

Rankings como un Problema de Orden Lineal Generalizado: Una aplicación a U-Ranking

Francisco J. Goerlich



Documento de trabajo U-Ranking 2/2020

Rankings como un Problema de Orden Lineal Generalizado: Una aplicación a U-Ranking

Francisco J. Goerlich

Universitat de València e Instituto Valenciano de Investigaciones Económicas (Ivie)

RESUMEN

Este trabajo plantea una solución alternativa a la comúnmente empleada en el problema de la agregación en la construcción de *rankings*. Dicha aproximación tiene ventajas e inconvenientes respecto a métodos más tradicionales, pero tiene un indudable atractivo, porque permite acomodar de forma alternativa cuestiones como la normalización de indicadores individuales o el tratamiento de la falta de información (*missing values*). Además, por construcción, no permite la compensación entre los indicadores individuales de partida. Aunque la literatura especializada presenta estos métodos como provenientes de la Teoría de la Elección Social, nuestra aproximación va directamente a la estructura matemática del problema, efectuando la conexión con dicha teoría a posteriori. Una aplicación a la construcción de *rankings* universitarios a partir de la base de datos de U-Ranking ilustra los métodos propuestos.

Palabras clave:

Rankings, Elección Social, Borda, Condorcet, Elección Multicriterio, Problema de Orden Lineal, U-Ranking, Universidades Clasificación JEL: D71, I23, C88

Versión (R1): Marzo/2020

Correspondencia:

Francisco J. Goerlich Gisbert, Universidad de Valencia, Departamento de Análisis Económico, Campus de Tarongers, Av. de Tarongers s/n, 46022-Valencia, España. E-mail: Francisco.J.Goerlich@uv.es.

Agradecimientos: Comentarios de Francisco Pérez, Irene Zaera, Inés Rosell y Rodrigo Aragón contribuyeron a mejorar la versión inicial de este trabajo. El autor agradece el apoyo financiero del Ministerio de Ciencia y Tecnología, proyecto ECO2015-70632-R, y del programa de investigación Fundación BBVA-Ivie. Resultados mencionados, pero no mostrados en el texto, y funciones en *R* para implementar los métodos expuestos en este trabajo, están disponibles a partir del autor. Los gráficos de este trabajo se ven mejor en la versión electrónica en color de este documento.

DOI: https://doi.org/10.12842/URANKING DT 22020

Este documento está disponible para su descarga en www.u-ranking.es

Índice

	ΓROE	

- 7 2. GUÍA BREVE PARA LA ELABORACIÓN DE INDICADORES COMPUESTOS
- 9 3. AGREGACIÓN: UN ENFOQUE ALTERNATIVO
 - 3.1. MARCO GENERAL
 - 3.2. ¿PODEMOS ACOMODAR PESOS?
 - 3.3. EMPATES (TIES) EN LOS INDICADORES INDIVIDUALES
 - 3.4. ¿QUÉ HACER CUÁNDO NO HAY INFORMACIÓN? (MISSING VALUES)
 - 3.5. VARIABLES CONTINUAS Y UMBRALES EN EL ESTABLECIMIENTO DE ORDENACIONES INDIVIDUALES
 - 3.6. EMPATES (TIES) EN EL ESPACIO DE SOLUCIONES
- 20 4. LA FORMULACIÓN MATEMÁTICA DEL PROBLEMA
 - 4.1. EL PROBLEMA COMPUTACIONAL
 - 4.2. SOLUCIONES MÚLTIPLES
- 23 5. ¿PODEMOS APRENDER ALGO DE LA TEORÍA DE LA ELECCIÓN SOCIAL?
- 26 6. AU-RANKING: U-RANKING ALTERNATIVO
 - 6.1. METODOLOGÍA DE U-RANKING
 - 6.2. U-RANKING COMO UN PROBLEMA DE ORDEN LINEAL
 - 6.3. ANÁLISIS EXPLORATORIO DE DATOS
 - 6.4. RESULTADOS Y COMPARACIÓN CON U-RANKING DE RENDIMIENTO
 - 6.5. RESULTADOS Y COMPARACIÓN CON U-RANKING PARA LAS DIMENSIONES
- 44 7. CONCLUSIONES
- 46 APÉNDICE
- 56 BIBLIOGRAFÍA

1. INTRODUCCIÓN

Los indicadores compuestos, y los *rankings* que de ellos se derivan, han ganado una tremenda popularidad en las últimas décadas. Los crecientes volúmenes de información y el énfasis en abordar los problemas desde una óptica multidimensional han forzado a los investigadores a buscar métodos de combinar esa información en un indicador único al que llamamos **indicador compuesto**. A partir de ellos es fácil derivar ordenaciones que traten de identificar a los mejores en un determinado ámbito. Muchos de estos indicadores han sido adoptados por instituciones como forma de monitorizar progresos en el logro de determinados objetivos.

Como consecuencia de este creciente uso, el interés, tanto por el público en general como por los investigadores en particular, en la elaboración de indicadores compuestos ha aumentado de forma notable. Una búsqueda en *Google Schoolar* del término 'composite indicator' en octubre de 2005 arrojó 992 resultados, en diciembre de 2010 esta cifra se había multiplicado por 5, hasta alcanzar los 5.340 resultados (Paruolo, Saisana y Saltelli 2013, p. 609), pero en la actualidad —octubre 2019— produce i1.6 millones de resultados! En el ámbito de las ciencias sociales existe incluso una revista académica expresamente dedicada a este tema: *Social Indicators Research*.¹

Esta creciente popularidad no ha ido acompañada por una convergencia en los métodos de producción de indicadores compuestos. Existe una serie de técnicas estadísticas aplicables en diferentes fases de la elaboración de los indicadores compuestos, algunas muy sencillas y otras terriblemente complejas, pero no existe una metodología básica acordada por los investigadores. Es posible, dada la naturaleza del problema, que dicho consenso no llegue a existir nunca.

Naturalmente esta falta de consenso dificulta las comparaciones, y es fácil encontrar *rankings* muy diferentes para el mismo conjunto de objetos a ordenar, ya sean países, ciudades o universidades, sin que sea posible saber de forma exacta de dónde proceden las diferencias: de los indicadores simples originales, de los pesos, de la normalización, de la agregación, de la imputación, del tratamiento de *outliers,...*

Son estas discrepancias las que, en gran parte, han motivado muchas veces el escepticismo del público, que se divide entre los partidarios entusiastas —además de los medios— que abogan por resumir realidades complejas multidimensionales en una sola dimensión, y profesionales —economistas y estadísticos en su mayoría— preocupados por el exceso de subjetividad en los diferentes pasos del proceso de elaboración de indicadores compuestos y sus *rankings* asociados.

Incluso académicos de reconocido prestigio, partidarios de una visión multidimensional de la medición económica, han manifestado su escepticismo respecto a lo que popularmente conocemos como indicadores compuestos. "On the whole, these composite indicators are better regarded as invitations to look more closely at the various components that underlie them." (Stiglitz, Sen y Fitoussi 2009, p. 65). Así pues, el mensaje parece ser que aunque los indicadores compuestos puedan ser útiles, no hay que perder de vista los indicadores individuales.

La presión por resumir y sintetizar la información en la era del *big data* es uno de los principales argumentos para la construcción de indicadores compuestos y *rankings* (Maggino 2017), que deberán ser elaborados de forma sensata, es decir siguiendo un cuidadoso proceso de control de calidad basado en consideraciones conceptuales y estadísticas, al tiempo que ser utilizados de forma responsable, por lo que las conclusiones y recomendaciones que de

5

¹ https://link.springer.com/journal/11205.

ellos se deriven deberán tener en cuenta el marco conceptual en el que los indicadores compuestos fueron desarrollados.

La creciente utilización de los indicadores compuestos en el ámbito de la monitorización de objetivos por parte de instituciones públicas, desde los Objetivos del Milenio de las Naciones Unidas hasta los indicadores de la Europa 2020, hizo que las autoridades europeas crearan, en el contexto del Joint Research Centre (JRC), una unidad específica de investigación en este tipo de indicadores (COIN2) que, junto con la Organización para la Cooperación y Desarrollo Económico (OCDE), ha desarrollado un manual de buenas prácticas para la elaboración de dichos indicadores (OCDE/Unión Europea/JRC 2008). Dicho manual no supone una homogenización completa de los métodos de producción, pero si representa un importante paso adelante al proponer una serie de pasos normalizados en la generación de indicadores compuestos. Pasos que pueden ser utilizados retrospectivamente para evaluar la calidad del indicador finalmente construido.

Este trabajo se centra en uno solo de los aspectos relacionados con la generación de *rankings*, el de la **agregación**. Es posible afirmar que lo que define un indicador compuesto son las propiedades subyacentes del método de agregación utilizado (Munda 2012). Por tanto la agregación es un elemento clave, aunque otros elementos puedan afectar sustancialmente al *ranking* final derivado del indicador compuesto.

La estructura del trabajo es la siguiente. El apartado 2 resume muy brevemente la guía propuesta por la OCDE (OCDE/Unión Europea/JRC 2008) y el grupo de Indicadores Compuestos del JRC (COIN) para la elaboración de indicadores compuestos. Aunque este trabajo solo aborda un aspecto relacionado con la elaboración de rankinas, resulta útil disponer de un marco general para ver donde encaja la agregación. Por otra parte, la agregación propuesta tiene efectos sobre otros aspectos, como la normalización o las alternativas al tratamiento de valores no disponibles (missing values). El siguiente apartado presenta un marco general al problema de la agregación radicalmente diferente al utilizado en la mayoría de aplicaciones prácticas, y lo hace en el contexto de lo que se conoce en matemáticas como un problema de orden lineal (Martí y Reinelt 2011). A continuación, se señalan los aspectos computacionales del problema, que para las

dimensiones habituales no son en absoluto despreciables y que han contribuido, sin duda, a que estos métodos no hayan ganado popularidad. El apartado 5 muestra las conexiones de nuestra aproximación a la agregación de rankings individuales con la teoría de la elección social. Es en este contexto en el que aparecen estos métodos en la literatura de los indicadores compuestos (Munda y Nardo 2009). Una vez clarificados los métodos, el apartado 6 ofrece una aplicación al ranking de Universidades conocido como U-Ranking (https://www.uranking.es/, Pérez y Aldás [dirs.] 2019), para lo cual se repasa detalladamente su metodología. Un apartado final ofrece las conclusiones y un largo apéndice aglutina los aspectos más técnicos que han sido suprimidos del texto principal.

² https://ec.europa.eu/jrc/en/coin.

2. GUÍA BREVE PARA LA ELABORACIÓN DE INDICADORES COMPUESTOS

Aunque sin proponer una metodología concreta, la guía de buenas prácticas de la OCDE (OCDE/Unión Europea/JRC 2008) establece 10 pasos a seguir para la elaboración de indicadores compuestos. Repasamos muy brevemente este esquema, que no debe entenderse de una forma unidireccional.

