

La Noción de Función de Pre-agregación

Humberto Bustince^{1,2}, Javier Fernandez^{1,2}, Laura De Miguel¹, and Carlos López-Molina¹

¹ Departamento of Automática y Computación, Universidad Publica de Navarra, Navarra, 31006 Spain

² Institute of Smart Cities, Universidad Publica de Navarra, Navarra, 31006 Spain

Resumen En este trabajo utilizamos la idea de monotonía direccional para introducir el concepto de pre-agregación. Presentamos además un método de construcción de tales funciones inspirado en las integrales Choquet.

Keywords: Función de pre-agregación, función de agregación, monotonía direccional, integral Choquet.

1. Introducción

Las funciones de agregación [1,2] son herramientas fundamentales hoy en día en muchos problemas de computación [3,4,5,6,7,8,9]. Sin embargo, dado que muchos operadores importantes como la moda no son monótonos, en los últimos años existe un interés creciente en relajar la condición de monotonía para recuperar operadores útiles desde el punto de vista de las aplicaciones. Por ejemplo, en [10], Wilkin y Beliakov proponen la idea de monotonía débil, que equivale a considerar monotonía a lo largo de la bisectriz del primer cuadrante. La consideración de direcciones arbitrarias llevó a Bustince et al. [11] a definir la monotonía direccional. Las funciones monótonas en sentido clásico son, en particular, débilmente monótonas y direccionalmente monótonas.

En este trabajo proponemos la definición de función de pre-agregación, que es una función que satisface las mismas condiciones de contorno de una función de agregación pero a la que sólo se exige monotonía direccional. En particular, discutimos un ejemplo particular de tales funciones, que se obtiene al reemplazar el producto por una agregación adecuada en la definición de la integral de Choquet [12,13,14,15]. La utilidad de este planteamiento queda clara en algunos trabajos recientes sobre clasificación [16,17].

El trabajo se organiza como sigue. En la Sección 2 recordamos algunos conceptos preliminares. En la Sección 3 presentamos la noción de función de pre-agregación y discutimos algunas de sus propiedades. En la Sección 4 introducimos un método de construcción. Terminamos con algunas conclusiones y referencias.

2. Preliminares

Empezamos recordando algunas nociones y resultados que son relevantes para este trabajo.

Definición 1 Una función $A : [0, 1]^n \rightarrow [0, 1]$ es una función de agregación (AGOF) n -dimensional si:

- (A1) A es creciente: para cada $i \in \{1, \dots, n\}$, si $x_i \leq y$, entonces $A(x_1, \dots, x_n) \leq A(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)$;
- (A2) $A(0, \dots, 0) = 0$ y $A(1, \dots, 1) = 1$.

Definición 2 Una AGOF $T : [0, 1]^2 \rightarrow [0, 1]$ es una t -norma si, para todo $x, y, z \in [0, 1]$, se verifican las siguientes propiedades:

- (T1) *Commutatividad:* $T(x, y) = T(y, x)$;
- (T2) *Asociatividad:* $T(x, T(y, z)) = T(T(x, y), z)$;
- (T3) *Condiciones de contorno:* $T(x, 1) = x$.

Si T sólo satisface (T3) y $T(1, x) = x$, se denomina *semi-cópula*.

Algunos ejemplos de t -normas son:

1. Mínimo $T_M(x, y) = \min\{x, y\}$
2. Producto algebraico $T_P(x, y) = xy$
3. t -norma de Łukasiewicz $T_L(x, y) = \max\{0, x + y - 1\}$
4. Producto Drástico $T_{DP}(x, y) = \begin{cases} x & \text{si } y = 1 \\ y & \text{si } x = 1 \\ 0 & \text{en otro caso} \end{cases}$
5. Mínimo nilpotente $T_{NM}(x, y) = \begin{cases} \min\{x, y\} & \text{si } x + y > 1 \\ 0 & \text{en otro caso} \end{cases}$
6. Producto de Hamacher $T_{HP}(x, y) = \begin{cases} 0 & \text{si } x = y = 0 \\ \frac{xy}{x+y-xy} & \text{en otro caso} \end{cases}$

Sea $N = \{1, \dots, n\}$ para un entero positivo arbitrario n .

