervalo-valorada pueden ser la información proveniente de los sensores, que es convertida a un intervalo de confianza teniendo en cuenta la precisión del propio instrumento; o en toma de decisiones, los intervalos pueden ser una herramienta

muy útil para los expertos cuando éstos no son capaces de asignar una puntuación exacta a una alternativa/criterio.

Sin embargo, la extensión de algunas funciones de agregación para manejar información intervalo-valorada no es trivial. Concretamente, tenemos en mente funciones como los operadores OWA o la integral de Choquet, donde es necesario ordenar los datos de entrada antes de ser fusionados. El problema surge debido a que el orden natural entre intervalos es un orden parcial y, por tanto, existen elementos incomparables que no pueden ser ordenados.

Para solucionar el problema del orden en las funciones de agregación intervalo-valoradas, en [5] se propuso una solución basada en la construcción de órdenes admisibles a partir de una pareja de funciones. Basándose en el concepto de orden admisible, en [6] se propuso una nueva definición de la integral de Choquet intervalo-valorada en la que los datos se ordenaban mediante la elección de un orden admisible. Sin embargo, aunque el problema del orden estaba solucionado, surge un nuevo problema: la elección del orden admisible más apropiado. Parece razonable pensar que si en una aplicación elegimos un mal orden admisible, los resultados de la fusión de información pueden ser contraproducentes para el resultado final de la aplicación. Se podría pensar que probando muchos órdenes admisibles se soluciona dicho problema, pero debemos tener en cuenta que: (1) podemos definir infinitos órdenes admisibles y (2) muchos órdenes admisibles son equivalentes entre sí, produciendo la misma ordenación de los datos.

Para solucionar el problema de la elección de un orden admisible concreto, en este trabajo proponemos el concepto de permutación admisible. Este concepto nos permite identificar de cuántas maneras diferentes podemos ordenar un conjunto finito de intervalos respetando en todo momento las restricciones impuestas por el orden parcial. Así, si tenemos en cuenta únicamente las permutaciones admisibles de un conjunto de intervalos podemos aislarnos de la elección de un orden admisible concreto. Basándonos en este concepto,



en este trabajo proponemos además una nueva aproximación para la integral de Choquet intervalo-valorada que tiene en cuenta todas las formas posibles de ordenar el conjunto de intervalos a ordenar y obtiene el resultado final realizando un promedio de todas ellos.

Este trabajo se organiza de la siguiente manera: en la Sección 2 recordamos los conceptos preliminares necesarios para entender el resto del trabajo. En la Sección 3 recordamos el concepto de orden admisible y de integral de Choquet basada en órdenes admisibles. En la Sección 4 proponemos el concepto de permutación admisible y de integral de Choquet basada en permutaciones admisibles y estudiamos algunas de sus propiedades. Finalmente, comentamos las conclusiones y líneas futuras en la Sección 5.

II. PRELIMINARES

Comenzamos esta sección recordando el concepto de función de agregación en conjuntos parcialmente ordenados.

Definición 1: [12] Sea (L, \preceq) un conjunto parcialmente ordenado acotado con un elemento mínimo 0_L y un elemento máximo 1_L . Una función $M : L^n \rightarrow L$ es una función de agregación si satisface las siguientes propiedades:

- (i) $M(0_L, \dots, 0_L) = 0_L$ y $M(1_L, \dots, 1_L) = 1_L$;
- (ii) es creciente en cada argumento, es decir, para todo $(x_1, \dots, x_n), (y_1, \dots, y_n) \in L^n$, $M(x_1, \dots, x_n) \preceq M(y_1, \dots, y_n)$ si $x_1 \preceq y_1, \dots, x_n \preceq y_n$.

Obsérvese que si L es el intervalo unidad junto con el orden habitual de los números reales, entonces obtenemos la definición usual de función de agregación [2], [3], [11].

En este trabajo nos centramos en la agregación de información intervalo-valorada. Por eso, tomamos $L = L([0, 1])$ como el conjunto de todos los subintervalos cerrados del intervalo $[0, 1]$:

$$L([0, 1]) = \{\mathbf{x} = [x, \bar{x}] \mid 0 \leq x \leq \bar{x} \leq 1\}.$$

Nótese que $L([0, 1])$ es un conjunto parcialmente ordenado con respecto a la relación de orden \leq_L definida de la siguiente manera: para cada $\mathbf{x}, \mathbf{y} \in L([0, 1])$,

$$\mathbf{x} \leq_L \mathbf{y} \text{ y si y solo si } \underline{x} \leq \underline{y} \text{ y } \bar{x} \leq \bar{y}.$$

De hecho, $(L([0, 1]), \leq_L)$ es un retículo completo donde el elemento mínimo es $0_L = [0, 0]$ y el máximo es $1_L = [1, 1]$ ([10]). En este retículo, el ínfimo y supremo de cada pareja de elementos viene dado, respectivamente, por

$$\begin{aligned} \mathbf{x}_1 \wedge \mathbf{x}_2 &= [\min(\underline{x}_1, \underline{x}_2), \min(\bar{x}_1, \bar{x}_2)] \\ \mathbf{x}_1 \vee \mathbf{x}_2 &= [\max(\underline{x}_1, \underline{x}_2), \max(\bar{x}_1, \bar{x}_2)]. \end{aligned}$$

Ejemplo 1: Los siguientes son ejemplos de funciones de agregación intervalo-valorada:

$$\begin{aligned} \mathbf{M}_{arith}(\mathbf{x}_1, \dots, \mathbf{x}_n) &= \left[\frac{1}{n} \sum_{i=1}^n \underline{x}_i, \frac{1}{n} \sum_{i=1}^n \bar{x}_i \right]; \\ \mathbf{M}_{min}(\mathbf{x}_1, \dots, \mathbf{x}_n) &= \left[\min_{i=1, \dots, n} \underline{x}_i, \min_{i=1, \dots, n} \bar{x}_i \right]; \\ \mathbf{M}_{max}(\mathbf{x}_1, \dots, \mathbf{x}_n) &= \left[\max_{i=1, \dots, n} \underline{x}_i, \max_{i=1, \dots, n} \bar{x}_i \right]. \end{aligned}$$

A. Medidas difusas y la integral de Choquet

Antes de definir el concepto de integral de Choquet, recordamos el concepto de medida difusa (ver [14], [15]).

Definición 2: Sea $X = \{1, \dots, n\}$. Una medida difusa definida sobre X es una función $m : 2^X \rightarrow [0, 1]$ tal que

- (i) $m(\emptyset) = 0$ y $m(X) = 1$;
- (ii) si $E \subset F$, entonces $m(E) \leq m(F)$.

Ejemplo 2:

- La medida difusa más pequeña viene dada por

$$m_*(E) = \begin{cases} 1 & \text{si } E = X; \\ 0 & \text{en otro caso.} \end{cases}$$

Obsérvese que para cualquier medida difusa m sobre X se cumple que $m_*(E) \leq m(E)$, para todo $E \subseteq X$.

- La medida difusa más grande viene dada por

$$m^*(E) = \begin{cases} 0 & \text{si } E = \emptyset; \\ 1 & \text{en otro caso.} \end{cases}$$

Obsérvese que para cualquier medida difusa m sobre X , se cumple que $m(E) \leq m^*(E)$ para todo $E \subseteq X$.

