CRITERIOS DE COHERENCIA EN PREFERENCIAS INTENSAS

José Luis García Lapresta

RESUMEN.— Este trabajo consta de dos partes bien diferenciadas. En la primera se explican las nociones de subconjunto borroso y de relación binaria borrosa, se fundamenta la teoría de subconjuntos borrosos de Zadeh en la lógica multivalente de Łukasiewicz y se marcan las diferencias entre esta teoría y la de la probabilidad. Se definen varias nociones acerca de las relaciones binarias borrosas, con el fin de que éstas sirvan de modelo de las preferencias de los agentes cuando se manifiestan con intensidad. En la segunda parte se presentan varios criterios de coherencia interna en preferencia borrosas, en el supuesto de que se verifique el axioma de reciprocidad, destacándose de cada uno de ellos tanto los aspectos positivos como los negativos. Por fin se propone aquí un nuevo criterio: el de transitividad máx-mín moderada estricta, que permite capturar un comportamiento racional básico dentro del contexto estudiado.

1. INTRODUCCIÓN

Uno de los pilares en los que se fundamenta la teoría de la elección social, así como otras teorías normativas sobre la toma de decisiones, es el supuesto de que los agentes tienen un comportamiento racional. Ha sido costumbre considerar que los agentes son capaces de mostrar sus preferencias sobre los pares de objetos existentes mediante relaciones binarias que cumplen diversos axiomas. Algunos de ellos han sido reconsiderados a lo largo del tiempo atendiendo a experimentos hechos sobre la coherencia en las predilecciones humanas. Podría decirse que la teoría positiva ha influido en la normativa, haciendo desaparecer o debilitar axiomas demasiado restrictivos por los cuales pocos individuos resultaban ser racionales o coherentes.

Cuando se pretende que un individuo descubra su disposición entre pares de objetos o situaciones, de forma que dados dos de ellos tenga que mostrar cuál de ellos prefiere o bien no decantarse por ninguno y, en ese caso, manifestarse indiferente, puede ocurrir que pronunciamientos similares provengan de tendencias muy distintas. Por ejemplo, ante dos modelos de automóvil, un individuo puede preferir uno de ellos tras una larga duda, después de sopesar concienzudamente los pros y los contras de ambos modelos. El mismo individuo ante una cámara de vídeo y una casa de campo puede confesar su preferencia de forma clara y rotunda. Sin embargo, en ambos casos sólo trasciende el hecho de que uno de los bienes es preferido al otro. Parece lógico desear conocer, no sólo qué objetos son preferidos a otros, sino la intensidad de estas preferencias. Está claro que las relaciones binarias ordinarias sólo modelizan preferencias absolutas y no recogen las intensidades con las que unas opciones son preferidas a otras. Para poder corregir este defecto y contemplar así este aspecto destacado de las inclinaciones humanas, las relaciones binarias borrosas constituyen un instrumento eficaz.

2. RELACIONES BINARIAS BORROSAS

En la teoría ordinaria de conjuntos las relaciones binarias son subconjuntos del producto cartesiano de un conjunto por sí mismo. Por analogía, las relaciones binarias borrosas se definen como subconjuntos borrosos del producto cartesiano de un conjunto ordinario por sí mismo. Por ello, si se desea modelizar la intensidad de las preferencias de los agentes mediante relaciones binarias borrosas, se hace necesario conocer en profundidad la noción de subconjunto borroso.

2.1. Subconjuntos borrosos

La teoría de los *subconjuntos borrosos (fuzzy sets)*, creada por el matemático Lofti A. Zadeh en 1965, cuenta en la actualidad con centenares de artículos publicados y una revista: *Fuzzy Sets and Systems*, dedicada en su totalidad al desarrollo de esta teoría, así como a sus aplicaciones a las ciencias, la técnica y las humanidades.

En la lógica clásica toda proposición es o bien verdadera o bien falsa; cuando es verdadera se le asocia valor de verdad 1 y cuando es falsa 0. Esta lógica es la usada con mayor asiduidad en las ciencias exactas. Ahora bien, en el terreno de las ciencias humanas algunas proposiciones no se someten a esta bivalencia, so pena de simplificar y alterar en exceso su naturaleza[1].

1 Bertrand Russell ya advertía que usar lógica clásica en las ciencias humanas supone una idealización de la realidad; según él, esta lógica trata no de esta «vida terrenal», sino de una imaginada «existencia celestial».