Definición 3 Una función $\mathbf{m} : 2^N \rightarrow [0, 1]$ es una medida difusa si, para todo $X, Y \subseteq N$, se tiene que:

- (m1) *Monotonía:* si $X \subseteq Y$, entonces $\mathbf{m}(X) \leq \mathbf{m}(Y)$;
- (m2) *Condiciones de contorno:* $\mathbf{m}(\emptyset) = 0$ y $\mathbf{m}(N) = 1$.

Algunos ejemplos de medidas son los siguientes.

Medida uniforme:

$$\mathbf{m}_U(A) = \frac{|A|}{n}. \quad (1)$$

Medida de Dirac: Para un $i \in N$ fijado,

$$m_D^i(A) = \begin{cases} 1 & \text{si } i \in A \\ 0 & \text{si } i \notin A. \end{cases} \quad (2)$$

Medida aditiva (Wmean): Sea $W = (w_1, \dots, w_n) \in [0, 1]^n$ tal que $\sum_{i=1}^n w_i = 1$. Sea

$$m_W(\{i\}) = w_i$$

Entonces, para $|A| \geq 1$, definimos:

$$m_W(A) = \sum_{i \in A} w_i. \quad (3)$$

Medida simétrica (OWA): Sea $W = (w_1, \dots, w_n) \in [0, 1]^n$ tal que $\sum_{i=1}^n w_i = 1$. Entonces, para cada subconjunto no vacío A , definimos:

$$m_{sW}(A) = \sum_{i=1}^{|A|} w_i. \quad (4)$$

Medida potencia:

$$m_{PM}(A) = \left(\frac{|A|}{n}\right)^q, \text{ con } q > 0. \quad (5)$$

Recordamos también la definición de integral Choquet.

Definición 4 [1, Definition 1.74] Sea $m : 2^N \rightarrow [0, 1]$ una medida difusa. La integral discreta de Choquet de $x = (x_1, \dots, x_n) \in [0, 1]^n$ respecto a m es la función $C_m : [0, 1]^n \rightarrow [0, 1]$, dada por

$$C_m(x) = \sum_{i=1}^n (x_{(i)} - x_{(i-1)}) \cdot m(A_{(i)}), \quad (6)$$

donde $(x_{(1)}, \dots, x_{(n)})$ es una permutación creciente de las componentes de x , es decir, $0 \leq x_{(1)} \leq \dots \leq x_{(n)}$, donde $x_{(0)} = 0$, y $A_{(i)} = \{(i), \dots, (n)\}$ es el subconjunto de índices que corresponde a las $n - i + 1$ mayores componentes de x .

Recordamos ahora la definición de monotonía direccional [11].

Definición 5 Sea $r = (r_1, \dots, r_n)$ un vector real n -dimensional, $r \neq \mathbf{0}$. Una función $F : [0, 1]^n \rightarrow [0, 1]$ es r -creciente si para todo $(x_1, \dots, x_n) \in [0, 1]^n$ y todo $c > 0$ tal que $(x_1 + cr_1, \dots, x_n + cr_n) \in [0, 1]^n$, se tiene que

$$F(x_1 + cr_1, \dots, x_n + cr_n) \geq F(x_1, \dots, x_n).$$

Example 1. ■ Las funciones de implicación (ver [18]) son $(-1, 1)$ -crecientes.

- Las funciones débilmente crecientes son un caso particular de funciones direccionalmente crecientes, con $r = (1, \dots, 1)$.

3. Funciones de Pre-agregación

Definición 6 Una función $F : [0, 1]^n \rightarrow [0, 1]$ es una función de pre-agregación n -dimensional si se tiene que:

(PA1) Existe un vector real $\mathbf{r} \in [0, 1]^n$ ($\mathbf{r} \neq \mathbf{0}$) tal que F es \mathbf{r} -creciente.

(PA2) $F(0, \dots, 0) = 0$ and $F(1, \dots, 1) = 1$.

Example 2. Algunos ejemplos de funciones de pre-agregación son los siguientes.