Definición 3: [9] Sea $m : 2^X \rightarrow [0, 1]$ una medida difusa. La integral de Choquet discreta de $x_1, \dots, x_n \in [0, 1]$ con respecto a m viene dada por

$$C_m(x_1, \dots, x_n) = \sum_{i=1}^n x_{\sigma(i)} (m(\{\sigma(i), \dots, \sigma(n)\}) - m(\{\sigma(i+1), \dots, \sigma(n)\})) \quad (1)$$

donde $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ es una permutación tal que $x_{\sigma(1)} \leq \dots \leq x_{\sigma(n)}$ y donde tomamos $\{x_{\sigma(n+1)}, x_{\sigma(n)}\} = \emptyset$.

III. EL CONCEPTO DE ORDEN ADMISIBLE Y LA INTEGRAL DE CHOQUET BASADA EN ÓRDENES ADMISIBLES

Como hemos visto en la Sección 2, el orden \leq_L definido en $L([0, 1])$ es un orden parcial. Esto significa que no siempre es posible comparar (ordenar) dos intervalos cualesquiera. Sin embargo, sabemos que muchas funciones de agregación, como la integral de Choquet, los operadores OWA o la integral de Sugeno, están basadas en la ordenación total de los datos de entrada. De esta manera, si queremos extender dichas funciones para agregar información intervalo-valorada, necesitamos solucionar el problema del orden.

Una primera solución a este problema fue dada en [5] mediante el concepto de orden admisible. Un orden admisible

es un orden lineal definido en $L([0, 1])$ que refina el orden parcial \leq_L .

Definición 4: [5] Sea $(L([0, 1]), \preceq)$ un conjunto parcialmente ordenado. Decimos que el orden \preceq definido en $L([0, 1])$ es un orden admisible si

- (i) \preceq es un orden lineal en $L([0, 1])$;
- (ii) para todo $\mathbf{x}, \mathbf{y} \in L([0, 1])$ con $\mathbf{x} \leq_L \mathbf{y}$, se tiene $\mathbf{x} \preceq \mathbf{y}$.

Ejemplo 3: Los siguientes son ejemplos de órdenes admisibles:

- (i) $\mathbf{x} \preceq_{Lex1}$ y (orden lexicográfico habitual en \mathcal{R}^2) si y solo si $\underline{x} < \underline{y}$ o $(\underline{x} = \underline{y} \text{ y } \bar{x} \leq \bar{y})$;
- (ii) $\mathbf{x} \preceq_{Lex2}$ y si y solo si $\bar{x} < \bar{y}$ o $(\bar{x} = \bar{y} \text{ y } \underline{x} \leq \underline{y})$;
- (iii) $\mathbf{x} \preceq_{XY}$ y (orden de Xu-Yager dado en [17]) si y solo si $\underline{x} + \bar{x} < \underline{y} + \bar{y}$ o $(\underline{x} + \bar{x} = \underline{y} + \bar{y} \text{ y } \bar{y} - \underline{y} \leq \bar{x} - \underline{x})$.

Una de las posibles formas de construcción de órdenes admisibles es mediante la utilización de dos funciones de agregación en $[0, 1]$ que satisfacen determinadas propiedades. Sea $K([0, 1]) = \{(\underline{x}, \bar{x}) \in [0, 1]^2 | \underline{x} \leq \bar{x}\}$.

Proposición 1: [5] Sean $A, B : [0, 1]^2 \rightarrow [0, 1]$ dos funciones de agregación definidas en $[0, 1]$ tales que, para todo $(x, y), (u, v) \in K([0, 1])$, las igualdades $A(x, y) = A(u, v)$ y $B(x, y) = B(u, v)$ se cumplen simultáneamente solo si $(x, y) = (u, v)$. Definamos la relación $\preceq_{A,B}$ en $L([0, 1])$ dada por $\mathbf{x} \preceq_{A,B} \mathbf{y}$ si y solo si

$$A(\underline{x}, \bar{x}) < A(\underline{y}, \bar{y}) \quad \text{o} \quad (2)$$

$$(A(\underline{x}, \bar{x}) = A(\underline{y}, \bar{y}) \quad \text{y} \quad B(\underline{x}, \bar{x}) \leq B(\underline{y}, \bar{y})). \quad (3)$$

Entonces, $\preceq_{A,B}$ es un orden admisible de $L([0, 1])$.

A. La integral de Choquet basada en órdenes admisibles

A raíz de la obtención de un método de construcción de órdenes admisibles entre intervalos, en [6] los autores propusieron una nueva definición de operador OWA y de integral de Choquet intervalo-valorados basados en órdenes admisibles. La principal característica de esta nueva aproximación es que los datos de entrada son ordenados en función de un orden admisible fijado previamente. En este sentido, si fijamos los datos a agregar, podemos obtener tantas integrales de Choquet (operadores OWA) diferentes como órdenes admisibles.

Definición 5: Sea $\mathbf{x}_1, \dots, \mathbf{x}_n \in L([0, 1])$ y sea $m : 2^X \rightarrow [0, 1]$ una medida difusa. La integral de Choquet respecto al orden admisible $\preceq_{A,B}$, con notación $\mathbf{C}_m^{\preceq_{A,B}}$, viene dada por

$$\mathbf{C}_m^{\preceq_{A,B}}(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i=1}^n \mathbf{x}_{\sigma_{A,B}(i)} (m(\{\sigma_{A,B}(i), \dots, \sigma_{A,B}(n)\}) - m(\{\sigma_{A,B}(i+1), \dots, \sigma_{A,B}(n)\}))$$

donde $\sigma_{A,B} : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ es una permutación tal que

$$\mathbf{x}_{\sigma_{A,B}(1)} \preceq_{A,B} \dots \preceq_{A,B} \mathbf{x}_{\sigma_{A,B}(n)}$$

y donde tomamos $\{\sigma_{A,B}(n), \sigma_{A,B}(n+1)\} = \emptyset$.

Una de las principales ventajas de esta nueva definición es el hecho de que $\mathbf{C}_m^{\preceq_{A,B}}$ generaliza la definición usual de integral de Choquet definida en $[0, 1]$, ya que

$$\mathbf{C}_m^{\preceq_{A,B}}([x_1, x_1], \dots, [x_n, x_n]) = [C_m(x_1, \dots, x_n), C_m(x_1, \dots, x_n)].$$

Además, si los datos de entrada son comparables con respecto al orden usual \leq_L , entonces obtenemos los mismos resultados que si aplicamos la integral de Choquet sobre los extremos inferiores y superiores.

Proposición 2: [6] Si $(\underline{x}_1, \dots, \underline{x}_n), (\bar{x}_1, \dots, \bar{x}_n)$ son comonótonos, es decir, para todo $i, j \in \{1, \dots, n\}$

$$(\underline{x}_i - \underline{x}_j)(\bar{x}_i - \bar{x}_j) \geq 0$$

entonces se cumple que $\mathbf{C}_m^{\preceq_{A,B}}(\mathbf{x}_1, \dots, \mathbf{x}_n) = \mathbf{C}_m(\mathbf{x}_1, \dots, \mathbf{x}_n)$ para cualquier pareja admisible de funciones de agregación (A, B) .

Sin embargo, aunque el problema se resuelve parcialmente con esta aproximación, también tiene una desventaja: debemos fijar un orden admisible. Recordemos que existen infinitos órdenes admisibles y que, en general, no es sencillo identificar el orden admisible apropiado para cada problema. Por estas razones, la elección del orden admisible es un punto crucial de esta aproximación que intentamos resolver mediante el concepto de permutación admisible.