Paso 1: Marco teórico: Define el concepto a ser medido, los objetivos y los usuarios finales a los que va destinado el indicador.

Lo que está mal definido, probablemente se medirá mal. Como punto inicial es necesario comprender y disponer de una definición precisa de aquello que se quiere medir. Algunas decisiones posteriores dependen de esta definición. Como el objetivo es resumir una realidad multidimensional, suele ser útil estructurar el concepto a medir en diferentes dimensiones.

Paso 2: Selección de indicadores.

La calidad de los indicadores compuestos depende en gran medida de la calidad de los indicadores individuales. Idealmente, los indicadores deberían seleccionarse en función de su solidez analítica, capacidad de ser medidos, credibilidad, relevancia para el fenómeno que se está midiendo y relación entre ellos. Desde un punto de vista práctico la selección de indicadores está limitada fundamentalmente por la disponibilidad estadística.

Paso 3: Análisis y tratamiento de datos, si es necesario.

El primer paso en todo ejercicio cuantitativo es el análisis de datos y su tratamiento cuando sea necesario. Debemos, por tanto, visualizar los datos y calcular estadísticos descriptivos para cada uno de los indicadores. También la correlaciones entre ellos. Hay que comprobar la existencia de valores perdidos (*missing values*) y de valores atípicos (*outliers*). Esta tarea deberá tener en cuenta la naturaleza de la variable, cardinal *versus* ordinal, y también si deseamos utilizar los indicadores de forma extensiva, indicadores de tamaño, o intensiva, indicadores escalados. En este último caso habrá que decidir sobre la variable de escala.

Tres **aspectos**, referentes al proceso de datos, suelen ser **importantes** en este paso.

(i) El tratamiento de los **valores perdidos** (*missing values*) si existen. Algunos procedimientos estadísticos no aceptan valores perdidos por lo que en muchos casos es necesario proceder a su imputación. Una idea tan sugerente como peligrosa, ya que el efecto final de la imputación sobre el indicador compuesto no suele ser conocido.

(ii) El tratamiento de **valores atípicos** (outliers) si existen. Su origen debe ser investigado. En muchos casos los valores atípicos no tienen nada de atípico, simplemente un objeto destaca de forma extraordinaria sobre el resto, por arriba o por debajo. Incluso en estos casos es posible que sea necesario su tratamiento, ya que de otra forma una sola observación puede dominar el indicador final.

(iii) Posibles **transformaciones** de los indicadores originales. Más allá de su tratamiento como indicadores extensivos o intensivos habrá que decidir si hay que someter estos a alguna transformación. Por ejemplo, en el caso de la renta esta se suele considerar en muchos casos en términos logarítmicos. Las transformaciones suelen tener efectos colaterales que habrá que tratar, por ejemplo, la transformación logarítmica no admite valores nulos o negativos.

Paso 4: Normalización.

No podemos sumar peras y manzanas. Los indicadores individuales suelen estar en escalas diferentes, muchas de ellas no comparables. La construcción de un indicador compuesto requiere un criterio de normalización de forma que todos los indicadores individuales tengan la misma unidad de medida, lo que en ocasiones significa eliminar dichas unidades. La selección de un método de normalización a aplicar al problema en cuestión no es una decisión trivial y merece un cuidado especial. El método de normalización debe tener en cuenta las propiedades de los datos y los objetivos finales del indicador compuesto.

Finalmente todos los indicadores deben ser ajustados en la misma dirección, de forma que un valor mayor del indicador indique un mejor desempeño del concepto medido.

Paso 5: Ponderaciones de los indicadores y de sus dimensiones.

Lo que más importa debe pesar más, pero sobre cómo asignar estos pesos no existe ningún acuerdo, y la elección de los pesos en los indicadores individuales suele tener un importante componente subjetivo. Por otra parte, es sabido que las ponderaciones suelen tener un efecto importante, tanto en el valor final del indicador compuesto, como en el *ranking* derivado a partir de él. Sea cual sea el método se utilice, los pesos son esencialmente juicios de valor y tienen la propiedad de hacer explícitos los objetivos subyacentes a la construcción de un indicador compuesto.

Aunque el objetivo explícito de las ponderaciones es dar más peso a aquello que más importa, el método de agregación puede enmascarar este objetivo y convertir a los pesos en relaciones de intercambio entre los indicadores simples. Será necesario considerar también si las correlaciones entre los indicadores deben ser tenidas en cuenta en la determinación de los pesos.

Paso 6: Agregación.

Los indicadores individuales deben ser agregados en un único indicador final, al igual que las dimensiones si es que estas existen. Sin agregación no hay indicador compuesto, y por tanto tampoco *ranking* asociado.

La definición técnica de un indicador compuesto es una combinación (agregación) matemática de un conjunto de indicadores simples, que no tienen por qué tener una unidad de medida común, y para los que no existe un esquema obvio de ponderación. Pero de estos tres elementos, la agregación es el que define las propiedades básicas del indicador (Munda 2012).

El método de agregación debe respetar los objetivos del indicador compuesto y su *ranking*, en particular si debemos permitir la compensabilidad entre indicadores individuales y en qué grado, o por el contrario es mejor adoptar un método de agregación no compensatorio.

La agregación es pues esencial: sin agregación no hay indicador compuesto, aunque quizá no sea este el aspecto que mayor influencia tenga sobre el resultado final.

Paso 7: Coherencia conceptual y estadística.

El objetivo es explorar si las diferentes dimensiones del fenómeno de interés están representadas de forma equilibrada en el indicador compuesto. Si las dimensiones estadísticas no

coinciden con las dimensiones teóricas, entonces debería considerarse una revisión del conjunto de los indicadores individuales. Es necesario investigar las correlaciones entre el indicador compuesto y los indicadores individuales y examinar si se ha introducido algún sesgo. Técnicas de análisis multivariante pueden ser útiles en este cometido.

Paso 8: Análisis de sensibilidad y robustez: Control de calidad.

Por su propia naturaleza es difícil ofrecer medidas de incertidumbre estándar en la construcción de indicadores compuestos. Por ello, la única posibilidad es realizar un análisis de sensibilidad para comprobar la robustez del indicador a supuestos alternativos. Esta es la única posibilidad de despejar algunas de las controversias relacionadas con el indicador y el ranking que de él se desprende, al mismo tiempo que se aumenta la transparencia y se acotan las recomendaciones de política que se desprenden del indicador (Saisana, Saltelli y Tarantola 2005).

Paso 9: Relación con otros indicadores: Construyendo una historia consistente con los datos.

Es necesario examinar la correlación del indicador compuesto no solo con los indicadores individuales que han participado en su construcción, sino también con otras medidas relevantes para el fenómeno que se desea medir. iSi el propósito es buscar historias en los números, las herramientas estadísticas no conocen esas historias, esto es misión del investigador!

Paso 10: Visualización y comunicación.

Un buen gráfico vale más que mil palabras. Presentar un indicador compuesto no es un tema trivial. Los indicadores compuestos deben poder comunicar una imagen a los responsables de la toma de decisiones y a otros usuarios finales de forma rápida y precisa. Las actividades de difusión deben ir acompañadas de una descripción, lo más completa posible, de la metodología, así como de la información que ha servido de base para la elaboración del indicador.

3. AGREGACIÓN: UN ENFOQUE ALTERNATIVO

3.1 MARCO GENERAL

De todos los pasos mencionados en el esquema conceptual anterior este trabajo solo hace hincapié en uno de ellos, el de la agregación. Este paso es esencial en la construcción de indicadores compuestos y *rankings*, en el sentido de que no hay escapatoria a su aplicación, y por ello creemos que merece una especial atención. Es posible que haya otros cálculos que tengan mayor incidencia sobre el resultado final, pero el método de agregación define, en gran parte, las propiedades del indicador compuesto y su *ranking* asociado (Munda 2012).

Partimos del supuesto de que el objetivo final es construir un *ranking* de un conjunto de objetos, por ejemplo universidades, más que un indicador compuesto de su desempeño. Un *ranking* no es más que una ordenación (completa) de los objetos, universidades en nuestro ejemplo, de acuerdo con un determinado criterio. Supongamos que las universidades a ordenar son 4, (*A*, *B*, *C*, *D*), y que el criterio de ordenación es el número de *Publicaciones científicas* que producen dichas universidades, medidas con arreglo a un cierto criterio. La figura 1 muestra el *ranking* resultante en este caso, y enfatiza **dos cuestiones clave**.

La primera cuestión es que con un solo criterio el *ranking* viene determinado por la simple ordenación que el único indicador disponible proporciona: $\langle B, D, A, C \rangle$ en el ejemplo de la figura 1. La segunda cuestión es que en el proceso de construcción del *ranking* perdemos la información cardinal, la distancia que separa unas universidades de otras, ya que todo lo que importa es el orden de las universidades. Ciertamente esta es una pérdida de información relevante, y que puede llegar a ser inadmisible en determinados contextos. En cualquier caso, debe quedar claro que la construcción de un ranking no plantea ningún problema de agregación porque, con un solo indicador, no tenemos que agregar nada.

La situación es radicalmente diferente si queremos construir el *ranking* a partir de 3 indicadores distintos: *Publicaciones científicas, Profesores por cada 100 alumnos* e *Ingresos por contratos en miles de* \in , cada uno de ellos medido en una escala diferente. La figura 2 ilustra esta situación.

Figura 1. Elaboración de un ranking a partir de un solo indicador

¿Qué es un ranking?

Universidad	Publicaciones científicas	Universidad	Publicaciones científicas	Ranking
Α	3,581	 Α	3	В
В	8,691	В	1	D
С	2,587	С	4	Α
D	7,985	D	2	С

Información cardinal

Información ordinal

Con un solo indicador, la simple ordenación que dicho indicador proporciona.

Figura 2. Elaboración de un ranking a partir de múltiples indicadores

¿Y con más indicadores?

Universidad	Publicaciones científicas	Profesores 100 alumnos	Ingresos por contratos (€)	Información cardinal
A	3,581	8	6,897	
В	8,691	2	5,784	
С	2,587	4	9,152	,
D	7,985	6	2,005	

Información ordinal

Universidad	Publicaciones científicas	Profesores 100 alumnos	Ingresos por contratos (€)	
Α	3	1	2	
В	1	4	3	
С	4	3	1	
D	2	2	4	

¡Tenemos un problema de agregación!

Fuente: Elaboración propia.

De la información cardinal no se deriva automáticamente un *ranking*, puesto que no parece razonable promediar las publicaciones científicas con los ingresos por contratos, que además podríamos medirlos en €, en lugar de en miles de €. Tampoco parece razonable promediar estos indicadores con el número de profesores por cada 100 alumnos, de importancia cuantitativa mucho más reducida. En consecuencia, antes de proceder a elaborar el indicador compuesto que dé lugar al ranking que vamos buscando debemos poner todos los indicadores en una "unidad común", lo que en la mayoría de la ocasiones significa eliminarles a todos ellos las unidades de medida. Es decir, los indicadores deben ser normalizados antes de proceder a su agregación. Existen numerosos procedimientos para ello ampliamente utilizados en la literatura de indicadores compuestos, como la estandarización, la normalización por la mediana o la reducción de todos los indicadores al intervalo [0, 1] (OCDE/Unión Europea/JRC 2008). Es importante que esta normalización deje intacta la información ordinal, es decir, no altere el orden generado por cada uno de los indicadores individuales.