En 1920 el lógico polaco Jan Lukasiewicz abría paso a la consideración de una lógica trivalente en la que además de los valores de verdad clásicos, verdadero-falso, había un tercero: *posible* o *indeterminado.* En 1930 el propio Lukasiewicz, junto a Alfred Tarski, generalizaron la lógica trivalente a una lógica multivalente con toda una gama ordenada de valores de verdad entre 0 (falso) y 1 (verdadero). La teoría de los subconjuntos borrosos puede considerarse, en un primer análisis, como la aplicación de la lógica multivalente al campo de la teoría de conjuntos.

Dado un conjunto no vacío X, llamado *universo,* en la teoría ordinaria de conjuntos, cualquiera que sea el subconjunto A de X, todo elemento de X o bien es de A o, por el contrario, no lo es; no existen situaciones intermedias. Esto permite que los conjuntos finitos puedan definirse por extensión, detallando en una lista los elementos que forman parte de él. Puede ocurrir que en un momento dado no se sepa si un elemento pertenece a un subconjunto; pero este problema afecta al estado de conocimiento, no al elemento que, sin duda, será o no será de ese subconjunto. En esta teoría todo subconjunto A del universo X tiene asociada su *función característica:* $\mu_A: X \rightarrow \{0, 1\}$, definida por:

$$\mu_A(x) = \begin{cases} 1, & si\ x \in A \\ 0, & si\ x \notin A. \end{cases}$$

Mientras que aquí la transición entre pertenencia y no pertenencia es abrupta, en la teoría de subconjuntos borrosos es gradual. La función característica de un *subconjunto borroso*[2] A del universo X es una aplicación $\mu_A: X \rightarrow [0, 1]$, donde dado $x \in X$, $\mu_A(x)$ es el *grado de pertenencia*[3] de x a A. Queda claro que no es posible definir un subconjunto borroso por extensión, ya que hay muchas maneras de «pertenecer» a un subconjunto borroso. Estos vienen dados directamente por sus funciones características. Obviamente, todo subconjunto ordinario es también borroso: $\mu_A(X) \subseteq \{0, 1\} \subseteq [0, 1]$. Además, un subconjunto borroso A de X es ordinario si y sólo si $\mu_A(X) \subseteq \{0, 1\}$. Por todo ello, puede afirmarse que la noción de subconjunto borroso generaliza la de subconjunto ordinario[4].

La utilización de esta escala numérica entre 0 y 1 para indicar el grado de pertenencia, por lo general asignado subjetivamente por un individuo o colectivo, permite reflejar un cierto ordenamiento en los objetos comparados, tan importante o más que los propios valores que señalan el grado de

2 De la propia definición se sigue que todo subconjunto borroso lo es de un conjunto ordinario (el universo); en consecuencia, parece más apropiado hablar de «subconjuntos» borrosos, en lugar de «conjuntos» borrosos.

3 Podría interpretarse $\mu_A(x)$ como el valor de verdad de la proposición «$x \in A$» en la lógica multivalente de Łukasiewicz.

4 Sobre la generalización de las operaciones y relaciones conjuntistas en el marco de la teoría de subconjuntos borrosos puede consultarse [Dubois-Prade].

pertenencia. Otra tradición, iniciada por Goguen en 1967, consiste en valorar las funciones de pertenencia sobre retículos o sobre conjuntos ordenados de etiquetas lingüísticas (por ejemplo, «nada», «poco», «ligeramente», «bastante», «mucho», «totalmente», etc.).

La noción de subconjunto borroso puede traer resonancias probabilísticas. Si se toma un álgebra de sucesos correspondiente a un experimento aleatorio como universo, entonces toda probabilidad sobre él puede considerarse como la función característica de un subconjunto borroso. Ahora bien, no todo subconjunto borroso puede interpretarse en estos términos, como puede observarse en el siguiente ejemplo.

Una sociedad filantrópica de un país tiene cuatro tipos de miembros:

- Directivos.
- Afiliados.
- Colaboradores.
- Simpatizantes.

Con la asignación de grados de pertenencia 1, 3/4, 1/2 y 1/4, respectivamente, a estas cuatro modalidades de formar parte de la citada sociedad y 0 en caso de no formar parte de ella, queda definida la función característica de un subconjunto borroso del conjunto de ciudadanos de ese país: el de las personas que constituyen la sociedad filantrópica. No es lo mismo que una persona sea de la citada sociedad en grado 1/4, es decir que lo es en la categoría de miembro simpatizante, que pertenecer a esa sociedad con probabilidad 1/4. En el primer caso habría certeza de que esa persona era miembro de la sociedad; en el segundo, no.