IV. EL CONCEPTO DE PERMUTACIÓN ADMISIBLE Y LA INTEGRAL DE CHOQUET BASADA EN PERMUTACIONES ADMISIBLES

En esta sección proponemos una nueva aproximación a la integral de Choquet intervalo-valorada que permita resolver el problema de $\mathbf{C}_m^{\preceq_{A,B}}$. De hecho, nuestra propuesta no se basa en la elección de un orden específico, sino que trata de tener en cuenta todos los órdenes admisibles posibles. Nótese que aunque exista la posibilidad de definir infinitos órdenes admisibles, cuando ordenamos un conjunto finito de intervalos, el número de permutaciones de dichas entradas es finito. Esto es debido a que muchos órdenes admisibles diferentes producen la misma ordenación de los datos y, por tanto, son equivalentes.

Ejemplo 4: Supongamos el orden admisible $\preceq_{A,B}$ dado por $A(x, y) = 0.5x + 0.5y$ y $B(x, y) = 0.4x + 0.6y$. La pareja de funciones (A, B) es admisible y, por tanto, el orden también. Supongamos ahora una segunda pareja de funciones admisibles (A, B') donde $B'(x, y) = 0.3x + 0.7y$. Bajo estas condiciones, si $\mathbf{x}_1 \preceq_{A,B} \mathbf{x}_2$, entonces también ocurre que $\mathbf{x}_1 \preceq_{A,B'} \mathbf{x}_2$; si $\mathbf{x}_1 \prec_{A,B} \mathbf{x}_2$, debemos distinguir dos casos: (1) Si $A(\underline{x}_1, \bar{x}_1) < A(\underline{x}_2, \bar{x}_2)$, entonces también $\mathbf{x}_1 \prec_{A,B'} \mathbf{x}_2$; (2) Si $A(\underline{x}_1, \bar{x}_1) = A(\underline{x}_2, \bar{x}_2)$, entonces por la definición de B' tenemos que $B(\underline{x}_1, \bar{x}_1) < B(\underline{x}_2, \bar{x}_2)$ implica que $B'(\underline{x}_1, \bar{x}_1) < B'(\underline{x}_2, \bar{x}_2)$ y por tanto $\mathbf{x}_1 \prec_{A',B'} \mathbf{x}_2$. De esta manera podemos ver que ambos órdenes son equivalentes. De hecho, cualquier otro orden $\preceq_{(A,B')}$ dado por $B'(x, y) = \alpha x + (1 - \alpha)y$ con $\alpha < 0.5$ es equivalente al orden $\preceq_{A,B}$.

Además, es importante mencionar que, dados n intervalos, no todas las $n!$ potenciales formas de ordenar dichos intervalos son admisibles. Recordemos que los órdenes admisibles refinan el orden parcial habitual \leq_L . De esta manera, dados $\mathbf{x}_1, \dots, \mathbf{x}_n \in L([0, 1])$ y una permutación $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$, si $\mathbf{x}_i <_L \mathbf{x}_j$, entonces debemos exigir que $\sigma^{-1}(i) < \sigma^{-1}(j)$. Más aún, si $\mathbf{x}_i = \mathbf{x}_j$, entonces para cada k entre $\sigma^{-1}(i)$ y $\sigma^{-1}(j)$, debemos exigir que $\mathbf{x}_i = \mathbf{x}_{\sigma(k)} = \mathbf{x}_j$.



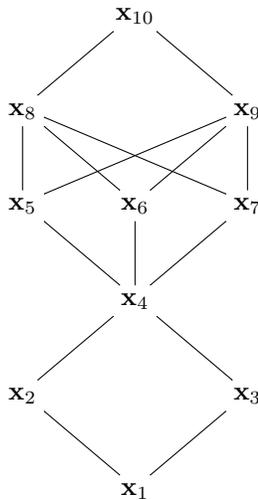
Basándonos en estas dos ideas, a continuación presentamos la definición de permutación admisible de un conjunto de intervalos.

Definición 6: Sean $\mathbf{x}_1, \dots, \mathbf{x}_n \in L([0, 1])$. Decimos que una permutación $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ de $\mathbf{x}_1, \dots, \mathbf{x}_n$ es una permutación admisible respecto al orden parcial \leq_L si

- (i) para todo $\mathbf{x}_i <_L \mathbf{x}_j$, tenemos que $\sigma^{-1}(i) < \sigma^{-1}(j)$;
- (ii) para cada \mathbf{x}_i , el conjunto $\{\sigma^{-1}(j) | j \in \{1, \dots, n\} \text{ with } \mathbf{x}_i = \mathbf{x}_j\}$ forma un intervalo en \mathbb{N} .

Obsérvese que la primera propiedad de la Definición 6 sirve para que no se altere el orden de los intervalos que sean comparables mediante el orden parcial, mientras que la segunda propiedad sirve para que los intervalos iguales sean ordenados de manera consecutiva.

Ejemplo 5: Consideremos un conjunto de 10 intervalos cuyas relaciones de orden vienen dadas en el siguiente diagrama de Hasse:



Para este conjunto de intervalos existe un total de 24 permutaciones admisibles: 2 formas de ordenar \mathbf{x}_2 y \mathbf{x}_3 , 6 formas de ordenar \mathbf{x}_5 , \mathbf{x}_6 y \mathbf{x}_7 y otras 2 formas de ordenar \mathbf{x}_8 y \mathbf{x}_9 . Finalmente, $24 = 2 \times 6 \times 2$. Nótese que las restricciones impuestas a las permutaciones admisibles reducen las potenciales $10! = 3628800$ permutaciones de 10 elementos a únicamente 24 permutaciones admisibles.

A. La integral de Choquet basada en permutaciones admisibles

En la sección anterior hemos visto que la propuesta de integral de Choquet intervalo-valorada se basaba en la elección de un orden admisible concreto que permite ordenar completamente el conjunto de intervalos a agregar. La propuesta que hacemos en esta sección es algo diferente ya que, en lugar de fijar un orden concreto, tiene en cuenta todas las posibles formas de ordenar dicho conjunto de intervalos. De hecho, lo que proponemos es una agregación en dos pasos: en el primero, se calcula el conjunto de permutaciones admisibles de los datos; una vez calculados, se aplica la integral de Choquet utilizando dicha permutación de datos y el resultado final se

agrega realizando un promedio que utiliza la media aritmética. De esta manera, la integral de Choquet que proponemos tiene en cuenta, de alguna manera, todas las posibles ordenaciones de los datos.

Definición 7: Sean $\mathbf{x}_1, \dots, \mathbf{x}_n \in L([0, 1])$ y sea $m : 2^n \rightarrow [0, 1]$ una medida difusa. Sean $\sigma_1, \dots, \sigma_p : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ el conjunto de las permutaciones admisibles de $\mathbf{x}_1, \dots, \mathbf{x}_n$. La integral de Choquet promedio C_m^{arith} respecto a m viene dada por

$$C_m^{arith}(\mathbf{x}_1, \dots, \mathbf{x}_n) = M_{arith}(C_m^{\sigma_1}(\mathbf{x}_1, \dots, \mathbf{x}_n), \dots, C_m^{\sigma_p}(\mathbf{x}_1, \dots, \mathbf{x}_n)) \quad (4)$$

donde

$$C_m^{\sigma_j}(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i=1}^n \mathbf{x}_{\sigma_j(i)} (m(\{\sigma_j(i), \dots, \sigma_j(n)\}) - m(\{\sigma_j(i+1), \dots, \sigma_j(n)\}))$$

para todo $j \in \{1, \dots, p\}$.