Una vez los indicadores han sido convertidos a una "escala común", la cardinalidad de cada uno de ellos debe ser reinterpretada, lo que no siempre es sencillo, pero en cualquier caso ya podemos proceder a la agregación en un indicador compuesto único, que es el que determinará el *ranking* final.

¿Ranking?

D

В

El procedimiento habitual aquí consiste en agregar mediante promedios, ya sean aritméticos, geométricos o utilizando medias generalizadas de orden- α (Goerlich y Villar 2009, apéndice A.1.1), que determinen el indicador compuesto, y a partir del mismo el *ranking* asociado. En el proceso de promediación es posible introducir pesos diferentes para cada indicador, que traten de mostrar la importancia de cada uno de ellos en el resultado final.

El principal inconveniente de construir indicadores compuestos mediante promedios es que permiten cierto grado de compensación entre los indicadores individuales, una vez han sido todos ellos reducidos a una escala común, para alcanzar un cierto valor del indicador compuesto. Por tanto siempre es posible compensar una actuación muy deficiente en un indicador con una actuación muy brillante en otro. El grado de compensación entre indicadores viene medido por los pesos asignados a los indicadores individuales que, en lugar de reflejar la importancia relativa de los mismos, muestra la relación de intercambio entre indicadores para alcanzar un valor dado del indicador compuesto (Munda y Nardo 2009).

Los promedios aritméticos permiten un grado de compensación a una tasa constante, mientras que los promedios geométricos permiten un grado de compensación que varía con el nivel de los indicadores, es decir cada vez se necesita un incremento mayor de un indicador para compensar una actuación cada vez más deficiente en otro indicador. Las medias geométricas, muy utilizadas en la práctica, no admiten valores nulos, y son muy sensibles a valores extremos. Un comportamiento similar exhiben las medias generalizadas de orden- α (Decancq y Lugo 2013).

En el ejemplo de las universidades de la figura 2, cada indicador puede considerarse que representa uno de los ámbitos de actuación de la Universidad: Investigación -Publicaciones científicas—, Docencia —Profesores por cada 100 alumnos— y Transferencia de Conocimiento —Ingresos por contratos—. La construcción de un indicador compuesto mediante un promedio de estos tres indicadores -una vez normalizados— implica asumir que podemos compensar una puntuación muy baja en Investigación con una puntuación muy elevada en Docencia o Transferencia de Conocimiento. Naturalmente si esta compensación es admisible o no dependerá, en gran parte, del problema que estemos tratando, pero hay numerosos contextos en los que esta compensación entre indicadores individuales no es, en absoluto, deseable. Así, por ejemplo, en los índices de calidad de vida es difícil justificar la compensación entre la renta per cápita y la calidad del aire, porque un deterioro excesivo del aire puede ser mortal para la salud, y en estas condiciones no es posible compensarlo con una mayor cuantía de renta, por muy grande que sea dicha cuantía (Munda y Saisana 2011; Munda 2015). Obsérvese, además, que el grado de compensación se mide a partir de los indicadores normalizados, lo que en muchas ocasiones dificulta esta interpretación de la compensabilidad (Decancq y Lugo 2013).

Así pues, el mensaje fundamental es que la agregación mediante promedios implica aceptar, en mayor o menor grado, la compensación entre indicadores individuales. Esta compensación se produce a una tasa constante en el caso de promedios aritméticos, que normalmente no se recomiendan por este motivo.

Frente a este tipo de agregación es posible construir *rankings*, pero no indicadores compuestos, si estamos dispuestos a abandonar la información cardinal, y quedarnos solo con la información ordinal que estos proporcionan. Se trata de métodos de agregación no compensatorios, en el sentido de que no se permite la compensación entre indicadores para obtener un determinado resultado (Munda y Nardo

2009; Munda 2012).³ Tal y como muestra la figura 2, cada uno de los indicadores originales constituye un *ranking* de universidades. El problema ahora consiste en agregar estos *rankings* individuales en un *ranking* de consenso con algún criterio razonable. Naturalmente podríamos promediar de nuevo estos *rankings* individuales y generar un *ranking* medio o mediano, pero esto no nos libraría del problema de la compensabilidad entre indicadores, y habríamos perdido, además, la información cardinal.

Ciertamente sacrificar la información cardinal en los indicadores originales supone una pérdida importante, ya que solo nos permite una ordenación final sin precisar ninguna idea de la distancia entre los objetos. Esta pérdida de información es, en muchos casos, ilusoria, puesto que la normalización, al igual que otros tratamientos previos de los datos - missing values, outliers,...-, afecta a la interpretación cardinal del indicador compuesto, de forma que la interpretación final de cuanto mejor es una universidad que otra es en muchos casos dudosa. No es posible discernir cuánto mejor en términos de qué, realmente. Por otra parte es posible que algunos indicadores sean ordinales por naturaleza.

En cualquier caso hay que ser conscientes de que los métodos de agregación propuestos a continuación se basan solo en información ordinal, y no es posible su aplicación si no estamos dispuestos a perder la información cardinal de los indicadores originales. Este sacrificio tiene sus compensaciones. Trabajar solo con información ordinal produce una normalización implícita, de la que por tanto ya no hay de qué preocuparse. Tampoco es necesario considerar el tratamiento de los outliers y, como veremos más adelante, podemos acomodar el caso en el que no se dispone de información sobre todos los objetos a ordenar, missing values, de forma que no es necesario proceder a su imputación. Finalmente deberemos tener en cuenta de que se trata de métodos no compensatorios.

Veamos la esencia de estos métodos a través de un ejemplo. Supongamos que disponemos de 21 indicadores individuales, cada uno con sus unidades de medida, para ordenar 4 universidades, tal y como se representan en la figura 3. Estos indicadores representan, todos ellos, 'cuanto más mejor', de forma que el mayor valor de cada indicador representa la primera universidad en el *ranking*. Cada indicador

-

³ Existe, sin embargo, otra literatura que si permite mantener la información cardinal y aplicar métodos de agregación no compensatorios, aunque no ha ganado mucha popularidad (Mazziota y Pareto 2016a, 2016b, 2018).

supone un *ranking* particular para nuestras universidades, lo que se representa en la figura 4. Como ya hemos indicado, el problema ahora es como agregar estos 21 *rankings* individuales en un *ranking* de consenso, es decir, en un *ranking* que en cierta forma minimice la distancia, como quiera que ésta sea medida, respecto a los *rankings* individuales.

Comencemos por contar quién domina a quién en cada *ranking* individual, y representar estas dominancias en formato matricial. Así, por ejemplo, el ranking derivado del indicador 1 implica la ordenación $\langle A, B, C, D \rangle$, de forma que A domina a B, C y D, B domina a C y D, y C domina a D. Estas dominancias pueden representarse en forma matricial —apéndice— tal y como aparecen en la figura 5, en la que las filas representan las universidades que van delante en el ranking respecto a las universidades que van detrás en el ranking, que se representan en las columnas. Lo mismo sucede, por ejemplo, con el indicador 3, cuyo ranking viene dado por $\langle B, D, C, A \rangle$, de forma que B domina a D, Cy A, D domina a Cy A, y C domina a A, y cuya representación matricial aparece al final de la figura 5.

Si representamos de esta forma cada uno de los *rankings* individuales y sumamos las 21 matrices resultantes, obtenemos la siguiente matriz —figura 6—, conocida, entre otros muchos nombres, como matriz de Condorcet o *outranking matrix*.⁴

Esta matriz es un adecuado resumen de nuestra información, aparece en la literatura con diferentes nombres, según el campo de conocimiento en el que estemos inmersos, y tiene una lectura directa: representa cuántas veces una universidad de las que aparecen en las filas domina a una universidad de las que aparecen en las columnas, teniendo en cuenta todos los *rankings* individuales de partida. Así, por ejemplo *A* va por delante de *C* en los *rankings* individuales 8 veces, y en consecuencia *C* va por delante de *A* en dichos *rankings* el resto de veces, es decir 13. O por ejemplo, B se sitúa por delante de D en todos los *rankings* individuales, es decir 21 veces.

Si llamamos a esta matriz ${\bf A}$, entonces resulta obvio que $a_{ij}+a_{ji}=21, \ \forall i\neq j, \ {\bf y} \ a_{ii}=0, \ \forall i$ Diferentes formas de manipular esta matriz implican diferentes procedimientos para generar un ranking de consenso determinado. Es decir, implementan diferentes criterios de agregación, de forma que nuestro problema ahora es cómo extraer un ranking de consenso, que cumpla ciertas propiedades razonables, a partir de la matriz de la figura 6.

Una forma intuitiva de proceder es la siguiente. Dado un *ranking* cualquiera, siempre es posible determinar, a partir de la matriz de Condorcet o *outranking matrix*, qué apoyo recibe dicho *ranking* de nuestra información de partida. Así, por ejemplo, consideremos aleatoriamente el *ranking* $\langle B, D, C, A \rangle$. Este *ranking* implica las siguientes dominancias entre cada par de universidades, BD, BC, BA, DC, DAyCA.

Cuántas veces una universidad está por delante de otra en comparaciones bilaterales podemos leerlo directamente de la matriz de Condorcet, $B \odot D = 21$, $B \odot C = 10$, $B \odot A = 13$, $D \odot C = 7$, $D \odot A = 13$ y $C \odot A = 13$. Por tanto, el apoyo de nuestra información a este *ranking* viene dado por la suma de estas dominancias, y es 77.

Una solución razonable para derivar un ranking de consenso —aunque no la única— consiste en calcular, para todos los rankings posibles, el apoyo que nuestra información otorga a cada uno de ellos, y elegir como ranking óptimo aquel para el que obtengamos el mayor apoyo.

Esta forma de proceder, originariamente debida a Kemeny (1959), descansa sobre el principio de la máxima verosimilitud, y en consecuencia tiene una base estadística evidente.

_

⁴ La función *outranking*, en *R*, calcula esta matriz a partir de *rankings*, o indicadores numéricos, individuales. El ejemplo numérico del texto procede de Moulin (1989), página 228.

Figura 3. Ejemplo: 21 indicadores cardinales y 4 Universidades a ordenar

Disponemos de 21 indicadores y 4 Universidades a ordenar

	Universidad A	Universidad B	Universidad C	Universidad D
Indicador 1	8,784	7,154	4,127	1,458
Indicador 2 (€)	10,475	5,478	8,457	2,458
Indicador 3	6	18	10	16
Indicador 4 (%)	0	60	100	40
i	i	i	i	i
Indicador 21	451	126	845	218

Fuente: Elaboración propia.