Por otra parte, si se considera como universo el conjunto formado por los máximos líderes de los actuales partidos políticos con representación parlamentaria, los ciudadanos interesados en la política tienen sus propias opiniones sobre la honradez de estos políticos. Si se le pide a uno de ellos que califique a cada uno de estos políticos con un número de 1 a 0, según piense que son más o menos honrados, resultará que esta asignación define un subconjunto borroso del conjunto de los líderes políticos mencionados. En este sentido podría asegurarse que la noción de honradez es *borrosa*. No menos podría decirse del concepto de preferencia, referido a un individuo en una situación y momento concretos. Las preferencias e inclinaciones individuales se manifiestan con mayor o menor intensidad, matizando sus tendencias en una amplia gama de grados. Sólo ocasionalmente se presentan de forma taxativa; no obstante, lo taxativo constituye un caso particular de lo intenso: aquél en el que todo se polariza en los dos extremos —todo o nada—, de la misma manera que los subconjuntos ordinarios son también borrosos.

2.2. Relaciones binarias borrosas

En concordancia con la noción de subconjunto borroso puede definirse la de relación binaria borrosa. Si en la teoría ordinaria de conjuntos una relación binaria R sobre X es un subconjunto de X x X, una *relación binaria borrosa* sobre X es un subconjunto borroso de X x X; así, μ_R: X x $X \rightarrow [0, 1\}$ será la función característica de la relación binaria borrosa R y $\mu_R(x, y)$ la *intensidad* con la que x está relacionado con y.

En ocasiones se ha relacionado la noción de borrosidad con la de imprecisión, como si todo aquello que estuviera ligado a conceptos borrosos conllevara un conocimiento imperfecto, cargado de inexactitud. Todo lo contrario. El comportamiento humano, cuyas tendencias e inclinaciones están provistas de intensidad, no puede ser fielmente reflejado mediante corsés excesivamente rígidos; esto resulta mucho más impreciso, por su excesiva simplificación, que hacerlo mediante minuciosos conceptos de la teoría de subconjuntos borrosos, que capturan con rigor las intensidades de las predilecciones humanas. Es más preciso conocer el nivel de intensidad con el que un individuo prefiere un bien a otro que saber tan sólo qué objeto es preferido por este agente. Las preferencias humanas son intrínsecamente intensas, por tanto parece lógico que sean modelizadas mediante relaciones binarias borrosas.

2.2.1. *Representación matricial*

A toda relación binaria borrosa R sobre un universo finito $X = \{x_1, ..., x_n\}$ se le puede asociar una matriz n x n sobre [0, 1]: $M_R = (r_{ij})_{1 \leq i, j \leq n}$, donde $r_{ij} = \mu_R(x_i, x_j) \in [0, 1]$. Recíprocamente, dada una matriz $M = (r_{ij})_{1 \leq i, j \leq n}$ de elementos de [0, 1], le corresponde de forma natural una relación binaria borrosa R_M sobre X cuya función característica viene definida por $\mu_{R_M}(x_i, x_j) = r_{ij}$. Ambas asignaciones son inversas la una de la otra. Hay, por tanto, una biyección entre el conjunto de relaciones binarias borrosas sobre X y el de matrices $n \times n$ sobre [0, 1]: $\mathcal{M}_{n \times n}([0, 1])$.

2.2.2. *Reciprocidad*

Si R es una relación binaria borrosa que representa las preferencias de un agente individual o colectivo sobre un conjunto de opciones $X = \{x_1, ..., x_n\}$, cuanto mayor sea la intensidad $r_{ij} = \mu_R(x_i, x_j)$ con la que x_i es preferido a x_j, tanto menor será la intensidad con la que x_j es preferido a x_i: $r_{ji} = \mu_R(x_j, x_i)$. Es más, podría interpretarse que el agente distribuye toda su capacidad de preferir: 1, entre cada par de opciones. Esta idea puede recogerse en el siguiente axioma que, a partir de ahora, se supondrá que es cumplido por toda relación binaria borrosa que modelice las preferencias individuales o colectivas.

Axioma de reciprocidad: $r_{ij} + r_{ji} = 1 \;\; \forall i, j \in \{1, ..., n\}$.