Ejemplo 6: Consideremos los siguientes intervalos que deber ser agregados mediante la integral de Choquet: $\mathbf{x}_1 = [0.1, 0.9]$, $\mathbf{x}_2 = [0.3, 0.5]$ y $\mathbf{x}_3 = [0.4, 0.8]$. Sea m una medida difusa dada por $m(\{\emptyset\}) = 0$, $m(\{1\}) = 0.3$, $m(\{2\}) = 0.5$, $m(\{3\}) = 0.4$, $m(\{1, 2\}) = m(\{1, 3\}) = 0.6$ y $m(\{2, 3\}) = 0.8$.

Vamos a calcular la integral de Choquet respecto a todas las permutaciones admisibles. Obsérvese que \mathbf{x}_1 no es comparable a $\mathbf{x}_2, \mathbf{x}_3$ y que $\mathbf{x}_2 <_L \mathbf{x}_3$. Por tanto, existen 3 permutaciones admisibles, representadas en la siguiente matriz donde el elemento (i, j) representa $\sigma_i(j)$:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Ahora, tenemos que

$$\begin{aligned} C_m^{\sigma_1}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) &= [0.3, 0.7] \\ C_m^{\sigma_2}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) &= [0.3, 0.7] \\ C_m^{\sigma_3}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) &= [0.27, 0.71]. \end{aligned}$$

Por último, calculamos la media aritmética de todos los valores por cada permutación admisible y tenemos que

$$C_m^{arith}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = [0.29, 0.703].$$

Para completar el estudio, vamos a analizar algunas de las propiedades más importantes de esta nueva integral de Choquet basada en permutaciones admisibles.

Proposición 3: Son ciertas las siguientes afirmaciones :

- (i) $C_m^{arith}(\mathbf{x}, \dots, \mathbf{x}) = \mathbf{x}$ para todo $\mathbf{x} \in L([0, 1])$;
- (ii) $\inf\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \leq_L C_m^{arith}(\mathbf{x}_1, \dots, \mathbf{x}_n) \leq_L \sup\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ para todo $\mathbf{x}_1, \dots, \mathbf{x}_n \in L([0, 1])$;
- (iii) $C_m^{arith}([x_1, x_1], \dots, [x_n, x_n]) = [C_m(x_1, \dots, x_n), C_m(x_1, \dots, x_n)]$ para todo $x_1, \dots, x_n \in L([0, 1])$;
- (iv) si $\mathbf{x}_1, \dots, \mathbf{x}_n$ forman una anticadena, entonces para cualquier medida difusa m , $C_m^{arith}(\mathbf{x}_1, \dots, \mathbf{x}_n) = M_{arith}(\mathbf{x}_1, \dots, \mathbf{x}_n)$;

- (v) si $\mathbf{x}_1, \dots, \mathbf{x}_n$ forman una cadena, entonces $\mathbf{C}_m^{arith}(\mathbf{x}_1, \dots, \mathbf{x}_n) = \mathbf{C}_m(\mathbf{x}_1, \dots, \mathbf{x}_n)$.

En [7], [16], el concepto de monotonía ha sido sustituido por los conceptos más generales de monotonía débil y monotonía direccional. Utilizando estas generalizaciones de monotonía, el concepto de preagregación ha sido dado en [13]. Aunque la nueva definición de integral de Choquet que hemos definido en este trabajo no satisface la propiedad de monotonía tradicional, sí que satisface un tipo de monotonía direccional.

Proposición 4: La integral de Choquet basada en permutaciones admisibles es una preagregación respecto al vector $\vec{r} = \vec{1}$.

V. CONCLUSIONES

En este trabajo hemos propuesto un nuevo mecanismo de fusión de intervalos basado en la integral de Choquet y en el concepto de permutación admisible. Esta nueva definición permite solucionar algunos problemas encontrados en soluciones anteriores, como la integral de Choquet basada en órdenes admisibles.

Como líneas futuras, nuestra intención es continuar estudiando las propiedades que cumple nuestra nueva definición y su relación con otro tipo de agregaciones intervalo-valoradas. Además, pretendemos extender esta definición a otros retículos y a otro tipo de funciones de agregación, como son los operadores OWA o la integral de Sugeno.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por MINECO, AEI/FEDER,UE bajo el proyecto TIN2016-77356-P y por los proyectos APVV-14-0013 y APVV-17-0066.

REFERENCES

- [1] K. Atanassov. Intuitionistic fuzzy sets. *Fuzzy Sets and Systems* 20 (1986) 87-96.
- [2] G. Beliakov, H. Bustince, T. Calvo. *A Practical guide to Averaging Functions*. Studies in Fuzziness and Soft Computing 329, Springer, 2016.
- [3] G. Beliakov, A. Pradera, T. Calvo. *Aggregation Functions: A Guide for Practitioners*. Studies in Fuzziness and Soft Computing 221, Springer, 2007.
- [4] H. Bustince, Interval-valued fuzzy sets in soft computing. *International Journal of Computational Intelligent Systems* 3 (2010) 215-222.
- [5] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar, Generation of linear orders for intervals by means of aggregation functions. *Fuzzy Sets and Systems* 220 (2013) 69-77.
- [6] H. Bustince, M. Galar, B. Bedregal, A. Kolesárová, R. Mesiar, A New Approach to Interval-Valued Choquet Integrals and the Problem of Ordering in Interval-Valued Fuzzy Set Applications. *IEEE Transactions on Fuzzy Systems* 21 (2013), 1150-1162.
- [7] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar. Directional monotonicity of fusion functions. *European Journal of Operational Research* 244 (2015) 300-308.
- [8] H. Bustince, E. Barrenechea, M. Pagola, J. Fernandez, Z. Xu, B. Bedregal, J. Montero, H. hagrás, F. Herrera, B. De Baets. A Historical Account of Types of Fuzzy Sets and Their Relationships. *IEEE Transactions on Fuzzy Systems* 24 (2016) 179-194.
- [9] G. Choquet, Theory of capacities. *Annales de l'Institute Fourier* 5 (1954) 131-292.
- [10] G. Deschrijver, C. Cornelus, E.E. Kerre. On the representation of intuitionistic fuzzy t-norms and t-conorms. *IEEE Transactions on Fuzzy Systems* 12 (2004) 45-61.
- [11] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, *Aggregation Functions*. Cambridge University Press, 2009.
- [12] M. Komorníková, R. Mesiar. Aggregation functions on bounded partially ordered sets and their classification. *Fuzzy Sets and Systems* 175 (2011) 48-56.
- [13] G. Lucca, J. Sanz, G. P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, H. Bustince. Preaggregation Functions: Construction and an Application. *IEEE Transactions on Fuzzy Systems* 24 (2016) 260-272.
- [14] D. Paternain, H. Bustince, M. Pagola, P. Sussner, A. Kolesárová and R. Mesiar. Capacities and overlap indexes with an application in fuzzy rule-based classification systems. *Fuzzy Sets and Systems* 305 (2016) 70-94.
- [15] V. Torra, Y. Narukawa and M. Sugeno, Eds., *Non-Additive Measures, Theory and Applications (Studies in Fuzziness and Soft Computing Series 310)*, Springer, 2014.
- [16] T. Wilkin, G. Beliakov. Weakly Monotonic Averaging Functions. *International Journal of Intelligent Systems* 30 (2015) 144-169.
- [17] Z. S. Xu, R. R. Yager. Some geometric aggregation operators based on intuitionistic fuzzy sets. *International Journal of General Systems* 35 (2006) 417-433.