Figura 4. Ejemplo: 21 indicadores ordinales — rankings— y 4 universidades a ordenar

• Partimos de los rankings individuales:

	Universidad A	Universidad B	Universidad C	Universidad D
Indicador 1	1	2	3	4
Indicador 2	1	3	2	4
Indicador 3	4	1	3	2
Indicador 4	4	2	1	3
ı	ı	ı	ı	ı
Indicador 21	2	4	1	3

• ¡Solo información ordinal! ¡No necesitamos normalizar!

Figura 5. Representación matricial de cada ranking individual

• El indicador 1 nos dice que ABCD

	Universidad A	Universidad B	Universidad C	Universidad D
Indicador 1	1	2	3	4
	ı	ı	ı	1

• Una forma útil de representar esta información es

	Α	В	С	D
А	0	1	1	1
В	0	0	1	1
С	0	0	0	1
D	0	0	0	0

• El indicador 3 nos dice que BDCA

	Universidad A	Universidad B	Universidad C	Universidad D
I	I	I	I	i
Indicador 3	4	1	3	2
ŧ	ı	i	ı	I

• En la representación anterior

	Α	В	С	D
Α	0	0	0	0
В	1	0	1	1
С	1	0	0	0
D	1	0	1	0

Fuente: Elaboración propia.

Figura 6. Matriz de Condorcet o outranking

	А	В	С	D
Α	0	8	8	8
В	13	0	10	21
С	13	11	0	14
D	13	0	7	0

Si efectuáramos este cálculo en nuestro ejemplo llegaríamos a la conclusión de que el *ran-king* óptimo, o de consenso, es $\langle C, B, D, A \rangle$. Este es el *ranking* que recibe mayor apoyo de nuestra información, entendido como el que presenta un mayor número de dominancias en todas las comparaciones bilaterales posibles, como si se tratara de un torneo entre universidades. Puede comprobarse que el apoyo de nuestra información a este *ranking* alcanza un valor de 85.

Resulta interesante comprobar que la determinación del *ranking* óptimo equivale a encontrar la permutación de filas y columnas en la matriz de Condorcet tal que la suma de los elementos por encima de la diagonal principal sea la mayor posible. En este sentido se trata de un problema de triangularización de matrices (Martí y Reinelt 2011), y tiene evidentes conexiones con el problema de la determinación de las jerarquías en el análisis *input-output* (Chenery y Watanabe 1958).

Es fácil observar que, para el *ranking* óptimo, de la reordenación de filas y columnas de la matriz de Condorcet ofrecida en la figura 6 resulta la matriz ofrecida en la figura 7, y que, efectivamente, si sumamos los elementos por encima de la diagonal principal obtenemos un valor de 85.

Figura 7. Matriz de Condorcet ordenada

	С	В	D	Α
С	0	11	14	13
В	10	0	21	13
D	7	0	0	13
Α	8	8	8	0

Fuente: Elaboración propia.

Definido de esta forma, el problema de encontrar un ranking óptimo es lo que se conoce, en la literatura de la Investigación Operativa, como un problema de orden lineal (LOP), sobre el que existe una ingente bibliografía desde múltiples perspectivas, centrada en muchas ocasiones en los métodos para encontrar esta ordenación óptima, o un sustituto razonable de ella cuando la dimensión del problema así lo aconseje (Martí y Reinelt 2011). Por razones que se harán evidentes más adelante, desde el punto de vista de la teoría de la elección social, o la evaluación multicriterio, este concepto de solución se le conoce con el nombre de Condorcet-Kemeny-Young-Levenglick (Munda 2008).

Figura 8. Ejemplo: Incorporación de pesos en los indicadores individuales

• Si el indicador 3 nos dice que BDCA

	Universidad A	Universidad A Universidad B		Universidad D
I	I	i	I	i
Indicador 3	4	1	3	2
I	ı	ı	ı	i

Y tiene un peso de 30, su representación será

	Α	В	С	D
Α	0	0	0	0
В	30	0	30	30
С	30	0	0	0
D	30	0	30	0

Una vez definido el problema básico de agregación en el caso más sencillo posible, *rankings* individuales que tienen todos ellos la misma importancia —pesan lo mismo—, son fuertes — no hay empates o relaciones de indiferencia entre dos objetos— y completos —todos los objetos están ordenados—, debemos extender el principio general anterior a casos más realistas.

Como indican Emond y Mason (2002) la práctica habitual en estos casos está plagada de situaciones en las que a los *rankings* individuales se les otorga diferente importancia, y por tanto debemos acomodar pesos, se producen empates frecuentes, sobre todo si algunos de los indicadores originales son ordinales, y en muchas ocasiones los *rankings* individuales no incluyen a todos los individuos, bien porque los individuos no indican preferencias ordenadas por todos los objetos, como suele ser el caso en las votaciones, o bien porque podemos tener valores faltantes, *missing values*, en el caso de la construcción de indicadores compuestos.

3.2. ¿PODEMOS ATRIBUIR PESOS A LOS INDICADORES INDIVIDUALES?

Ciertamente es posible sin ninguna complicación adicional. Si, por ejemplo, al indicador 3, cuyo *ranking* viene dado por $\langle B, D, C, A \rangle$, le otorgamos un peso de 30, entonces su representación en formato matricial viene dada simplemente por la representación anterior — figura 5— multiplicada por 30 —figura 8—.

Asignando el peso correspondiente a cada indicador individual, y sumando la representación matricial correspondiente a todos ellos obtenemos la matriz de Condorcet, o *outranking matrix*, en términos ponderados. Para un *ranking* cualquiera determinamos el apoyo de nuestros datos a dicho *ranking* por el mismo procedimiento que antes, sumando las dominancias ponderadas que se derivan del mismo. El *ranking* óptimo será aquel que obtenga mayor apoyo por parte de la información disponible.

La introducción de pesos en este tipo de agregación solo debe tener en cuenta una **precaución**. En general, es esencial que ninguno de los indicadores individuales tenga asignado un peso superior al 50%, ya que de otra forma el *ranking* óptimo coincidirá con dicho indicador, que se convierte de esta forma en un 'dictador' en la terminología de Arrow (1963) (Munda y Nardo 2009).⁵

3.3. EMPATES (*TIES*) EN LOS INDI-CADORES INDIVIDUALES

En la práctica es frecuente que los indicadores individuales presenten el mismo valor para varios de los objetos a ordenar y, en consecuencia, incluso a este nivel inicial, no sea posible ordenar de forma completa todos los objetos. Así, por ejemplo, el indicador 8 muestra el mismo valor para las universidades C y D, de forma que no podemos decidir cuál de las 2 va delante de la otra —figura 9—. En términos del ranking generado, ambas universidades ocupan la misma posición —figura 9—.

La representación matricial habitual de este *ranking* —aunque no la única posible— es repartir la dominancia entre ambos objetos. De esta forma la representación matricial del indicador 8 vendría dada por la matriz ofrecida en la figura 10. Esta es una forma conveniente de proceder cuando parte de la información de partida es de naturaleza ordinal.

⁵ Aunque este es el criterio general, el apéndice ofrece un contra-ejemplo, es decir un caso en el que el peso de un indicador individual es de ²/₃, pero este no resulta ser el indicador óptimo.

Figura 9. Ejemplo: Incorporación de empates (ties) en los indicadores individuales

Si el indicador 8 es

	Universidad A	Universidad B	Universidad C	Universidad D
I	I	ı	I	ı
Indicador 8	3,214	8,526	5,560	5,560
i	ı	ı	ı	ı

• En términos de rankings, B D ~ C A

	Universidad A Universidad B		Universidad C	Universidad D
I	I	I	i	i
Indicador 8	4	1	2	2
ı	ı	ı	ı	I

Fuente: Elaboración propia.

Figura 10. Representación matricial de un ranking individual con empates

	Α	В	С	D
Α	0	0	0	0
В	1	0	1	1
С	1	0	0	0,5
D	1	0	0,5	0

Fuente: Elaboración propia.

Una vez acomodados los empates de esta forma es posible construir la matriz de Condorcet y evaluar el apoyo de cualquier *ranking*, en la forma usual de contar dominancias en comparaciones bilaterales.

Si llamamos **A** a la matriz de Condorcet, entonces también se cumple en este caso que $a_{ij}+a_{ji}=21, \ \forall i\neq j, \ \ y \ a_{ii}=0, \ \forall i$, si no consideramos pesos.

3.4. ¿QUÉ HACER CUANDO NO HAY INFORMACIÓN? (MISSING VALUES)

Con frecuencia encontramos situaciones en las que no disponemos de información en los *rankings* individuales acerca de todos los objetos a ordenar. En el caso de elecciones o concursos no se suele votar a todos los candidatos, y parece natural pensar que estos van por detrás de todos aquellos candidatos que si han sido votados. Esta es la aproximación más utilizada por la literatura (Ailon 2010).

Sin embargo, en el caso de indicadores compuestos, que es nuestro centro de atención, esta no parece una opción razonable ya que no hay motivo para pensar que si no disponemos de la observación de un indicador para una determinada universidad ésta debe ponerse a la cola en dicho indicador. Parece mucho más natural asumir que no disponemos de información en ese caso concreto, y efectuar la comparación sólo entre los objetos que son comparables.

Por ejemplo, en el indicador 6 de nuestro ejemplo no se dispone de la información para la universidad B—figura 11—, de forma que para ese indicador individual el *ranking* solo compara el resto de universidades para las que disponemos de información, $\langle C, D, A \rangle$.

La representación matricial de este *ranking* añade ceros cuando encontramos dos objetos no comparables, es decir, cuando debemos comparar cualquiera de las universidades con la universidad *B*. De esta forma la fila y la columna de dicha matriz estará constituida por ceros —figura 12—.

El motivo para no quitar esta fila y columna es que para otros indicadores si disponemos de información sobre la universidad *B*, y en consecuencia la necesitamos en el proceso de agregación de los *rankings* individuales para obtener la matriz de Condorcet. Una vez disponemos de esta matriz evaluamos el apoyo de cualquier *ranking* en la forma usual de contar dominancias en comparaciones bilaterales.

Figura 11. Ejemplo: Incorporación de valores faltantes en los indicadores individuales.

Si el indicador 6 es

	Universidad A Universidad B Universidad C		Universidad C	Universidad D
ı	i	I	i	i
Indicador 6	3,104	NA	7,620	5,530
I	ı	I	I	Ī

Solo podemos comparar A, C y D. El ranking es C D A

	Universidad A	Universidad B	versidad B Universidad C Universi	
ı	ı	ı	ı	i
Indicador 6	3	NA	1	2
•	ſ	ı	ı	I

Fuente: Elaboración propia.

Figura 12. Representación matricial de un ranking individual con valores no disponibles

	A	В	С	D
Α	0	0	0	0
В	0	0	0	0
C	1	0	0	1
D	1	0	0	0

Fuente: Elaboración propia.