Resulta inmediato comprobar que el conjunto de las relaciones binarias borrosas sobre X que satisfacen el axioma de reciprocidad está en biyección con el conjunto de matrices

$$\{M \in \mathcal{M}_{nxn}([0, 1]) \mid M + M' = (1)\},$$

donde M' es la matriz traspuesta de M y (1) la matriz de $\mathcal{M}_{nxn}([0, 1])$ cuyos elementos son todos 1 (obsérvese que ambos conjuntos son convexos).

En consecuencia, para cada par de opciones x_i, x_j se pueden presentar tres casos:

1. $r_{ij} > 1/2$: el agente inclina sus preferencias hacia x_i en detrimento de x_j o, simplemente, prefiere x_i a x_j.
2. $r_{ij} < 1/2$: por el axioma de reciprocidad, este hecho equivale a $r_{ij} > 1/2$, es decir, prefiere x_j a x_i.
3. $r_{ij} = 1/2$: no se da ninguna de las situaciones anteriores, por tanto se muestra indiferente ante las opciones x_i y x_j.

En este sentido, cabe decir que el axioma de reciprocidad garantiza la *completitud:* cualesquiera que sean las opciones x_i, x_j, o bien x_i es preferida a x_j, o bien x_j es preferida a x_i o, por el contrario, son indiferentes.

Este axioma ha sido considerado en la literatura desde diferentes puntos de vista. En este sentido han de destacarse los trabajos pioneros de Davidson-Marschak (1959), Debreu (1958) y Georgescu-Roegen (1958). En ellos se establece el axioma de reciprocidad en términos probabilísticos. Se supone que, dadas dos opciones x_i y x_j, el agente se decide por una de las dos: x_i —con probabilidad r_{ij}— o x_j —con probabilidad r_{ji}—; al no existir más casos, se obtiene $r_{ij} + r_{ji} = 1$. Tanino, en 1984, recupera este axioma fuera ya del contexto probabilístico, enmarcándolo en la teoría de subconjuntos borrosos, con una perspectiva similar a la utilizada en el presente trabajo.

2.2.3. *Relaciones binarias ordinarias asociadas*

En concordancia con la terminología de la teoría de subconjuntos borrosos, dado $\alpha \in [0, 1]$, se define la *relación binaria ordinaria de nivel* α (o α-*corte*) asociada a R como

$$R_\alpha = \mu^{-1}_R([\alpha, 1]) = \{(x_i, x_j) \in X x X \mid r_{ij} \geq \alpha\}.$$

Asimismo, la *relación binaria ordinaria de nivel estricto* α viene definida por

$$R_\alpha = \mu^{-1}_R((\alpha, 1]) = \{(x_i, x_j) \in X x X \mid r_{ij} > \alpha\}.$$

De acuerdo con esta notación, pueden considerarse las siguientes relaciones binarias asociadas a R:

1. Preferencia absoluta
   $R_1 = \{(x_i, x_j) \in X \times X \mid r_{ij} = 1\}$.
2. Preferencia relativa
   $P = R_{\overline{1/2}} = \{(x_i, x_j) \in X \times X \mid r_{ij} > 1/2\}$.
3. Indiferencia[5]
   $I = R_{1/2} \cap R^{-1}_{1/2} = \mu^{-1}_R(\{1/2\}) = \{(x_i, x_j) \in X \times X \mid r_{ij} = 1/2\}$.

De esta definición se sigue que $<P, I>$ es una estructura preferencial sobre X (véase [García Lapresta (a)]):

1. P es asimétrica:
   $(x_i, x_j) \in P \Rightarrow r_{ij} > 1/2 \Rightarrow r_{ji} < 1/2 \Rightarrow (x_j, x_i) \notin P$.
2. I es reflexiva:
   $r_{ii} + r_{ii} = 1 \Rightarrow r_{ii} = 1/2 \Rightarrow (x_i, x_i) \in I$.
3. I es simétrica:
   $(x_i, x_j) \in I \Rightarrow r_{ij} = 1/2 \Rightarrow r_{ji} = 1/2 \Rightarrow (x_j, x_i) \in I$.
4. P y I son incompatibles:
   $(x_i, x_j) \in P \Rightarrow r_{ij} > 1/2 \Rightarrow r_{ij} \neq 1/2 \Rightarrow (x_i, x_j) \notin I$.
5. $P \cup P^{-1} \cup I = X \times X$:
   Corresponde a la completitud, ya estudiada.