Aggregation on relaxed indistinguishability operators based on different triangular norms

Tomasa Calvo Sánchez

University of Alcalá

Email: tomasa.calvo@uah.es

Pilar Fuster-Parra

University of Balearic Islands

Email: pilar.fuster@uib.es

Oscar Valero

University of Balearic Islands

Email: o.valero@uib.es

Abstract—We consider the aggregation problem of relaxed T_i -indistinguishability operators where each member of the collection to be merged is a relaxed indistinguishability operator with respect to a different t-norm T_i . Thus we provide sufficient conditions to merge a collection such an indistinguishability operators. Moreover, we give a characterization of those functions that allow to merge a collection of relaxed indistinguishability operators into a new one whenever all involved t-norms are continuous, Archimedean and strict. Namely, we characterize these functions in terms of the additive generator of the involved t-norms and those functions that transforms n -dimensional triangular triplets of the non-negative real line into a 1-dimensional triplet of the non-negative real line.

I. INTRODUCTION

Aggregation functions constitute an important tool in the field of information fusion (see [4], [5], [7]). The information can be given by means of fuzzy relations what depends on different applications (see, for instance, [1], [13], [19], [26], [27], [34]). Sometimes the aggregation procedure used to take a working decision requires that the nature of the aggregated data be preserved. In the last decades, the problem of aggregating fuzzy relations has grown considerably among researchers in fuzzy mathematics (see [14], [21], [22], [28], [29], [31]). Motivated by the preceding fact, in [9], [10], we have carried out the study of the class preservation of relaxed T -indistinguishability operators by means of aggregation.

Let us recall that, on account of [32], a T -indistinguishability operator on a (non-empty) set X is a fuzzy relation $E : X \times X \rightarrow [0, 1]$ satisfying for all $x, y, z \in X$ the following: (i) $E(x, x) = 1$, (ii) $E(x, y) = E(y, x)$, and (iii) $T(E(x, y), E(y, z)) \leq E(x, z)$. In general, a T -indistinguishability operator allows to classify objects when a measure presents some kind of uncertainty. They are also known as measures of similarity, in fact, the greater $E(x, y)$ the most similar are x and y . We assume that the reader is familiar with the basics of t-norms (see, for instance, [22]).

In the literature, we can find several authors that have analyzed the functions that allows to merge a collection of indistinguishability operators into a single one (see [28], [29], [31]). In particular, in [28] a characterization of those functions that aggregate indistinguishability operators have been given in terms of triangle triplets (compare [30]). Recently, the aforementioned description of the functions that aggregates indistinguishability operators have been extended by Mayor and Recasens in [24], now, in terms of T -triangular triplets.

Inspired by [24], recent results on the aggregation of relaxed T -indistinguishability operators have been developed in [9]. Now, let us recall that a relaxed T -indistinguishability operator on a (non-empty) set X is a fuzzy relation $E : X \times X \rightarrow [0, 1]$ verifying for all $x, y, z \in X$ the following properties: (i) $E(x, y) = E(y, x)$ and (ii) $T(E(x, y), E(y, z)) \leq E(x, z)$. A relaxed T -indistinguishability operator E on a nonempty set X will be called a relaxed T -equality provided that it satisfies the following additional property for any $x, y, z \in X$: (i') $E(x, y) = E(x, x) = E(y, y) \Leftrightarrow x = y$.

There exist another measures very useful to classify objects that are related to (relaxed) T -indistinguishability operators, the so-called (extended) pseudo-metrics (see [12]). However, the extended pseudo-metrics are measures of dissimilarity. In the recent literature, different techniques to generate extended pseudo-metrics from indistinguishability operators, and vice-versa, have been developed (see, for example, [2], [3], [17], [20], [22], [25], [31], [33]).

It is clear that T -indistinguishability operators are relaxed T -indistinguishability operators. Nevertheless, the following is an instance of relaxed and partial indistinguishability operator that is not an indistinguishability one.

Example 1. Let Σ be a non-empty alphabet. Denote by Σ^∞ the set of all finite and infinite sequences over Σ . Given $v \in \Sigma^\infty$ denote by $l(v)$ the length of v . Thus $l(v) \in \mathbb{N} \cup \{\infty\}$ for all $v \in \Sigma^\infty$. Moreover, if $\Sigma_F = \{v \in \Sigma^\infty : l(v) \in \mathbb{N}\}$ and $\Sigma_\infty = \{v \in \Sigma^\infty : l(v) = \infty\}$, then $\Sigma^\infty = \Sigma_F \cup \Sigma_\infty$ and we will write $v = v_1 v_2 \dots v_{l(v)}$ and $w = w_1 w_2 \dots$ whenever $v \in \Sigma_F$ and $w \in \Sigma_\infty$, respectively.

Next, given $x, y \in \Sigma_\infty$, denote by $l(x, y)$ the longest common prefix of x and y (of course if x and y have not a common prefix then $l(x, y) = 0$).

Define the fuzzy relation E_Σ on $\Sigma^\infty \times \Sigma^\infty$ by $E_\Sigma(u, v) = 1 - 2^{-l(u, v)}$ for all $u, v \in \Sigma^\infty$. Then it is a simple matter to check that E_Σ is a relaxed T_{\min} -indistinguishability operator on Σ^∞ . In fact, it is a relaxed T_{\min} -equality on Σ^∞ . Note that T_{\min} denotes the minimum t-norm.

Note that E_Σ is not a indistinguishability operator, since $E_\Sigma(u, u) = 1 \Leftrightarrow u \in \Sigma_\infty$.

As we have mentioned before, the study of the class preservation of relaxed T -indistinguishability operators by means of aggregation was carried out in [9], [10]. Now we focus our attention on the similar preservation problem when

a collection of T_i -indistinguishability operators belonging to the aforesaid class under consideration and in such a way that each member of the collection to be merged is a relaxed indistinguishability operator with respect to different t-norm T_i and all of them are defined on the same set.

Thus, in this work, we provide sufficient conditions to merge a collection of relaxed indistinguishability operators in Section II. Concretely, we see that a dominance condition is sufficient to guarantee that a function aggregates relaxed T_i -indistinguishability operators. Furthermore, we give a characterization of those functions that allow to merge a collection of relaxed indistinguishability operators into a new one whenever all involved t-norms are continuous, Archimedean and strict. Namely, we characterize these functions in terms of the additive generator of the involved t-norms and those functions that transforms n -dimensional triangular triplets of the positive real line into a 1-dimensional triplet of the positive real line. Finally, in Section III, some conclusions are given and future work is proposed.

II. THE AGGREGATION OF RELAXED T_i -INDISTINGUISHABILITY OPERATORS

First of all we introduce the notion of relaxed T_i -indistinguishability operator aggregation function as follows.

Definition 1. A function $F : [0, 1]^n \rightarrow [0, 1]$ aggregates relaxed T_i -indistinguishability operators into a relaxed T -indistinguishability operator if $F(E_1, \dots, E_n)$ is a relaxed T -indistinguishability operators on X for any set X and any collection (E_1, \dots, E_n) of relaxed T_i -indistinguishability operators on X , where $F(E_1, \dots, E_n)$ is the fuzzy binary relation given by $F(E_1, \dots, E_n)(x, y) = F(E_1(x, y), \dots, E_n(x, y))$.

Notice that Definition 1 extends the notion given in [9]. In the light of the preceding concept we have the next result.

Proposition 1. Let T be a t-norm and let $E_i, i = 1, \dots, n$, be a collection of relaxed T_i -indistinguishability operators. If a function $F : [0, 1]^n \rightarrow [0, 1]$ is increasing and satisfies $T(F(a), F(b)) \leq F(T_1(a_1, b_1), \dots, T_n(a_n, b_n))$ for all $\mathbf{a}, \mathbf{b} \in [0, 1]^n$, then $F(E_1, \dots, E_n)$ is a relaxed T -indistinguishability operator.