3.5. VARIABLES CONTINUAS Y UMBRALES EN EL ESTABLECI-MIENTO DE ORDENACIONES INDIVIDUALES

Todo lo que necesitamos para construir la matriz de Condorcet, o *outranking matrix*, — figura 6— es información ordinal, de la que extraemos directamente *rankings* individuales. Cuando las variables son continuas y ofrecen información cardinal, simplemente obtenemos el *ranking* individual que de ellas se derivan. Sin embargo, esto puede dar lugar a situaciones que podrían parecer injustas o, más generalmente, no deseables.

Veamos un ejemplo. Supongamos que en el indicador *Publicaciones Científicas* la universidad A ofrece un registro de 8,345 y la universidad B de 8,344, solo una publicación menos. Por su parte las universidades C y D ofrecen registros de 7,845 y de 2,356 publicaciones respectivamente. Aunque la universidad C

está claramente por delante de la universidad D, y las universidades A y B están claramente por delante de las universidades C y D, no está tan claro que la universidad A deba quedar situada por delante de la universidad B. Entre otras muchas razones porque nuestro instrumental de medida puede no ser tan preciso.

Para solucionar estas situaciones podemos recurrir a la discretización de las variables continuas, o establecer umbrales de forma que solo se establezca una relación de preferencia si la diferencia, en valor absoluto, entre dos indicadores es mayor de un determinado umbral. En caso contrario, tenemos un empate o relación de indiferencia (Munda 2008, capítulo 4). En este contexto el tratamiento de los empates (*ties*) es esencial.

3.6. EMPATES (*TIES*) EN EL ESPA-CIO DE SOLUCIONES

Una reflexión final en relación a los empates (ties) es necesaria. Hasta ahora hemos considerado la posibilidad de existencia de relaciones de indiferencia entre los objetos a ordenar en los rankings individuales, pero no en el ranking óptimo, o de consenso. Es decir, no hemos tenido en cuenta la posibilidad de rankings con empates en el espacio de soluciones. Esta es una situación común en la literatura (Athanassoglou 2015).

La determinación del *ranking* óptimo como un problema de triangularización de matrices, es

decir, como la búsqueda de la permutación de filas y columnas en la matriz de Condorcet tal que la suma de los elementos por encima de la diagonal principal sea la mayor posible, no admite un ranking con empates como solución, aunque puede dar lugar a soluciones múltiples que podrían ser interpretadas como tales. Sin embargo, esta es una situación que se da en la práctica con mucha frecuencia; por ejemplo U-Ranking ordena 62 universidades, pero el ranking final solo considera 11 grupos de universidades en su edición de 2019 (Pérez y Aldás [dirs.] 2019). Existen además argumentos que indican que es razonable que en el espacio de soluciones se incluyan rankings con empates (ties) apéndice—. Ello es posible solo a partir de los rankings en términos matriciales y a costa de aumentar notablemente la complejidad en la búsqueda de soluciones (Emond y Mason 2000, 2002).

4. LA FORMULACIÓN MATEMÁTICA DEL PROBLEMA

El apartado anterior ha introducido de forma intuitiva el problema que tratamos de resolver y cómo acomodar las situaciones más comunes con las que nos encontramos en la práctica a la hora de elaborar indicadores compuestos: pesos, empates (*ties*) en los indicadores individuales y valores faltantes (*missing values*).

En este apartado formalizamos el problema matemático que debemos resolver y comentamos ciertos aspectos computacionales para obtener el *ranking* óptimo o de consenso. Los detalles más técnicos se exponen en el apéndice.

Disponemos de *m rankings* individuales con la finalidad de ordenar *n* objetos o alternativas. Estos *rankings* individuales pueden ser débiles —presentar empates (*ties*)—, parciales —tener valores faltantes (*missing values*)— y llevar asociado un peso. El punto de partida que resume nuestra información en términos de las dominancias o fortalezas entre objetos viene dado por la *outranking matrix*,

$$\mathbf{A} = \sum_{k=1}^{m} \mathbf{\omega}_k \mathbf{A}^k \tag{1}$$

donde A^k es la representación matricial del ranking individual k, de orden $n \times n$, y ω_k es el peso asociado, un escalar, que puede estar normalizado o no. Los elementos de esta matriz para el ranking individual k, $\left\{a_{ii}^k\right\}$, son

$$a_{ij}^{k} = \begin{cases} 1 & \text{si el objeto } i \text{ precede al objeto } j \\ 0.5 & \text{si el objeto } i \text{ es equivalente al objeto } j \\ 0 & \text{si el objeto } i \text{ sigue al objeto } j \\ \text{o } i \text{ y } j \text{ no son comparables} \end{cases}$$

El objetivo es encontrar un *ranking* que reciba el mayor apoyo de nuestra información en todas las comparaciones bilaterales resumidas en la *outranking matrix* (1).

Consideremos un *ranking*, *r*, cuya representación matricial viene dada por la matriz *S*. En el

espacio de *rankings* completos los elementos de S, $\{s_{ij}\}$, vendrían dados solamente por elementos del conjunto $\{0, 1\}$, mientras que si ampliamos dicho espacio para incluir *rankings* débiles los elementos de S, $\{s_{ij}\}$, vendrían dados por elementos del conjunto $\{0, 0.5, 1\}$. El apoyo de nuestra información a este *ranking* viene dado por el producto escalar de las matrices S Y A,

$$S \cdot A = S \cdot (\sum_{k=1}^{m} \omega_k A^k) =$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij} \sum_{k=1}^{m} \omega_k a_{ij}^k$$
(3)

Si el espacio de *rankings* considerado es representado por el conjunto Ω , deseamos encontrar un *ranking*, representado por *S*, que maximice la expresión anterior,

$$\underset{S \in \Omega}{Max} \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij} c_{ij}$$
 (4)

donde $c_{ij} = \sum_{k=1}^{m} \omega_k a_{ij}^k$ son los elementos de la *outranking matrix*.

La ecuación (4) es básicamente la ecuación (5) en Emond y Mason (2002, p. 23), si bien nuestra definición de *outranking matrix* no coincide exactamente con la de estos autores en presencia de empates (*ties*) —apéndice—.

La *outranking matrix* se computa fácilmente sumando, elemento por elemento, la representación matricial de cada *ranking* individual convenientemente ponderado. La **solución** a nuestro problema es el *ranking* que, cuando es multiplicado escalarmente por la *outranking matrix*, proporciona el mayor valor.

4.1. EL PROBLEMA COMPUTACIONAL

El problema de optimización definido en (4) es sencillo de implementar en la práctica. La **enumeración exhaustiva** de todos los casos posibles ofrece un método de solución obvio. La *outranking matrix* solo se calcula una vez a partir de los datos de nuestro problema. Por lo que definido el espacio de *rankings* donde bus-

⁶ Estos *m* rankings individuales proceden de *m* indicadores, cardinales u ordinales, —figuras 3 y 4—.

car la solución óptima, solo necesitamos calcular el producto escalar entre cada uno de estos *rankings* y la *outranking matrix*, (3), y anotar el *ranking* para el que se obtiene el mayor valor posible, (4).

Desafortunadamente esta sencillez desaparece cuando consideramos la tasa a la que crecen los posibles rankings. Considerando como espacio de soluciones los rankings completos, los casos posibles son n!. Ello significa que con 4 objetos a ordenar, las universidades de nuestro ejemplo del apartado anterior, los casos posibles son 24, pero con 10 objetos a ordenar estos son ya 3,628,800. Con 24 objetos a ordenar los rankings posibles son del orden de 6.2×10^{23} , y con 62 objetos a ordenar, las universidades que hay en U-Ranking 2019 (Pérez y Aldás [dirs.] 2019), este número crece hasta 3.1×10^{85} (Goerlich 2019). iCientos de miles de billones de rankings posibles!

La situación es todavía mucho peor si consideramos como espacio de soluciones la posibilidad de incluir *rankings* débiles, es decir con empates (*ties*). En este escenario los casos posibles crecen todavía mucho más rápidamente (Gross 1962). Con 4 objetos a ordenar, las universidades de nuestro ejemplo del apartado anterior, los casos posibles son 75, 24 de estos casos son *rankings* completos, 36 son *rankings* con empates en los que existen 3 grupos, 14 son *rankings* con empates en los que existen 2 grupos y finalmente hay 1 *ranking* en el que todos los objetos están empatados (*all-tied solution*). Con 10 objetos a ordenar estos son

⁷ Aunque la mayoría de *rankings* no consideran esta opción, esto no es siempre así. Por ejemplo, U-Ranking reduce las 62 universidades a solo 11 grupos, y por tanto la solución adoptada incluye numerosos grupos de indiferencia entre un conjunto importante de universidades. Emond y Mason (2000, p. 4) ofrecen argumentos a favor de considerar *rankings* débiles en el espacio de soluciones.

⁸ Si solo consideramos rankings completos entonces los rankings posibles son el número de permutaciones de n elementos sin repetición, que viene dado por factorial de n, n! Si consideramos, además, rankings débiles habrá repeticiones al formar las permutaciones de *n* elementos. El número de permutaciones de *n* elementos con repetición viene dado por n^n . Sin embargo, en términos de rankings, habrá muchos rankings repetidos en el conjunto de todas esas permutaciones, por lo que en este caso el número de rankings posibles será mayor de n! y menor de n^n . Es posible determinar dicho número por medio de una ecuación en diferencias (Gross 1962; Emond 1997; Emond y Mason 2000). La función weak.rankings, en R, calcula dicho número, pudiendo determinar el número de rankings con un determinado número de grupos. Enumerar los rankings posibles en este caso no es trivial (Emond 1997, Anexo A).

ya 102,247,563, de los que solo 3,628,800 son *rankings* completos. Con 24 objetos a ordenar los *rankings* posibles son del orden de 3.0×10^{27} , y con 62 objetos a ordenar, las universidades que hay en U-Ranking 2019 (Pérez y Aldás [dirs.] 2019), este número crece hasta 1.7×10^{95} (Goerlich 2019).

La **enumeración exhaustiva** no es, pues, una alternativa factible cuando el número de objetos a ordenar excede de 10 o 12, dado el estado actual de la computación. De hecho es conocido que este problema es *NP-hard* (Bartholdi, Tovey y Trick 1989), por lo que la literatura ha desarrollado, desde diferentes perspectivas, una serie de algoritmos, exactos o aproximados, para encontrar el *ranking* óptimo o de consenso en el sentido definido por (4) (Emond y Mason 2002; Martí y Reinelt 2011; Ali y Meilă 2012; Athanassoglou 2015; Azzini y Munda 2020).