- (i) La moda, $Mod(x_1, \dots, x_n)$, definida como la función que devuelve el valor que aparece más veces en la tupla considerada, o el menor de tales valores, si es que hay más de uno, es $(1, \dots, 1)$ es $(1, 1)$ -creciente y, por tanto, un ejemplo de pre-agregación.
- (ii) $F(x, y) = x - (\max\{0, x - y\})^2$ es $(0, 1)$ -creciente, y un ejemplo de función de pre-agregación.

Si F es una función de pre-agregación respecto a \mathbf{r} , diremos simplemente que F es una \mathbf{r} -pre-agregación.

Comentario 1 Si $A : [0, 1]^n \rightarrow [0, 1]$ es una función de agregación, entonces A es también una pre-agregación.

Podemos utilizar funciones de agregación para construir funciones direccionalmente crecientes como sigue.

Proposición 1 Sea $A : [0, 1]^m \rightarrow [0, 1]$ una función de agregación. Sea $F_i : [0, 1]^n \rightarrow [0, 1]$ ($i \in \{1, \dots, m\}$) una familia de m \mathbf{r} -pre-agregaciones para el mismo vector $\mathbf{r} \in [0, 1]^n$. Entonces, la función $A(F_1, \dots, F_m) : [0, 1]^n \rightarrow [0, 1]$, definida como

$$A(F_1, \dots, F_m)(x_1, \dots, x_n) = A(F_1(x_1, \dots, x_n), \dots, F_m(x_1, \dots, x_n))$$

es también una \mathbf{r} -pre-agregación.

Demostración. Sea $\mathbf{r} = (r_1, \dots, r_n)$. Entonces, del \mathbf{r} -crecimiento, se sigue que, para cada $(x_1, \dots, x_n) \in [0, 1]^n$ y $c > 0$ tal que $(x_1 + cr_1, \dots, x_n + cr_n) \in [0, 1]^n$,

$$F_i(x_1 + cr_1, \dots, x_n + cr_n) \geq F_i(x_1, \dots, x_n)$$

para todo $i \in \{1, \dots, m\}$. Como toda función de agregación es creciente, se sigue el resultado.

El siguiente resultado es inmediato.

Proposición 2 Sean $F_1, F_2 : [0, 1]^n \rightarrow [0, 1]$ dos \mathbf{r} -pre-agregaciones para el mismo vector $\mathbf{r} \in [0, 1]^n$. Entonces:

- (i) $\frac{F_1 + F_2}{2}$ es una \mathbf{r} -pre-agregación.
- (ii) $F_1 F_2$ es una \mathbf{r} -pre-agregación.

Respect a la dualidad, podemos decir los siguiente.

Proposición 3 Sea $F : [0, 1]^n \rightarrow [0, 1]$ una r -pre-agregación para $r \in [0, 1]^n$. Entonces, la función

$$F^d(x_1, \dots, x_n) = 1 - F(1 - x_1, \dots, 1 - x_n)$$

es también una r -pre-agregación.

Demostración. Sea $r = (r_1, \dots, r_n)$. Tomemos $(x_1, \dots, x_n) \in [0, 1]^n$ y $c > 0$ tales que $(x_1 + cr_1, \dots, x_n + cr_n) \in [0, 1]^n$. Entonces

$$\begin{aligned} F^d(x_1 + cr_1, \dots, x_n + cr_n) &= 1 - F(1 - x_1 - cr_1, \dots, 1 - x_n - cr_n) \\ &\geq 1 - F(1 - x_1, \dots, 1 - x_n) = F^d(x_1, \dots, x_n) \end{aligned}$$

y tenemos el resultado.

4. Un método de construcción de pre-agregaciones

En esta sección presentamos un método de construcción de pre-agregaciones. El método se inspira en la construcción de integrales Choquet, reemplazando el producto en (6) por otras agregaciones.