Proof. To prove that $F(E_1, \dots, E_n)$ is a relaxed T -indistinguishability operator, we only need to show the T -transitivity condition of $F(E_1, \dots, E_n)$ because the symmetry follows directly from the symmetry of each E_i relaxed T_i -indistinguishability operator.

Now, as each E_i is a relaxed T_i -indistinguishability operator we have that $T_i(E_i(x, y), E_i(y, z)) \leq E_i(x, z)$ for all $i = 1, \dots, n$. Since F is an increasing function we get $F(T_1(E_1(x, y), E_1(y, z)), \dots, T_n(E_n(x, y), E_n(y, z))) \leq F(E_1(x, z), \dots, E_n(x, z))$.

Calling $\mathbf{a} = (E_1(x, y), \dots, E_n(x, y))$, $\mathbf{b} = (E_1(y, z), \dots, E_n(y, z))$ and $\mathbf{c} = (E_1(x, z), \dots, E_n(x, z))$ we obtain that $T(F(E_1, \dots, E_n)(x, y), F(E_1, \dots, E_n)(y, z)) = T(F(E_1(x, y), \dots, E_n(x, y)), F(E_1(y, z), \dots, E_n(y, z))) =$

$$\begin{aligned} T(F(a), F(b)) &\leq F(T_1(a_1, b_1), \dots, T_n(a_n, b_n)) = \\ &F(T_1(E_1(x, y), E_1(y, z)), \dots, T_n(E_n(x, y), E_n(y, z))) \leq \\ &F(E_1(x, z), \dots, E_n(x, z)). \end{aligned}$$

Therefore, $F(E_1, \dots, E_n)$ is a relaxed T -indistinguishability operator. \square

According to the previous result, it seems natural to wonder whether the converse of Proposition 1 is true in general. However, the next example, which was given in [9], shows that there are relaxed T_i -indistinguishability aggregation functions that are not increasing.

Example 2. Denote by T_D the Drastic t-norm. Consider the function $F : [0, 1]^2 \rightarrow [0, 1]$ defined by

$$F(\mathbf{a}) = \begin{cases} 1, & \mathbf{a} = (1, 1) \\ 0, & \mathbf{a} = (1/2, 1/2) \\ 1/2, & \text{otherwise} \end{cases}$$

for all $\mathbf{a} \in [0, 1]^2$. It is not hard to check that F aggregates T_D -indistinguishability operators. However F is not an increasing function due to $F(1/2, 0) \leq F(0, 0)$.

When the collection of $n + 1$ -t-norms T, T_i , are such that $T \leq T_i$, for all $i = 1, \dots, n$, the following result can be stated.

Proposition 2. Let T be a t-norm and let $E_i, i = 1, \dots, n$, be a collection of relaxed T_i -indistinguishability operators. If $T \leq T_i$ for all $i = 1, \dots, n$, then $T(E_1, \dots, E_n)$ is a relaxed T -indistinguishability operator.

Proof. First of all we show the transitivity condition. On the one hand, by the commutativity and associativity of T , we have that

$$T(T(E_1(x, y), \dots, E_n(x, y)), T(E_1(y, z), \dots, E_n(y, z))) = T(T(E_1(x, y), E_1(y, z)), \dots, T(E_n(x, y), E_n(y, z))).$$

On the other hand, the monotony of T and the fact that $T \leq T_i$ provide that

$$T(T(E_1(x, y), E_1(y, z)), \dots, T(E_n(x, y), E_n(y, z))) \leq T(T_1(E_1(x, y), E_1(y, z)), \dots, T_n(E_n(x, y), E_n(y, z))).$$

Since $T_i(E_i(x, y), E_i(y, z)) \leq E_i(x, z)$ for all $i = 1, \dots, n$ we have that

$$T(T_1(E_1(x, y), E_1(y, z)), \dots, T_n(E_n(x, y), E_n(y, z))) \leq T(E_1(x, z), \dots, E_n(x, z)).$$

It follows that

$$T(T(E_1, \dots, E_n)(x, y), T(E_1, \dots, E_n)(y, z)) \leq T(E_1, \dots, E_n)(x, z).$$

This means that $T(E_1, \dots, E_n)$ is a T -transitive relation. As the symmetry of this relation is always guaranteed by the symmetry of T , we conclude that $T(E_1, \dots, E_n)$ is a relaxed T -indistinguishability operator. \square

As an example, in the spirit of the above proposition, we can choose $T = T_D$ and $T_i = T_P$ for all $i = 1, \dots, n$. Then, by Proposition 2, $T_D(E_1, \dots, E_n)$ is a relaxed T_D -indistinguishability operator when E_i is a collection of T_P -indistinguishability operators, since $T_D \leq T_P$.



In order to introduce the announced characterization of those functions that merge relaxed T_i -indistinguishability operators, we need to recall the following notion which was provided in [6].

Definition 2. A triplet $(a, b, c) \in [0, \infty]^3$ is said to be (1-dimensional) triangular if and only if $a \leq b + c$, $b \leq a + c$ and $c \leq a + b$. Being $\mathbf{a}, \mathbf{b}, \mathbf{c} \in [0, \infty]^n$, $n \geq 1$, we say that $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is a (n -dimensional) triangular triplet if (a_i, b_i, c_i) is a triangular triplet for all $i = 1, \dots, n$, where $\mathbf{a} = (a_1, \dots, a_n)$, $\mathbf{b} = (b_1, \dots, b_n)$ and $\mathbf{c} = (c_1, \dots, c_n)$.

Taking into account the preceding notion we provide the following characterization.

Theorem 1. Let T, T_1, \dots, T_n be strict and continuous Archimedean t -norms. The following assertions are equivalent:

- 1) A function $F : [0, 1]^n \rightarrow [0, 1]$ aggregates a collection of n relaxed T_i -indistinguishability operators E_i into a relaxed T -indistinguishability operator.
- 2) There exists a function $G : [0, +\infty]^n \rightarrow [0, +\infty]$ which transforms n -dimensional triangular triplets into 1-dimensional triangular triplets and satisfies

$$G = t \circ F \circ (t_1^{-1} \times \dots \times t_n^{-1}),$$

where t and t_i are additive generators of t -norms T and T_i , respectively.

Proof. (1) \rightarrow (2) Let us suppose that F aggregates a collection of relaxed T_i -indistinguishability operators into a relaxed T -indistinguishability operator. We have to prove that G transforms n -dimensional triangular triplets into 1-dimensional triangular triplets. To this end, assume that $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \in [0, \infty]^n$ is a n -dimensional triplet. Define the fuzzy relations E_i , $i = 1, \dots, n$, on a non-empty set $X = \{x, y, z\}$ (of different elements) by $E_i(x, y) = E_i(y, x) = t_i^{-1}(b_i)$, $E_i(x, z) = E_i(z, x) = t_i^{-1}(a_i)$, $E_i(y, z) = E_i(z, y) = t_i^{-1}(c_i)$, $E_i(x, x) = E_i(y, y) = E_i(z, z) = 1$. It is not hard to check that each E_i is a relaxed T_i -indistinguishability operator, since $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is a n -dimensional triplet. Next, let us prove that $(G(\mathbf{a}), G(\mathbf{b}), G(\mathbf{c}))$ is a 1-dimensional triangular triplet.