Obsérvese que esta complejidad computacional no depende del número de rankings individuales, m, que pueden ser muy numerosos. La outranking matrix, $A = \sum_{k=1}^m \omega_k A^k$, se computa solo una vez y permanece fija en el proceso de optimización (4). La complejidad de cálculo complejidad de complejidad de complejidad de complejidad complejidad

Desde el punto de vista de la aplicación realizada en este trabajo, para resolver el problema de optimización (4) se ha implementado una modificación del algoritmo de *branch-and-bound* (rama y cota) desarrollado en Emond y Mason (2000, 2002) y que proporciona soluciones exactas. La modificación es necesaria puesto que el concepto de solución propuesto por Emond y Mason (2000, 2002) descansa sobre una métrica algo diferente de la captada por la *outranking matrix*, tal y como ha sido definida en la sección 3, y no es equivalente a la nuestra en el caso de la presencia de empates (*ties*) —apéndice—.

La aproximación de Emond y Mason (2000, 2002) se encuentra implementada en la librería *ConsRank* de R (D'Ambrosio, Amodio y Mazzeo 2019), que incluye además dos algoritmos heurísticos, *Quick* y *Fast* (D'Ambrosio, Amodio e Iorio 2015; Amodio, D'Ambrosio y Siciliano 2016), que proporcionan excelentes aproximaciones cuando no es posible obtener la solución exacta dada la dimensión del problema, como sucede en este caso con 62 universidades a ordenar. A partir de esta librería se introdujeron las modificaciones necesarias en los algoritmos, exactos y aproximados, para reformular el

problema de optimización en términos de la *outranking matrix*, (4) —apéndice—.

El algoritmo finalmente utilizado fue *Fast* (Amodio, D'Ambrosio y Siciliano 2016), con un número suficiente de iteraciones para asegurar la consecución de un óptimo con una elevada probabilidad.

4.2. SOLUCIONES MÚLTIPLES

No hay ninguna seguridad de que el *ranking* óptimo derivado de la resolución del problema de optimización sea único. Si hay poco consenso en los *rankings* individuales o muchos valores faltantes (*missing values*), es posible que el número de soluciones sea relativamente grande. En ocasiones puede ser admisible aceptar múltiples soluciones, pero normalmente no lo será, y deberemos adoptar un criterio para generar una solución única a partir de todas las soluciones óptimas encontradas. Es necesario reconocer que esto puede ser un problema práctico que haya que solucionar.

Puede derivarse un *ranking* medio o mediano a partir del conjunto de soluciones, puede incluso obtenerse un *ranking* 'optimo' del conjunto de soluciones óptimas. Pero es necesario reconocer que, cualquiera que sea la solución adoptada, el *ranking* resultante no será necesariamente óptimo, en el sentido de que no recibirá el mayor apoyo de nuestra información en todas las comparaciones bilaterales resumidas en la *outranking matrix* (1), aunque dicho apoyo no debería alejarse mucho del óptimo.

El analista debe ser consciente de esta posibilidad, la existencia de múltiples *rankings* óptimos, así como de la necesidad de arbitrar una solución práctica al respecto.

En la aplicación a U-Rankings se produjeron *rankings* múltiples en muy pocas ocasiones, y en ningún caso fueron muy numerosos. Cuando se dio esta situación se tomó como solución el *ranking* medio de los *rankings* óptimos. Obsérvese que el *ranking* medio no tiene por qué ser óptimo, aunque debe estar muy cerca del resultado óptimo.

5. ¿PODEMOS APRENDER ALGO DE LA TEORÍA DE LA ELECCIÓN SOCIAL?

La respuesta a esta pregunta es claramente **sí**. No solo de la teoría de la elección social, sino también de la teoría de la evaluación multicriterio (Munda 2008), ya que el trabajo de Arrow y Raynaud (1986) mostró formalmente la equivalencia entre ambos enfoques. En el contexto de la construcción de los indicadores compuestos, los métodos expuestos en los dos apartados anteriores son introducidos precisamente a partir de la teoría de la elección social y la evaluación multicriterio (OCDE/Unión Europea/JRC 2008), donde las alternativas son los objetos a ordenar y los indicadores individuales son los criterios de evaluación.

La teoría de la elección social se ocupa básicamente de la agregación de preferencias individuales en un resultado social óptimo en un cierto sentido. Este es, precisamente, el problema estudiado por Kemeny (1959), es decir un problema de agregación. Lógicamente, esta cuestión aparece de forma natural en el contexto de las votaciones, ya que estas son una forma de obtener las preferencias individuales con el objeto declarado de obtener un resultado social.

En el contexto de la elección social hay dos grandes líneas de pensamiento que se originan en los trabajos pioneros de Borda (1781) y Condorcet (1785). Vansnick (1990) mostró que las dos principales aproximaciones a la agregación en la teoría de la evaluación multicriterio, métodos compensatorios y no compensatorios, pueden derivarse directamente de las propuestas de estos dos autores.

La solución de Borda es sencilla de implementar en la práctica. Dados n objetos o alternativas a ordenar, para cada ranking individual computamos sus puntuaciones de la siguiente forma. El objeto en última posición recibe una puntuación de 0, el objeto en penúltima posición recibe una puntuación de 1, y así sucesivamente. El proceso continua hasta que el objeto en primera posición recibe una puntuación de $n\!-\!1$. Sumando estas puntuaciones para todos los rankings individuales obtenemos el ranking global. Esta forma de proceder se conoce generalmente como $scoring\ rules$.

Las puntuaciones adoptadas son arbitrarias y otro esquema de ponderaciones es perfectamente admisible. Si estas puntuaciones son una transformación lineal del scoring original de Borda, entonces ello no tiene efectos sobre el ranking global finalmente derivado. Naturalmente, transformaciones no lineales pueden alterar el resultado final de forma sustancial. Una generalización del procedimiento de Borda consiste en utilizar una secuencia nodecreciente de números reales, $p_0 \le p_1 \le p_2 \le ... \le p_{n-1}$, para ponderar los objetos de acuerdo con su ranking en cada uno de los indicadores individuales.

El scoring original de Borda tiene, sin embargo, una interpretación natural. Para cada ranking individual el objeto en cuestión recibe la puntuación que corresponde al número de objetos a los que precede. Así, el objeto en última posición no precede a nadie, y recibe 0 puntos, el objeto en penúltima posición precede a un solo objeto, y recibe 1 punto, el objeto en primera posición precede al resto de los n-1 objetos, y por tanto recibe n-1 puntos. Esta observación conecta el scoring original de Borda con la outranking matrix de un ranking individual. Sumando por filas la outranking matrix de un ranking individual obtenemos el scoring de Borda —apéndice—.9

Así pues la solución de Borda descansa sobre la intensidad de las preferencias, lo que se manifiesta en las ponderaciones asignadas. Ello implica que este criterio de agregación es compensatorio, en el sentido de permitir la compensabilidad entre los indicadores individuales para alcanzar un resultado global (Munda 2008, sec.-6.2) y no ha ganado popularidad en la construcción de indicadores compuestos.

La solución de Condorcet se basa en las comparaciones bilaterales introducidas en el apartado 3, y que conducen a la matriz de Condorcet o *outranking matrix*. Lo propuesta original de Condorcet para derivar de esta matriz un

⁹ Esta conexión entre el *scoring* de Borda y la matriz de Condorcet ha sido utilizada por algunos autores para proponer un criterio de agregación que combine ambos enfoques (Herrero 2016; Herrero y Villar 2019).

ranking consiste en quedarse solo con los ganadores e imponer la transitividad. Utilizando como ejemplo las 4 universidades del epígrafe 3.1, que conduce a la *outranking matrix* de la figura 6, quedarse con los ganadores implica quedarnos con las puntuaciones superiores a 10, es decir, la matriz de la figura 13.

Figura 13. Ganadores en la matriz de Condorcet o outranking

	Α	В	С	D
Α				
В	13			21
С	13	11		14
D	13			

Fuente: Elaboración propia.

Si imponemos transitividad sobre esta matriz, resulta obvio que el *ranking* viene dado por $\langle C, B, D, A \rangle$.

Sin embargo, este es un caso particular en el que la solución de Condorcet, obtenida por transitividad, existe. En general, es mucho más común que no exista una solución de este tipo, sino que se presenten ciclos que impidan la obtención de un *ranking* que cumpla con la propiedad de transitividad.

Por ejemplo, supongamos que tenemos 3 universidades a ordenar y 60 indicadores individuales. A partir de ellos construimos la *outranking matrix* representada en la figura 14 (Munda 2008, p. 117).

Quedarse con los ganadores da lugar a la matriz de la figura 15.

A partir de ella observamos que $A \\cappa B$, $B \\cappa C \\cappa A$, lo que viola el principio de transitividad y genera un ciclo que impide la obtención de un *ranking* de consenso. Es lo que se conoce, en el contexto de la elección social, como los **ciclos de Condorcet**.

Es conocido que la existencia de ciclos es muy frecuente en la práctica, y que su probabilidad crece con el número de objetos a ordenar y de *rankings* individuales (Fishburn, Gehrlein y Maskin 1979). En este contexto, derivar un *ranking* consistente con Condorcet deberá renunciar a la propiedad de transitividad en presencia de ciclos. ¹⁰

Figura 14. Matriz de Condorcet o outranking

	A	В	С
Α	0	33	25
В	27	0	42
С	35	18	0

Fuente: Elaboración propia.

Figura 15. Ganadores en la matriz de Condorcet o outranking

	Α	В	С
Α		33	
В			42
С	35		

Fuente: Elaboración propia.

La aproximación más natural es la expuesta en las secciones 3 y 4, que no se concentra solo en los ganadores sino que acaba seleccionando como *ranking* óptimo o de consenso aquel que recibe el mayor apoyo en todas las comparaciones bilaterales. Esta es la aproximación seguida por Kemeny (1959) y axiomatizada por Young y Levenglick (1978), cuya popularidad se ha visto mermada en la práctica por los problemas computacionales mencionados anteriormente.

Sin embargo, este no es el único método consistente con Condorcet que es posible utilizar para derivar un ranking a partir de la outranking matrix. La literatura de la elección social ha desarrollado múltiples reglas, algunas de ellas prácticas y muy rápidas, para seleccionar un ranking óptimo o cercano al óptimo, sin necesidad de resolver el problema de optimización (4).11 Una de las más rápidas y simples es la conocida como regla de Copeland (Moulin 1989), que determina la puntuación de una alternativa comparándola con el resto en la outranking matrix, asignando + 1 frente las alternativas a las que gana, - 1 frente las alternativas contra las que pierde, y sumando el resultado para obtener su valoración final. La puntuación de todas las alternativas, calculada de esta forma determina el ranking óptimo.

Otro procedimiento para generar un *ranking* a partir de comparaciones bilaterales es el de Arrow y Raynaud (1986), que resuelve los ciclos de Condorcet de forma que se minimiza el

¹⁰ Entendemos por una regla consistente con Condorcet un método para derivar un ranking a partir de la *outranking matrix* tal que, si no existen ciclos, proporcione la solución de Condorcet derivada mediante transitividad.