Así, siempre que se considere pertinente, se tomará la estructura preferencial ordinaria asociada a la relación binaria borrosa en cuestión. Ahora bien, dar este paso supone una pérdida de información, ya que la estructura preferencial sólo recoge qué opciones son preferidas a otras, no cómo son preferidas.

3. ANÁLISIS DE VARIOS CRITERIOS DE COHERENCIA

Pueden encontrarse en la literatura diversos supuestos sobre las relaciones binarias borrosas utilizadas en modelos de coherencia interna. Algunos de ellos han sido establecidos con el fin de modelizar el comportamiento racional de los agentes, cuando estos manifiestan sus predilecciones sobre pares de opciones, y otros fuera de este contexto. Tras las críticas pertinentes a varios criterios recogidos de la literatura, se propondrá aquí uno nuevo: el de «transitividad máx-mín moderada estricta». Este axioma, junto al de reciprocidad, permiten modelizar satisfactoriamente,

5 La relación binaria inversa de R (ordinaria o borrosa): R^{-1}, viene definida por:

$$\mu_{R^{-1}}(x_i, x_j) = \mu_R(x_j, x_i).$$

desde nuestro punto de vista, la coherencia interna de las preferencias que se manifiestan con intensidad.

En todos los casos se supondrá que se verifica el axioma de reciprocidad y, en su compañía, se analizarán los aspectos positivos y negativos de cada uno de los criterios expuestos.

3.1. CRITERIOS CLÁSICOS

La transitividad de la relación de preferencia no estricta, $P \cup I$, ha sido una hipótesis convencional de comportamiento racional en todo el análisis económico, principalmente desde su uso por von Neumann-Morgenstern y Arrow. Con el paso del tiempo, este supuesto ha sido seriamente criticado y contradicho por numerosos experimentos realizados. Una de las relajaciones más habituales ha consistido en abandonar la transitividad de I y retenerla para P. A este respecto puede consultarse el ameno trabajo de Blair-Pollack.

3.1.1. *Transitividad de* $P \cup I$

Este criterio viene determinado por el hecho de que $R_{1/2} = P \cup I$ es transitiva:

$$(r_{ij} \geq 1/2 \text{ y } r_{jk} \geq 1/2) \Rightarrow r_{ik} \geq 1/2 \quad \forall i, j, k \in \{1, ..., n\}.$$

Ha sido defendido, entre otros, por Billot y Luo, en ausencia del axioma de reciprocidad, como criterio de racionalidad cuando las preferencias son borrosas.

Puede comprobarse fácilmente que, cuando se supone el axioma de reciprocidad, la transitividad de $P \cup I$ entraña la de la relación de indiferencia I. Esto hace que el criterio sea muy restrictivo, ya que la capacidad humana de discernimiento está limitada por umbrales de percepción. Por ejemplo, un individuo al que le presentan tres trajes de igual factura pero de colores x_1, x_2 y x_3, ligeramente distintos, puede no captar diferencias entre x_1 y x_2 ni entre x_2 y x_3 y, por tanto, mostrarse indiferente ante ambos pares de trajes y, sin embargo, captar diferencias entre x_1 y x_3, de forma que prefiera uno a otro.

Por otra parte, el criterio es demasiado tolerante, ya que permite que haya caídas en el nivel de intensidad. Por ejemplo, si un ciudadano prefiere que en su país gobierne un partido socialdemócrata a que lo haga uno conservador y, a su vez, que lo haga un partido conservador a uno de extrema derecha, en ambos casos de forma absoluta, no destacaría por su coherencia si se mostrase indiferente ante gobiernos socialdemócratas o de extrema derecha. Pues bien, el presente criterio permitiría que se diera este caso ($r_{ij} = r_{jk} = 1$ y $r_{ik} = 1/2$).

3.1.2. *Transitividad de* P

Este nuevo criterio viene caracterizado por el hecho de que la relación de preferencia relativa, $R_{\overline{1/2}} = P$, es transitiva:

$$(r_{ij} > 1/2 \text{ y } r_{jk} > 1/2 \Rightarrow r_{ik} > 1/2 \quad \forall i, j, k \in \{1, ..., n\}.$$

Aunque de aquí no se deduce la transitividad de I, existen prácticamente los mismos inconvenientes que en el criterio anterior, pues también permite caídas en el nivel de intensidad. Por ejemplo, consentiría la siguiente situación: $r_{ij} = r_{jk} = 1$ y $r_{ik} = 0.6$.