Since F aggregates the collection of relaxed T_i -indistinguishability operators E_i into a relaxed T -indistinguishability operator we have that

$$\frac{T(F(E_1, \dots, E_n)(x, y), F(E_1, \dots, E_n)(y, z))}{F(E_1, \dots, E_n)(x, z)} \leq$$

It follows that

$$t^{-1}(t \circ F(E_1, \dots, E_n)(x, y)) + t \circ (F(E_1, \dots, E_n)(y, z)) \leq F(E_1, \dots, E_n)(x, z).$$

Whence we deduce that

$$t \circ (F(E_1, \dots, E_n)(x, y)) + t \circ (F(E_1, \dots, E_n)(y, z)) \geq t(F(E_1, \dots, E_n)(x, z)).$$

The fact that $E_i(x, y) = t^{-1}(b_i)$, $E_i(x, z) = t^{-1}(a_i)$ and $E_i(y, z) = t^{-1}(c_i)$ provides that

$$t \circ (F(t_1^{-1}(b_1), \dots, t_n^{-1}(b_n))) + t \circ (F(t_1^{-1}(c_1), \dots, t_n^{-1}(c_n))) \geq t \circ (F(t_1^{-1}(a_1), \dots, t_n^{-1}(a_n))).$$

Hence we conclude that $G(\mathbf{a}) \leq G(\mathbf{b}) + G(\mathbf{c})$ and, thus, that G transforms a n -dimensional triangular triplet into a 1-dimensional triangular triplet.

(2) \rightarrow (1) Assuming that G transforms n -dimensional triangular triplets into 1-dimensional triangle triplets, we must prove that $F(E_1, \dots, E_n)$ is a relaxed T -indistinguishability operator for all collection E_i , $i = 1, \dots, n$, of relaxed T_i -indistinguishability operators. To this end, consider a collection E_i , $i = 1, \dots, n$, of relaxed T -indistinguishability operators on a non-empty set X . Then, for each $x, y, z \in X$, we set $E_i(x, y) = E_i(y, x) = a_i$, $E_i(y, z) = E_i(z, y) = b_i$, and $E_i(x, z) = E_i(z, x) = c_i$. Since $T_i(E_i(x, y), E_i(y, z)) \leq E_i(x, z)$, $T_i(E_i(x, z), E_i(z, y)) \leq E_i(x, y)$ and $T_i(E_i(y, x), E_i(x, z)) \leq E_i(y, z)$ we have that $t_i(a_i) + t_i(b_i) \geq t_i(c_i)$, $t_i(b_i) + t_i(c_i) \geq t_i(a_i)$ and $t_i(a_i) + t_i(c_i) \geq t_i(b_i)$ and, thus, that $(t_i(a_i), t_i(b_i), t_i(c_i))$ is a 1-dimensional triangle triplet. So $(\mathbf{t}(\mathbf{a}), \mathbf{t}(\mathbf{b}), \mathbf{t}(\mathbf{c}))$ is a n -dimensional triangle triplet. The fact that G transforms n -dimensional triangular triplets into 1-dimensional triangle triplets provides that $(G(\mathbf{t}(\mathbf{a})), G(\mathbf{t}(\mathbf{b})), G(\mathbf{t}(\mathbf{c})))$ is a 1-dimensional triangle triplet, where $\mathbf{t}(\mathbf{a}) = (t_1(a_1), \dots, t_n(a_n))$, $\mathbf{t}(\mathbf{b}) = (t_1(b_1), \dots, t_n(b_n))$ and $\mathbf{t}(\mathbf{c}) = (t_1(c_1), \dots, t_n(c_n))$. Thus $G(\mathbf{t}(\mathbf{a})) \leq G(\mathbf{t}(\mathbf{b})) + G(\mathbf{t}(\mathbf{c}))$, $G(\mathbf{t}(\mathbf{b})) \leq G(\mathbf{t}(\mathbf{a})) + G(\mathbf{t}(\mathbf{c}))$ and $G(\mathbf{t}(\mathbf{c})) \leq G(\mathbf{t}(\mathbf{a})) + G(\mathbf{t}(\mathbf{b}))$.

Since $G = t \circ F \circ (t_1^{-1} \times \dots \times t_n^{-1})$ we have

$$\begin{aligned} t(F(\mathbf{a})) &= G(\mathbf{t}(\mathbf{a})) \leq G(\mathbf{t}(\mathbf{b})) + G(\mathbf{t}(\mathbf{c})) \\ &= t(F(\mathbf{b})) + t(F(\mathbf{c})). \end{aligned}$$

From the preceding inequality we deduce that

$$F(\mathbf{a}) \geq t^{-1}(t(F(\mathbf{b})) + t(F(\mathbf{c}))) = T(F(\mathbf{b}), F(\mathbf{c})).$$

Therefore we obtain that

$$\frac{T(F(E_1, \dots, E_n)(z, y), F(E_1, \dots, E_n)(x, z))}{F(E_1, \dots, E_n)(x, y)} \leq$$

Similarly we can show that

$$\frac{T(F(E_1, \dots, E_n)(x, y), F(E_1, \dots, E_n)(y, z))}{F(E_1, \dots, E_n)(x, z)} \leq$$

and

$$\frac{T(F(E_1, \dots, E_n)(y, x), F(E_1, \dots, E_n)(x, z))}{F(E_1, \dots, E_n)(y, z)} \leq$$

Moreover, $F(E_1, \dots, E_n)(x, y) = F(E_1, \dots, E_n)(y, x)$ due to every E_i is a relaxed T_i -indistinguishability operator. Consequently, we conclude that F aggregates relaxed T_i -indistinguishability operators into a relaxed T -indistinguishability operator. \square

In order to finish the paper we recall two pertinent properties of those functions that transform n -dimensional triangular triplets into 1-dimensional triangular triplets (see [6]).

Proposition 3. Consider a function $G : [0, \infty]^n \rightarrow [0, \infty]$, $n \geq 1$. Then:

- i) If G transforms n -dimensional triangular triplets into 1-dimensional triangular triplets, then it is subadditive.
- ii) If G is increasing and subadditive, then it transforms n -dimensional triangular triplets into 1-dimensional triangular triplets.

Taking into account Propositions 1 and 3 we can state the next result.

Corollary 1. Let T, T_1, \dots, T_n be strict and continuous Archimedean t -norms ($n \geq 1$) and let $F : [0, 1]^n \rightarrow [0, 1]$ be an increasing function. Then the following assertions are equivalent:

- 1) $F : [0, 1]^n \rightarrow [0, 1]$ aggregates relaxed T_i -indistinguishability operators into a relaxed T -indistinguishability operator.
- 2) The function $G : [0, +\infty]^n \rightarrow [0, +\infty]$ given by $G = t \circ F \circ (t_1^{-1} \times \dots \times t_n^{-1})$ is subadditive.

Proof. (1) \rightarrow (2) Taking into account Theorem 1 we know that G transforms n -dimensional triangular triplets into a 1-dimensional triangular triplet and, in addition, from Proposition 3 we can ensure that G is a subadditive function.