¹¹ Un gran conjunto de estas reglas puede verse en https://www.condorcet.vote/.

máximo desacuerdo entre todas las comparaciones bilaterales posibles. Y lo hace a partir de un procedimiento secuencial en el que va identificando iterativamente los objetos en la posición más baja del *ranking*. En la práctica el conjunto de *rankings* óptimos seleccionados por el procedimiento de Arrow y Raynaud (1986) puede ser muy grande (Munda 2008, Sec.-6.4).

En realidad, el enfoque de la elección social, seguido mayoritariamente por economistas y teóricos de la política de las votaciones, es una aproximación axiomática. Se formulan las propiedades que una función de agregación debe cumplir y, a partir de aquí, se caracterizan las diversas funciones agregación en términos de dichos axiomas (Arrow 1963; Fishburn 1977). Por otra parte, el enfoque matemático o de la teoría de la computación es lo que podríamos llamar una aproximación 'métrica'. Comienza con una métrica subyacente a los rankings individuales -cuanto difieren entre ellos- y define un problema de agregación — optimización— con el objeto de encontrar un ranking óptimo o de consenso en un cierto sentido (Fagin et al. 2006). Básicamente se trata el mismo problema desde dos perspectivas diferentes.

6. AU-Ranking: U-Ranking alternativo

6.1. METODOLOGÍA DE U-RANKING

U-Ranking, en su edición de 2019, analiza 62 universidades españolas (48 públicas y 14 privadas) a partir de un conjunto total de 25 indicadores¹² (Pérez y Aldás [dirs.] 2019, cuadro 3). Las universidades incluidas se indican en el cuadro 1 y los indicadores, su definición, código y peso utilizado en la agregación en el cuadro 2.

Los rasgos básicos de la metodología están bien descritos en Pérez y Aldás (dirs.) (2019, capítulo 2). Este epígrafe simplemente detalla los pasos seguidos en la construcción del *ranking* a partir del esquema conceptual expuesto en el apartado 2, e incluye muchos aspectos de cálculo que son necesarios para entender las diferencias entre la metodología original y la implementada en este trabajo.

U-Ranking es, en realidad, un sistema de *rankings*, ya que además del denominado *ranking* de rendimiento incluye un *ranking* de volumen, que tiene en cuenta el tamaño de las universidades y un sistema *web* que permite obtener al usuario *rankings* personalizados por titulaciones. El objeto de este trabajo es el *ranking* principal, que es el conocido como *ranking* de **rendimiento**, y que ajusta las universidades por su tamaño (Pérez y Aldás [dirs.] 2019).

U-Ranking tiene **dos fortalezas** importantes. En **primer lugar** cubre las **3 dimensiones** principales de las universidades: *Docencia, Investigación* e *Innovación y Desarrollo Tecnológico*. Construye un indicador para cada una de estas dimensiones y, por agregación, un indicador global. La evaluación de cada una de estas 3 dimensiones considera, a su vez, 4 **ámbitos de actividad**: Acceso a *recursos, Producción* obtenida, *Calidad* e *Internacionalización* de las actividades, para cada uno de los cuales disponemos de entre 1 y 3 indicadores.¹³

Así pues, los 25 indicadores de partida para las 62 universidades se agrupan en 4 ámbitos para cada una de las 3 dimensiones (Pérez y Aldás [dirs.] 2019, p. 21).

Partiendo de estos indicadores, a los que como veremos a continuación ya se les ha efectuado previamente el correspondiente tratamiento de datos v han sido normalizados, se procede a la construcción del ranking mediante un proceso de agregación en 3 niveles. Primero, los indicadores de cada ámbito, para cada dimensión, se agregan mediante una media aritmética simple, ya que existen algunos indicadores con valores nulos. Segundo, los ámbitos de cada dimensión se agregan mediante medias geométricas ponderadas. Los pesos, para cada ámbito y dimensión, proceden de consultas a expertos universitarios y pueden verse en Pérez y Aldás (dirs.) (2019, cuadro 4).14 Esto determina el indicador compuesto, y su ranking correspondiente, para cada una de las dimensiones consideradas: Docencia, Investigación e Innovación y Desarrollo Tecnológico. Tercero, las dimensiones se agregan mediante medias geométricas ponderadas con pesos del 56% para Docencia, 34% para Investigación y 10% para Innovación y Desarrollo Tecnológico. De nuevo los pesos son determinados mediante consultas a expertos. Este último paso determina el indicador compuesto que, redondeado a 1 decimal, se interpreta como un índice de rendimiento de las universidades. A partir de aquí se obtiene el *ranking* correspondiente que agrupa las 62 universidades en 11 grupos (Pérez y Aldás [dirs.] 2019, cuadro 6). Se trata pues de un *ranking* con un importante volumen de empates (ties): por eiemplo, el grupo más numeroso es el situado en torno a la media, que agrupa 18 universidades.

gico, de forma que para esta dimensión no existe dicho ámbito (Pérez y Aldás [dirs.] 2019, cuadro 3).

¹² El panel inicial que aspira a cubrir U-Ranking incluye 31 indicadores, pero por razones de disponibilidad estadística solo se consideran 25 de dichos indicadores en esta edición (Pérez y Aldás [dirs.] 2019, cuadro 3).

¹³ La disponibilidad de información hace que no exista información para el único indicador potencial de *Calidad* dentro de la dimensión *Innovación y Desarrollo Tecnoló-*

¹⁴ El peso asignado al ámbito de *Calidad* dentro de la dimensión *Innovación y Desarrollo Tecnológico* se reparte proporcionalmente entre el resto de ámbitos.

Cuadro 1. Universidades incluidas en U-Ranking 2019 Orden Código Universidad Tipo Universidad Pontificia Comillas COMILLAS Privada Universidad de Alicante 2 UA Publica 3 UAB Universitat Autònoma de Barcelona Publica UAH Universidad de Alcalá Publica Universidad de Almería 5 UAL Publica UAM Universidad Autónoma de Madrid Publica UANE Universidad Nebrija Privada UB Universitat de Barcelona 8 Publica UBU Universidad de Burgos Publica 9 **UC3M** Universidad Carlos III de Madrid Publica 10 11 UCA Universidad de Cádiz Publica UCEU Universidad San Pablo-CEU 12 Privada 13 UCH Universidad Cardenal Herrera-CEU Privada Universidad de Castilla-La Mancha 14 UCLM Publica 15 UCM Universidad Complutense de Madrid Publica UCO Universidad de Córdoba 16 Publica 17 UCV Universidad Católica de Valencia San Vicente Mártir Privada Universidade da Coruña UDC Publica 18 19 Universidad de Deusto Privada 20 UDG Universitat de Girona Publica Universidad a Distancia de Madrid **UDIMA** Privada 21 Universitat de Lleida Publica 22 UDL 23 UFV Universidad Francisco de Vitoria Privada Universidad de Granada UGR Publica 24 25 UHU Universidad de Huelva Publica 26 UIB Universitat de les Illes Balears Publica UIC Universitat Internacional de Catalunya 27 Privada 28 UJAEN Universidad de Jaén Publica 29 UJI Universitat Jaume I de Castellón Publica Universidad de La Laguna 30 ULL Publica 31 ULPGC Universidad de Las Palmas de Gran Canaria Publica 32 UM Universidad de Murcia Publica UMA Universidad de Málaga 33 Publica UMH Universidad Miquel Hernández de Elche Publica 34 35 **UMON** Mondragón Unibertsitatea Privada 36 UN Universidad de Navarra Privada Universidad Nacional de Educación a Distancia 37 UNED Publica 38 UNEX Universidad de Extremadura Publica UNICAN Universidad de Cantabria Publica 39 40 UNILEON Universidad de León Publica 41 UNIOVI Universidad de Oviedo Publica UNIRIOJA Universidad de La Rioja 42 Publica UNIZAR Universidad de Zaragoza Publica 43 UOC Universitat Oberta de Catalunya Privada 45 UPC Universitat Politècnica de Catalunya Publica 46 UPCT Universidad Politécnica de Cartagena Publica 47 UPF Universitat Pompeu Fabra Publica 48 **UPM** Universidad Politécnica de Madrid Publica 49 UPNA Universidad Pública de Navarra Publica 50 UPO Universidad Pablo de Olavide Publica UPV Universitat Politècnica de València 51 Publica UPV-EHU Universidad del País Vasco/Euskal Herriko Unibertsitatea 52 Publica Universidad Rey Juan Carlos 53 URJC Publica 54 URLL Universitat Ramon Llull Privada 55 URV Universitat Rovira i Virgili Publica Universidad de Sevilla 56 US Publica 57 USAL Universidad de Salamanca Publica Universidad de Santiago de Compostela 58 USC Publica 59 UV Universitat de València Publica 60 UVA Universidad de Valladolid Publica UVIC Vic-Universidad Central de Catalunya Privada 61 62 UVIGO Universidade de Vigo Publica

Fuente: Pérez y Aldás (dirs.) (2019).

				Indicador			Peso	
Dimensión	Ámbito		Código	Definición	Desagregación	Dimensión	Ámbito	Indicador
		1	DR1c	Profesor por cada 100 alumnos	Rama de enseñanza			4,7%
	Recursos	2	DR2c	Presupuesto / Alumno	Universidad		25,4%	4,7%
		3	DR3b	Profesor Doctor / Profesores	Universidad			4,7%
<u>.</u>		4	DP1d	Tasa de Éxito	Rama de enseñanza			5,7%
Docencia	Producción	5	DP2d	Tasa de Evaluación	Rama de enseñanza	F.C.O.	30,4%	5,7%
90		6	DP3d	Tasa de Abandono	Rama de enseñanza	56%		5,7%
Ď	Calidad	7	DC2d	% de estudiantes de posgrado	Rama de enseñanza		22 00/	6,7%
	Callada	8	DC3R	Notas de corte	Rama de enseñanza		23,9%	6,7%
	Internacionalización	9	DI1d	% de alumnos extranjeros	Rama de enseñanza		20.20/	5,7%
	internacionalizacion	10	DI2b	% de alumnos en programas de intercambio	Rama de enseñanza		20,3%	5,7%
		11	IR1a	Recursos públicos competitivos por profesor doctor	Rama de enseñanza			3,4%
	Recursos	12	IR2a	Contratos de personal doctor, becas de investigación y apoyo técnico sobre el presupuesto	Rama de enseñanza	anza	20,0%	3,4%
'n.		13	IP1a	Documentos citables con referencia ISI por profesor doctor	Rama de enseñanza			3,4%
Investigación	Producción	14	IP2b	Sexenios totales sobre sexenios posibles	Rama de enseñanza		30,0%	3,4%
Ę.		15	IP3b	Tesis doctorales leídas por cada 100 profesores doctores	Rama de enseñanza	34%		3,4%
is si		16	IC1ar	Factor medio de impacto	Rama de enseñanza			3,4%
Ĕ	Calidad	17	IC2ar	% de publicaciones en el primer cuartil	Rama de enseñanza		30,0%	3,4%
		18	IC3ar	Citas por documento	Rama de enseñanza			3,4%
	Internacionalización	19	II2g	Fondos de investigación internacionales por profesor doctor	Universidad		20.00/	3,4%
	internacionalizacion	20	II3ar	% de publicaciones en coautorías internacionales	Rama de enseñanza		20,0%	3,4%
≗		21	TR1b	Ingresos por licencias por cada 100 profesores doctores	Universidad			1,4%
Desarrollo igico	Recursos	22	TR2d	Ingresos por contratos de asesoramiento por cada 100 profesores doctores	Universidad		43,3%	1,4%
S S		23	TR3a	Ingresos por formación continua por profesor doctor	Universidad			1,4%
og. igi	Producción	24	TP1	Número de patentes por cada 100 profesores doctores	Universidad		33,3%	3,3%
Innovación y Deso Tecnológico	Internacionalización	25	TI1b	Patentes triádicas por cada 100 profesores doctores	Universidad	10%	23,3%	2,3%

Nota: El peso a nivel de ámbito para la dimensión de Innovación y Desarrollo Tecnológico incluye la redistribución del peso del ámbito de Calidad, para el que no existen indicadores en esta dimensión (Pérez y Aldás, 2019, pág.-24, Cuadro 4). Los pesos a nivel de indicador son pesos derivados para el indicador global, de forma que suman 1 a este nivel de agregación, y no son utilizados directamente en U-Ranking.