3.2. TRANSITIVIDAD MÁX-MÍN

Este criterio corresponde a la noción de transitividad borrosa, propuesta por Zadeh en 1971 y utilizada frecuentemente en la literatura. A continuación se presentan cuatro formas equivalentes de esta propiedad, si bien la última de ellas sólo coincide con las anteriores en presencia del axioma de reciprocidad.

PROPOSICIÓN

Son equivalentes:

1. R_α es transitiva $\forall \alpha \in [0, 1]$,
2. $r_{ik} \geq$ máx { mín $\{r_{ij}, r_{jk}\} \mid j \in \{1, ..., n\}\} \; \forall i, k \in \{1, ..., n\}$,
3. $r_{ik} \geq$ mín $\{r_{ij}, r_{jk}\} \; \forall i, j, k \in \{1, ..., n\}$,
4. mín $\{r_{ij}, r_{jk}\} \leq r_{ik} \leq$ máx $\{r_{ij}, r_{jk}\} \; \forall i, j, k \in \{1, ..., n\}$.

Demostración: es obvio que 2 y 3 son equivalentes.

$1 \Rightarrow 3$: si se toma $\alpha =$ mín $\{r_{ij}, r_{jk}\}$, se tiene $r_{ij}, r_{jk} \geq \alpha$ o, lo que es lo mismo, $(x_i, x_j) \in R_\alpha$ y $(x_j, x_k) \in R_\alpha$; entonces, de la hipótesis se sigue $(x_i, x_k) \in R_\alpha$, es decir, $r_{ik} \geq \alpha =$ mín $\{r_{ij}, r_{jk}\}$.

$3 \Rightarrow 4$: bastará probar la otra desigualdad. La hipótesis asegura que se verifica $r_{ki} \geq$ mín $\{r_{kj}, r_{ji}\}$ para cualesquiera $i, j, k \in \{1, ..., n\}$. Cambiando de signo se obtiene:

$$-r_{ki} \leq -\text{mín}\,\{r_{kj}, r_{ji}\} = \text{máx}\,\{r_{kj}, -r_{ji}\} \quad \forall i, j, k \in \{1, ..., n\}.$$

Sumando 1 a ambos términos de la desigualdad se tiene:

$$r_{ik} = 1 - r_{ki} \leq 1 + \text{máx}\,\{-r_{kj}, -r_{ji}\} = \text{máx}\,\{1 - r_{kj}, 1 - r_{ji}\} = \text{máx}\,\{r_{jk}, r_{ij}\}$$

para cualesquiera $i, j, k \in \{1, ..., n\}$, es decir,

$$r_{ik} \leq \text{máx}\,\{r_{ij}, r_{jk}\} \; \forall i, j, k \in \{1, ..., n\}.$$

4 $\Rightarrow$ 1: si $(x_i, x_j) \in R_\alpha$ y $(x_j, x_k) \in R_\alpha$, se tiene $r_{ij}, r_{jk} \geq \alpha$; entonces, por hipótesis, $r_{ik} \geq \text{mín}\ \{r_{ij}, r_{jk}\} \geq \alpha$; es decir, $(x_i, x_k) \in R_\alpha$.

Como aspectos positivos de este criterio cabe destacar el hecho de que P es necesariamente transitiva, así como la prohibición de que haya caídas de intensidad: si x_1 es preferido a x_2, con nivel de intensidad $r_{12} > 1/2$ y, a su vez, x_2 es preferido a x_3, con nivel de intensidad $r_{23} > 1/2$, entonces x_1 será preferido a x_3 con intensidad no menor que el más bajo de los niveles r_{12}, r_{23}.

Sin embargo, de este criterio también se sigue la transitividad de $R_{1/2}$ y, por tanto, la de I. Además, impide la acumulación de la intensidad de las preferencias. Por ejemplo, ante tres opciones x_1, x_2 y x_3, si un agente prefiere x_1 a x_2 y x_2 a x_3, en ambos casos con grado 0.8, la cuarta versión de este criterio impide que x_1 sea preferido a x_3 con grado mayor a 0.8. A este respecto, parece razonable admitir que si un trabajador prefiere con intensidad elevada tener turno de mañana a tenerlo de tarde y, con la misma intensidad, tenerlo de tarde a tenerlo de noche, desee a su vez, con total convicción, trabajar por la mañana a hacerlo en turno de noche, sin que por ello se le tache de incoherente.