(2) \rightarrow (1). The monotony of F yields that G is increasing. Proposition 3 guarantees that the function G transforms n -dimensional triangular triplets into 1-dimensional triangular triplets. Now, from Theorem 1, it follows that F aggregates relaxed T_i -indistinguishability operators into a relaxed T -indistinguishability operator. \square

III. CONCLUSIONS

We have addressed the aggregation problem of a collection of relaxed indistinguishability operators where each member of the collection to be merged is a relaxed indistinguishability operator with respect to a different t -norm. A sufficient condition to merge a collection of the indistinguishability operators under consideration has been given. Moreover, we have characterized those functions that allow to merge a collection of relaxed T_i -indistinguishability operators into a new with respect to a new t -norm T . Namely, the aforesaid characterization is provided in terms of the additive generator of the all involved t -norms and another functions that transform a n -dimensional triangular triplet of the positive real line into a 1-dimensional triplet of the positive real line. As future work we will discuss whether some concrete families of aggregation functions (OWA operators, overlap functions, etc) could be used as relaxed T -indistinguishability operators.

ACKNOWLEDGEMENTS

This research was funded by the Spanish Ministry of Economy and Competitiveness under Grants DPI2017-86372-C3-3-R, TIN2016-81731-REDT (LODISCO II) and AEI/FEDER,

UE funds, by the Programa Operatiu FEDER 2014-2020 de les Illes Balears, by project ref. PROCOE/4/2017 (Direcció General d' Innovació i Recerca, Govern de les Illes Balears), and by project ROBINS. The latter has received research funding from the EU H2020 framework under GA 779776. This publication reflects only the authors views and the European Union is not liable for any use that may be made of the information contained therein.

REFERENCES

- [1] B. De Baets and R. Mesiar, T -partitions, Fuzzy Sets and Syst., 97, (1998), 211–223.
- [2] B. De Baets and R. Mesiar, Pseudo-metrics and T -equivalences, J. Fuzzy Math., 5 (1997), 471–481.
- [3] B. De Baets and R. Mesiar, Metrics and T -equalities, J. Math. Anal. Appl., 267 (2002), 531–547.
- [4] G. Beliakov, H. Bustince and T. Calvo-Sánchez, A Practical Guide to Averaging Functions, Springer, Heidelberg, Berlin, New York, 2016.
- [5] G. Beliakov, A. Pradera and T. Calvo, Aggregation Functions: a guide for practitioners, Springer, Heidelberg, Berlin, New York, 2007.
- [6] J. Borsík and J. Doboš, On a product of metric spaces. Math. Slovaca 31 (1981), 193–205.
- [7] B. Bouchon-Meunier, (Ed.), Aggregation and Fusion of Imperfect Information, Studies in Fuzziness and Soft Computing, Vol. 12, Physica-Verlag, Heidelberg, 1998.
- [8] M. Bukatin, R. Kopperman and S.G. Matthews, Some corollaries of the correspondence between partial metric and multivalued equalities, Fuzzy Set. Syst. 256 (2014), 57–72.
- [9] T. Calvo, P. Fuster-Parra and O. Valero, On the problem of relaxed indistinguishability aggregation operators, Proc. of the Workshop on Applied Topological Structures, WATS'17, (2017), 19–26.
- [10] T. Calvo, P. Fuster-Parra and O. Valero, On the Problem of Aggregation of Partial T -Indistinguishability Operators, Jesús Medina, Manuel Ojeda-Aciego, José Luis Verdegay, David A. Pelta, Inma P. Cabrera, Bernadette Bouchon-Meunier, Ronald R. Yager (Eds.), Proc. Part I. Information Processing and Management of Uncertainty in Knowledge-Based Systems, Theory and Foundations, IPMU'18, (2018), 207–218.
- [11] M. Demicri, The order-theoretic duality and relations between partial metrics and local equalities, Fuzzy Set. Syst. 192 (2011) 45–57.
- [12] M.M. Deza and E. Deza, Encyclopedia of Distances, Springer-Verlag, 2016.
- [13] J. Drewniak and U. Dudziak, Aggregation preserving classes of fuzzy relations, Kybernetika, 41 (3), (2005), 265–284.
- [14] U. Dudziak, Preservation of t -norm and t -conorm based properties of fuzzy relations during aggregation process, Proc. 8th Conference of European Society for Fuzzy Logic and Technology, EUSFLAT'2013, (2013), 376–383.
- [15] P. Fuster-Parra, J. Martín, J. J. Miñana and O. Valero, A study on the relationship between relaxed metrics and indistinguishability operators, 2017, (submitted).
- [16] P. Fuster-Parra, J. Martín, J. Recasens and O. Valero, On the metric behavior of partial indistinguishability operators, preprint.
- [17] S. Gottwald, On t -norms which are related to distances of fuzzy sets, BUSEFAL 50 (1992), 25–30.
- [18] M. Grabisch, J.L. Marichal, R. Mesiar and E. Pap, Aggregation Functions, Cambridge University Press, New York, 2009.
- [19] O. Grigorenko and J. Lebedinska, On another view of aggregation of fuzzy relations, In: S. Galichet, J. Montero and G. Mauris, (eds), Proc. 7th. Conf. EUSFLAT'2011 and LFA'2011, Atlantis Press, (2011), 21–27.
- [20] U. Höhle, Fuzzy equalities and indistinguishability, Proc. of EUFIT'93, Vol. 1, Aachen, 1993 pp. 358–363.
- [21] J. Jacas and J. Recasens, Aggregation of T -Transitive Relations, Int. J. Intell. Syst. 18 (2003), 1193–1214.
- [22] P. Klement, R. Mesiar and E. Pap, Triangular Norms, Kluwer, Dordrecht, 2000.
- [23] C. M. Ling, Representation of associative function, Publ. Math. Debrecen 12, (1965), 189–212.
- [24] G. Mayor and J. Recasens, Preserving T -transitivity in: Artificial Intelligence Research and Development, A. Nebot et al. (eds), IOS Press, Vol. 288, 79–87, 2016.



- [25] S. Ovchinnikov, Representation of transitive fuzzy relations, Aspects of Vageness (H. Skala, S. Termini and E. Trillas, eds.), Reidel, Dordrecht, 1984, pp. 105-118.
- [26] S. Ovchinnikov, Similarity relations, fuzzy partitions, and fuzzy orderings, *Fuzzy Sets and Syst.*, 40 (1991), 107-126.
- [27] V. Peneva and I. Popchev, Properties of the aggregation operators related with fuzzy relations, *Fuzzy Sets and Syst.*, 139 (3), (2003), 615-633.
- [28] A. Pradera, E. Trillas, and E. Castiñeira. On distances aggregation, in: *Proc. of Information Processing and Management of Uncertainty in Knowledge-Based Systems International Conference*, B. Bouchon-Meunier et al. (eds), Universidad Politécnica de Madrid Press, Vol. II, 693-700, 2000.
- [29] A. Pradera, E. Trillas and E. Castiñeira, On the aggregation of some classes of fuzzy relations, in: *Technologies for constructing intelligent systems*, B. Bouchon-Meunier, J. Gutierrez, L. Magdalena and R. Yager (eds), Springer-Verlag, 125-147, 2002.
- [30] A. Pradera, E. Trillas, A note on pseudo-metrics aggregation, *Int. J. Gen. Syst.* 31 (2002), 41-51.
- [31] J. Recasens, *Indistinguishability Operators: Modelling Fuzzy Equalities and Fuzzy Equivalence Relations*, Springer, Berlin, 2010.
- [32] E. Trillas, Assaig sobre les relacions d'indistingibilitat, *Proc. Primer Congrés Català de Lògica Matemàtica*, Barcelona, 1982, pp. 51-59.
- [33] L. Valverde, On the structure of F-indistinguishability operators, *Fuzzy Set. Syst.* 17 (1985) 313-328.
- [34] L. A. Zadeh, Similarity relations and fuzzy orderings, *Inform. Sci.* 3(2), (1971), 177-200.