Sea $m : 2^N \rightarrow [0, 1]$ una medida difusa y $M : [0, 1]^2 \rightarrow [0, 1]$ una función. Definimos $C_m^M : [0, 1]^n \rightarrow [0, n]$ como

$$C_m^M(x) = \sum_{i=1}^n M(x_{(i)} - x_{(i-1)}, m(A_{(i)})), \tag{7}$$

donde $N = \{1, \dots, n\}$, y $(x_{(1)}, \dots, x_{(n)})$ es una permutación creciente de las componentes de x , i.e., $0 \leq x_{(1)} \leq \dots \leq x_{(n)}$, con la convención de que $x_{(0)} = 0$, y $A_{(i)} = \{(i), \dots, (n)\}$ es el subconjunto de índices que corresponden a las $n - i + 1$ mayores componentes de x .

Tenemos en primer lugar lo siguiente.

Proposición 4 Sea $M : [0, 1]^2 \rightarrow [0, 1]$ una función tal que $M(x, y) \leq x$ para todo $x, y \in [0, 1]$. Entonces

$$C_m^M(x_1, \dots, x_n) \leq \max(x_1, \dots, x_n)$$

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

Demostración. Tenemos que

$$\begin{aligned} C_m^M(x_1, \dots, x_n) &= \sum_{i=1}^n M(x_{(i)} - x_{(i-1)}, m(A_{(i)})) \\ &\leq \sum_{i=1}^n (x_{(i)} - x_{(i-1)}) = x_{(n)} = \max(x_1, \dots, x_n) \end{aligned}$$

de donde se sigue el resultado.

Proposición 5 Sea $M : [0, 1]^2 \rightarrow [0, 1]$ una función tal que $M(x, 1) = x$ para todo $x \in [0, 1]$. Entonces

$$C_{\mathbf{m}}^M(x_1, \dots, x_n) \geq \overline{\text{mín}}(x_1, \dots, x_n)$$

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

Demostración. Tenemos que

$$\overline{\text{mín}}(x_1, \dots, x_n) = x_{(1)} = M(x_{(1)} - x_{(0)}, \mathbf{m}(A_{(1)})) \leq C_{\mathbf{m}}^M(x_1, \dots, x_n)$$

de donde se sigue el resultado.

Proposición 6 Sea $M : [0, 1]^2 \rightarrow [0, 1]$ una función tal que $M(x, 1) = x$ and $M(0, y) = 0$ para todo $x, y \in [0, 1]$. Entonces, la función $C_{\mathbf{m}}^M$ es idempotente.

Demostración. Para cada $(x, \dots, x) \in [0, 1]^n$ tenemos que

$$C_{\mathbf{m}}^M(x, \dots, x) = M(x, 1) + \sum_{i=2}^n M(0, m(A_{(i)})) = x.$$

Proposición 7 Para toda $M : [0, 1]^2 \rightarrow [0, 1]$ tal que $M(0, 1) = 0$, $M(1, 1) = 1$ que es $(1, 0)$ -increasing, $C_{\mathbf{m}}^M$ es $(1, \dots, 1)$ -creciente.

Demostración. Sea $\mathbf{r} = (1, \dots, 1)$. En la ecuación (7), para $i \geq 2$, se tiene que, para todo $c > 0$

$$M(x_{(i)} + c - (x_{(i-1)} + c), \mathbf{m}(A_{(i)})) = M(x_{(i)} - x_{(i-1)}, \mathbf{m}(A_{(i)}))$$

mientras que, para $i = 1$

$$M(x_{(1)} + c - x_{(0)}, \mathbf{m}(A_{(1)})) = M(x_{(1)} + c, \mathbf{m}(A_{(1)})) \geq M(x_{(1)}, \mathbf{m}(A_{(1)}))$$

luego $C_{\mathbf{m}}^M$ es \mathbf{r} -creciente.

El siguiente resultado es ahora directo.

Theorem 1. Sea $M : [0, 1]^2 \rightarrow [0, 1]$ una función tal que, para todo $x, y \in [0, 1]$, se verifica que $M(x, y) \leq x$, $M(x, 1) = x$, $M(0, y) = 0$ y M es $(1, 0)$ -creciente. Entonces, para cualquier medida difusa \mathbf{m} , $C_{\mathbf{m}}^M$ es una pre-agregación idempotente y compensativa, i.e., $\overline{\text{mín}}(x_1, \dots, x_n) \leq C_{\mathbf{m}}^M(x_1, \dots, x_n) \leq \overline{\text{máx}}(x_1, \dots, x_n)$.