3.3. TRANSITIVIDAD MÁX-MÍN MODERADA

El criterio de transitividad máx-mín ha sido desechado en parte por impedir que se acumulen las intensidades de las preferencias. Este aspecto indeseable desaparece cuando se restringe su aplicación a los casos en los que hay preferencia relativa (o indiferencia) entre las opciones.

3.3.1. *Transitividad máx-mín moderada no estricta*

Este criterio ha sido considerado, junto al axioma de reciprocidad, por Tanino en 1984. Con anterioridad, Luce y Suppes lo tuvieron en cuenta, en 1965, dentro del marco de la teoría de la elección probabilística.

En la siguiente proposición, que se demuestra de forma análoga a la incluida en 3.2, se exponen tres formas equivalentes de este criterio.

PROPOSICIÓN

Son equivalentes:

1. R_α es transitiva $\forall \alpha \in [1/2, 1]$,
2. $r_{ik} \geq \text{máx}\ \{\text{mín}\ \{r_{ij}, r_{jk}\} \mid j \in \{1, ..., n\},\ r_{ij} \geq 1/2\ ,\ r_{jk} \geq 1/2\}$ $\forall i, k \in \{1, ..., n\}$
3. $(r_{ij} \geq 1/2$ y $r_{jk} \geq 1/2) \Rightarrow r_{ik} \geq \text{mín}\ \{r_{ij}, r_{jk}\}\ \forall i, j, k \in \{1, ..., n\}$.

Bajo este supuesto se hace posible la acumulación de intensidad en las preferencias, ya que con la moderación impuesta no es posible caracterizar este criterio de manera análoga al apartado 4 de la proposición demostrada en 3.2. Así mismo, la relación de preferencia relativa asociada, *P*, resulta transitiva. Sin embargo, resulta demasiado restrictivo, puesto que conlleva la transitividad de la relación de indiferencia.

Han sido estudiados varios supuestos de transitividad. En todos ellos se han detectado aspectos negativos que los hacen poco o nada plausibles como criterios de coherencia interna en la toma de decisiones, cuando éstas se basan en preferencias intensas. A continuación se propone un nuevo criterio que goza de todas las propiedades deseables, aparecidas hasta ahora y, en cambio, elimina aquellas que han sido desechadas para modelizar la racionalidad de los agentes en el contexto presentado.

3.3.2. *Transitividad máx-mín moderada estricta*

Este nuevo criterio, expuesto también a través de tres formas equivalentes, conserva los aspectos positivos del anterior y, además, elimina la obligatoriedad de que la relación de indiferencia sea transitiva.

PROPOSICIÓN

Son equivalentes:

1. R_α es transitiva $\forall \alpha \in (1/2, 1]$,
2. $r_{ik} \geq \text{máx}\ \{\text{mín}\ \{r_{ij}, r_{jk}\} \mid j \in \{1, ..., n\}\ r_{ij} > 1/2,\ r_{jk} > 1/2\}$ $\forall i, k \in \{1, ..., n\}$
3. $(r_{ij} > 1/2 \text{ y } r_{jk} > 1/2) \Rightarrow r_{ik} \geq \text{mín}\ \{r_{ij}, r_{jk}\}\ \forall i, j, k \in \{1, ..., n\}$.

Por las razones expuestas anteriormente, este supuesto junto al de reciprocidad permiten modelizar de forma satisfactoria un comportamiento racional básico de los agentes que manifiestan sus preferencias con intensidad.

Podría desearse un criterio de coherencia que exigiera un mayor compromiso por parte de los agentes, de forma que no sólo se impidieran las caídas en los niveles de intensidad de las preferencias, sino que además obligara a que se acumulasen en cierta medida. En [García Lapresta (b)] se estudian varios criterios más restrictivos que el de transitividad máx-mín moderada estricta, en los que se imponen acumulaciones crecientes en los niveles de intensidad de las preferencias. Aunque todos ellos parecen plausibles, desde un punto de vista normativo, análisis experimentales efectuados muestran un elevado incumplimiento de estos criterios. En ello intervienen, sin duda, las limitadas capacidades humanas de concentración, memoria y decisión.

BIBLIOGRAFÍA

Armstrong, W. E.: «A note on the theory of consumer's behavior». *Oxford Economic Papers* 2. 1950, pp. 119-122.

Arrow, K. J.: *Social Choice and Individual Values*. Segunda edición: Yale University Press. New Hagen, 1963 (existe traducción al castellano con una introducción de Andreu Mas-Colell: *Elección Social y Valores Individuales*. Instituto de Estudios Fiscales. Madrid, 1974).

Barrett, C. R.; Pattanaik, P. K.; Salles, M.: «On the structure of fuzzy social welfare functions». *Fuzzy Sets and Systems* 19. 1986, pp. 1-10.

Basu, K.: «Fuzzy revealed preference theory». *Journal of Economic Theory* 32. 1984, pp. 212-227.

Billot, A.: «Aggregation of preferences: the fuzzy case». *Theory and Decision* 30. 1991, pp. 51-93.

Blair, D. H.; Pollack, R. A.: «Decisiones racionales colectivas». *Investigación y Ciencia*. Octubre, 1983, pp. 64-72.

Davidson, D.; Marschak, J.: «Experimental tests of stochastic decision theory» en *Measurement Definitions and Theories*. Editor C. West Churchman. John Wiley and Sons. New York, 1959.

Debreu, G.: «Stochastic choice and cardinality utility». *Econometrica* 26. 1958, pp. 440-444.

Dubois, D.; Prade, H.: *Fuzzy Sets and Systems. Theory and Applications*. Academic Press. New York, 1980.

García Lapresta, J. L. (a): «Preferencia e indiferencia en Teoría de la Elección Social». *Anales de Estudios Económicos y Empresariales* 7, 1992, pp. 247-254.

García Lapresta, J. L. (b): «Un estudio teórico-experimental sobre la coherencia en la toma de decisiones intensas». *Estudios de Economía Aplicada*. VII Réunión Anual ASEPELT-ESPAÑA. 1993, pp. 401-410.

Georgescu-Roegen, N.: «Threshold in choice and theory of demand». *Econometrica* 26. 1958, pp. 157-168.

Goguen, J. A.: «L- fuzzy sets». *Journal of Mathematical Analysis and Applications* 18. 1967, pp. 145-174.

Luce, R. D.; Suppes, P.: «Preferences, utility and subjective probability» en *Handbook of Mathematical Psychology* III. Wiley. New York, 1965.

Łukasiewicz, J.: *Selected Works*. Editor L. Borkowski. North Holland. Amsterdam, 1970.

Łukasiewicz, J.; Tarski, A.: «Untersuchungen über den Aussagenkalkül». *Comptes rendus de la Société et des Lettres de Varsovie* 23. 1930, pp. 1-21.

Luo, C. Z.: «Fuzzy relation equation on infinite sets». *Busefal* 26. 1986, pp. 57-66.

Marschak, J.: «Norms and habits of decision making under certainty» en *Mathematical Models of Human Behavior*. Dunlap and Associates. Stamford, Connecticut. 1955, pp. 44-53.

May, K. O.: «Intransitivity, utility, and the aggregation of preference patterns». *Econometrica* 22. 1954, págs. 1-13.

Montero, F. J.; Tejada, J.: «Some problems on the definition of fuzzy preference relations». *Fuzzy Sets and Systems* 20. 1986, pp. 45-53.

Tanino, T.: «Fuzy preference orderings in group decision making». *Fuzzy Sets and Systems* 12. 1984, pp. 117-131.

Trillas, E.: *Conjuntos borrosos*. Vicens Vives. Barcelona, 1980.

Uriarte Ayo, J. R.: «Teoría de la decisión normativa versus Teoría de la decisión normativa versus Teoría de la decisión descriptiva». *Revista de Economía Pública* 13. 1991, pp. 85-112.

Von Neumann, J.; Morgenstern, O.: *Theory of Games and Economic Behavior*. Princeton University Press. Princeton, 1944.

Zadeh, L. A.: «Fuzzy sets». *Information and Control* 8. 1965, pp. 338-353.
Zadeh, L. A.: «Similitarity relations and fuzzy orderings». *Information Sciences* 3. 1971, pp. 177-200.
Zadeh, L. A.: «The concept of a lingüistic variable and its application to approximate reasoning». *Information Sciences*. Part I: 8, 1975, pp. 199-249. Part II: 8, 1975, pp. 301-357. Part III: 9, 1975, pp. 43-80.