Demostración. Se sigue de las Proposiciones 4, 5 y 7.

Dado que una semi-cópula es una función de agregación M tal que $M(1, x) = M(x, 1) = x$ para cada $x \in [0, 1]$, tenemos el siguiente corolario.

Corolario 1 Sea $M : [0, 1]^2 \rightarrow [0, 1]$ una semi-cópula. Entonces, para cualquier medida difusa \mathbf{m} , $C_{\mathbf{m}}^M$ es una pre-agregación idempotente y compensativa.

Comentario 2 Con las hipótesis del Teorema 1, no podemos garantizar la monotonía de C_m^M , i.e., C_m^M es una pre-agregación propia. Para comprobarlo, observemos lo siguiente:

- (i) Sea $M(x, y) = T_M(x, y)$. Tomemos $N = \{1, 2, 3, 4\}$ y la medida uniforme $\mathbf{m} = \mathbf{m}_U$ dada en (1). Entonces

$$C_m^{T_M}(0,05, 0,1, 0,7, 0,9) = 0,8 ,$$

mientras que

$$C_m^{T_M}(0,05, 0,1, 0,8, 0,9) = 0,7 ,$$

luego $C_m^{T_M}$ no es creciente, y, por tanto, no es una función de agregación.

- (ii) Consideremos la t -norma de Lukasiewicz $T_L(x, y) = \max\{0, x + y - 1\}$. De nuevo, para $N = \{1, 2, 3, 4\}$ y la medida uniforme $\mathbf{m} = \mathbf{m}_U$, tenemos que

$$C_m^{T_L}(0,05, 0,1, 0,7, 0,9) = 0,15 ,$$

mientras que

$$C_m^{T_L}(0,05, 0,2, 0,7, 0,9) = 0,05 ,$$

luego $C_m^{T_L}$ no es creciente, y, por tanto, no es una función de agregación.

- (iii) Respecto al producto drástico $T_{DP}(x, y) = \min\{x, y\}$, si $\max\{x, y\} = 1$ y 0 en otro caso, sea $N = \{1, 2, 3\}$ y la medida uniforme $\mathbf{m} = \mathbf{m}_U$. Entonces

$$C_m^{T_{DP}}(0, 0, 1) = 0,33 ,$$

mientras que

$$C_m^{T_{NM}}(0, 0,5, 1) = 0 ,$$

luego $C_m^{T_{NM}}$ no es creciente, y, por tanto, no es una función de agregación.

- (iv) Consideremos la t -norma nilpotente $T_{NM}(x, y) = \min\{x, y\}$ si $x + y > 1$ y 0 en otro caso. De nuevo, para $N = \{1, 2, 3, 4\}$ y la medida uniforme $\mathbf{m} = \mathbf{m}_U$ tenemos que

$$C_m^{T_{NM}}(0,05, 0,1, 0,7, 0,9) = 0,55 ,$$

mientras que

$$C_m^{T_{NM}}(0,05, 0,2, 0,7, 0,9) = 0,5 ,$$

luego $C_m^{T_{NM}}$ no es creciente, y, por tanto, no es una función de agregación.

- (v) Consideremos el producto de Hamacher $T_{HP}(x, y) = \frac{xy}{x+y-xy}$ si $(x, y) \neq (0, 0)$ y $T(0, 0) = 0$. Si consideramos $N = \{1, 2, 3, 4\}$ y la medida uniforme $\mathbf{m} = \mathbf{m}_U$ dada en (1), vemos que

$$C_m^{T_{HP}}(0,05, 0,1, 0,7, 0,9) = 0,5991 ,$$

mientras que

$$C_m^{T_{HP}}(0,05, 0,1, 0,8, 0,9) = 0,5877 ,$$

luego $C_m^{T_{HP}}$ no es creciente, y, por tanto, no es una función de agregación.

5. Conclusiones

En este trabajo, hemos introducido el concepto de función de pre-agregación. Hemos presentado un método de construcción de tales funciones basado en el de la integral Choquet utilizando otras t-normas en lugar del producto.

En [16] se ha demostrado experimentalmente que el uso de este tipo de pre-agregaciones permite mejorar bajo determinadas circunstancias el comportamiento de métodos de clasificación como el FARC-HD. En el futuro, pretendemos estudiar métodos de construcción de pre-agregaciones más generales, con la vista puesta en su posible utilidad para aplicaciones.

6. Agradecimientos

Este trabajo ha sido financiado por los proyectos TIN2013-40765-P y TIN2014-56381-REDT del Gobierno de España.

Referencias

1. Beliakov, G., Pradera, A., Calvo, T.: *Aggregation Functions: A Guide for Practitioners*. Springer, Berlin (2007)
2. Grabisch, M., Marichal, J., Mesiar, R., Pap, E.: *Aggregation Functions*. Cambridge University Press, Cambridge (2009)
3. Amo, A., Montero, J., Biging, G., Cutello, V.: Fuzzy classification systems. *European Journal of Operational Research* **156**(2) (2004) 495 – 507
4. Bustince, H., Barrenechea, E., Calvo, T., James, S., Beliakov, G.: Consensus in multi-expert decision making problems using penalty functions defined over a cartesian product of lattices. *Information Fusion* **17** (2014) 56–64
5. Bustince, H., Pagola, M., Mesiar, R., Hüllermeier, E., Herrera, F.: Grouping, overlaps, and generalized bientropic functions for fuzzy modeling of pairwise comparisons. *IEEE Transactions on Fuzzy Systems* **20**(3) (2012) 405–415
6. Bustince, H., Montero, J., Mesiar, R.: Migrativity of aggregation operators. *Fuzzy Sets and Systems* **160**(6) (2009) 766–777
7. Bustince, H., Fernandez, J., Mesiar, R., Montero, J., Orduna, R.: Overlap functions. *Nonlinear Analysis* **72**(3-4) (2010) 1488–1499
8. Alcalá-Fdez, J., Alcalá, R., Herrera, F.: A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. *IEEE Transactions on Fuzzy Systems* **19**(5) (2011) 857–872
9. Jurio, A., Bustince, H., Pagola, M., Pradera, A., Yager, R.: Some properties of overlap and grouping functions and their application to image thresholding. *Fuzzy Sets and Systems* **229** (2013) 69 – 90
10. Wilkin, T., Beliakov, G.: Weakly monotone aggregation functions. *International Journal of Intelligent Systems* **30** (2015) 144–169
11. Bustince, H., Kolesárová, A., Fernandez, J., Mesiar, R.: Directional monotonicity of fusion functions. *European Journal of Operational Research* **244**(1) (2015) 1150–1162
12. Choquet, G.: Theory of capacities. *Annales de l'Institut Fourier* **5** (1953–1954) 131–295

13. Denneberg, D.: Representation of the Choquet integral with the 6-additive mobius transform. *Fuzzy Sets and Systems* **92**(2) (1997) 139 – 156
14. Grabisch, M., Labreuche, C.: A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *Annals of Operations Research* **175**(1) (2010) 247–286
15. Gilboa, I., Schmeidler, D.: Additive representations of non-additive measures and the Choquet integral. *Annals of Operations Research* **52**(1) (1994) 43–65
16. Lucca, G., Sanz, J., Dimuro, G.P., Bedregal, B., Mesiar, R., Kolesárová, A., Bustince, H.: Pre-aggregation functions: construction and an application. *IEEE Transactions on Fuzzy Systems* (accepted for publication)
17. Barrenechea, E., Bustince, H., Fernandez, J., Paternain, D., Sanz, J.A.: Using the Choquet integral in the fuzzy reasoning method of fuzzy rule-based classification systems. *Axioms* **2**(2) (2013) 208–223
18. Bustince, H., Burillo, P., Soria, F.: Automorphisms, negations and implication operators. *Fuzzy Sets and Systems* **134**(2) (2013) 209–229