Fuente: elaboración propia a partir de Pérez y Aldás (dirs.) (2019), (p.24 cuadro 4).

La segunda fortaleza importante de U-Ranking es que toma en cuenta la especialización de las universidades en las diferentes ramas de conocimiento: Artes y Humanidades, Ciencias Sociales y Jurídicas, Ciencias, Ingeniería y Arquitectura y Ciencias de la Salud. Por esta razón, los indicadores de partida que integran el proceso de agregación que acabamos de describir están construidos a partir de indicadores iniciales a nivel de Rama de Conocimiento para cada universidad, siempre que esto sea posible, lo que sucede en 17 de los 25 indicadores utilizados. Los 8 indicadores restantes solo están disponibles a nivel de universidad, pero su valor se replica para todas las ramas de la universidad al objeto de poder realizar el análisis de datos para cada indicador a nivel de rama y universidad. Es en esta fase previa donde se realiza todo el tratamiento de datos y la normalización.

Puesto que el indicador final es un índice de rendimiento, los indicadores iniciales, a nivel de rama, son ratios o porcentajes construidos a partir de estadísticas de diferentes años, y promediados hasta un máximo de 6 años con intención de dotar de cierta estabilidad al indicador final. Una vez disponemos de estos indicadores iniciales tenemos de una tabla de 62 universidades \times 5 ramas \times 25 indicadores = 7,750 celdas, algunas de ellas con valores no disponibles y sobre la que se efectúa el tratamiento de datos que se describe a continuación. Dicho tratamiento es a nivel de indicador y rama, 125 en total, para cada universidad, suponiendo que todas las universidades tengan actividad en todas las ramas.

El orden en el que se realiza el **tratamiento de datos** es importante, puesto que el resultado final no es conmutativo.

En **primer** lugar, se pasa un **filtro** para eliminar algunas ramas de ciertas universidades cuya importancia resulta insignificante. En *Docencia* una rama no se considera si el porcentaje de alumnos de grado y máster en dicha rama es inferior al 3% del total de su universidad y si dichos alumnos son menos de 500. En *Investigación* e *Innovación y Desarrollo Tecnológico* una rama no se considera si el número de profesores doctores en dicha rama es cero. ¹⁵ El resultado de aplicar

¹⁵ Tanto los alumnos como los profesores deben considerarse equivalentes a tiempo completo (ETC).

dicho filtro se muestra en el cuadro 3, que muestra las ramas eliminadas en las Universidades en la que los filtros mencionados son de aplicación.

En **segundo** lugar se realiza el tratamiento de *outliers*. Este se realiza a partir de los datos anteriores ya filtrados, y por tanto con un conjunto de datos todavía incompleto. Se considera como *outlier* toda aquella observación que se desvía 3 desviaciones típicas, σ , o más de su media, μ , tanto por encima o como por debajo. Cuando este es el caso la observación se trunca a $\mu + 3 \times \sigma$ o $\mu - 3 \times \sigma$ según sea el caso.

En **tercer** lugar se estiman los **valores faltantes** (*missing values*). Los valores faltantes se estiman mediante en un proceso en dos etapas con la restricción de que solo se consideran si finalmente la universidad no mejora tras el proceso de imputación. El método seguido procede de la siguiente forma

(*i*) Con los datos efectivamente disponibles se calcula la matriz de correlaciones entre indicadores para cada rama, y distinguiendo entre las dimensiones de *Docencia*, por una parte, e *Investigación* e *Innovación* y *Desarrollo Tecnológico*, por otra. De esta forma disponemos de 10 matrices de correlación con dimensión igual al número de indicadores disponibles en las dimensiones consideradas.¹⁶

(ii)Para imputar el valor de un indicador de una universidad en una rama concreta, dentro de las dimensiones consideradas, se seleccionan, de la matriz de correlaciones correspondiente, los dos indicadores con mayor correlación —en valor absoluto— con el indicador a imputar, y se realiza una regresión lineal por mínimos cuadrados ordinarios. Y a partir de los indicadores de la universidad se imputa el indicador correspondiente. Es posible que, para dicha universidad, no existan valores de los indicadores que actúan como

29

 $^{^{16}}$ Es decir, para *Docencia* disponemos de 5 matrices de correlación, una para cada rama, de dimensión 10×10 , que son los indicadores disponibles en esta dimensión, mientras que para *Investigación* e *Innovación y Desarrollo Tecnológico* disponemos de 5 matrices de correlación, una para cada rama, de dimensión 15×15 , que son los indicadores disponibles en el conjunto de estas dos dimensiones (Pérez y Aldás [dirs.] 2019, cuadro 3).

Cuadro 3. Ramas no consideradas en determinadas Universidades. Docencia Investigación e Innovación y Desarrollo Tecnológico Tipo Universidad Total **Total** 7 Privada COMILLAS Privada UANE 2 Pública UC3M 3 Privada UCEU Privada UCH 2 Privada UCV Privada UDE Privada UDIMA 2 Privada UFV Privada UIC 2 Pública UII 1 Pública ULPGC Privada UMON Pública UNIRIOJA Privada UOC 1 Pública UPC 4 Pública UPCT 3 Pública UPF 1 Pública UPM 3 Pública UPNA 2 Pública UPO Pública UPV Privada URLL Pública UVIGO Total 1 18 2 12

Rama 1: Artes y Humanidades. Nota:

Rama 2: Ciencias Sociales y Jurídicas.

Rama 3: Ciencias.

Rama 4: Ingeniería y Arquitectura.

Rama 5: Ciencias de la Salud.

Fuente: elaboración propia a partir de la base de datos de U-Ranking.

predictores, en cuyo caso se mantiene el valor faltante.17

(iii) El último paso es comprobar que la imputación no mejora a la universidad en cuestión. Lo incluimos aquí, aunque se ejecuta después del proceso de normalización que comentamos a continuación.

¹⁷ Esto hace que la matriz de 25 indicadores a nivel de universidad pueda no estar efectivamente completa, y los valores faltantes no se hayan rellenado totalmente. Una alternativa para que esto no suceda sería coger, de la matriz de correlaciones correspondiente, los dos indicadores con mayor correlación para los que si se dispone de información de predictores de la universidad en cuestión.

En cuarto lugar se procede a la normalización. La normalización se efectúa por la mediana de cada indicador a nivel de rama. Esto centra los indicadores en uno.

Por las razones explicadas anteriormente el proceso de normalización se efectúa dos veces, antes y después de imputar los valores faltantes, de forma que los valores imputados solo se mantienen si la universidad no mejora su posición. Una vez decidido qué valores imputados se mantienen y cuáles no se procede de nuevo al proceso de normalización final con el mismo criterio.

En quinto y último lugar, se procede a la obtención de los 25 indicadores a nivel de universidad mediante agregación de las ramas. Estos indicadores están ya normalizados, sin outliers y con los valores faltantes imputados.

Para la dimensión de Docencia los indicadores se agregan mediante el porcentaje de alumnos de grado, máster y doctorado de la universidad en cada rama. En Investigación e Innovación y Desarrollo Tecnológico la agregación utiliza el porcentaje de profesores doctores de la universidad. En ambos casos para los últimos 6 años.

En consecuencia el proceso de elaboración de U-Ranking incluye 4 niveles de agregación y no 3, como parece desprenderse de Pérez y Aldás (dirs.) (2019, p. 21), ya que existe un primer nivel de agregación, mediante medias aritméticas ponderadas a nivel de rama, para obtener los indicadores a nivel de universidad.

Ello permite finalmente obtener la matriz de 62 universidades × 25 indicadores que es la que se utiliza en el proceso de agregación en los 3 niveles descritos al principio de este apartado. A nivel de ámbito mediante medias aritméticas simples, y a nivel de dimensión y global mediante medias geométricas ponderadas.

6.2. U-RANKING COMO UN PRO-**BLEMA DE ORDEN LINEAL: AU-RANKING**

El objeto de nuestro ejercicio es ver el efecto que sobre los rankings de dimensión y el ranking global de rendimiento de las universidades tiene sustituir los criterios de agregación utilizados por U-Ranking, promedios aritméticos y geométricos, por otros basados en la matriz de Condorcet, o *outranking matrix*; manteniendo sin embargo el resto de supuestos intactos, fundamentalmente los pesos, y siendo lo más fieles posible a los datos originales. Naturalmente, como ya hemos visto en el apartado 4, los métodos de agregación de rankings individuales en un ranking óptimo o de consenso alteran sustancialmente el tratamiento de datos, puesto que suponen una normalización implícita, evitan el tratamiento de outliers y permiten obviar la imputación de valores faltantes. El ejercicio, fundamentalmente de carácter metodológico, toma como referencia los datos de la edición de U-Ranking de 2019 (Pérez y Aldás [dirs.] 2019).

Por tanto, los resultados de nuestro ejercicio, en relación al ranking derivado en U-Ranking, se deben no solo al criterio de agregación, sino también al resto de cuestiones que afectan al tratamiento de datos. Separar qué parte de los cambios observados se debe solo al criterio de agregación, y qué parte a los diferentes pasos en el tratamiento de datos exigiría un análisis de sensibilidad que está fuera del alcance del presente trabajo. 18

Adaptar el esquema metodológico descrito en el apartado anterior a nuestro propósito no es totalmente directo. El proceso seguido, y puesto que el objetivo sigue siendo construir rankings de universidades teniendo en cuenta su heterogeneidad a nivel de Rama de Conocimiento, cuando ello sea posible, es el siguiente.

El punto de partida son los indicadores iniciales, ratios o porcentajes construidos a partir de estadísticas, ya promediados hasta un máximo de 6 años. Estos indicadores originales pueden tener *outliers*, valores faltantes y cada uno tiene unas unidades de medida (Pérez y Aldás [dirs.] 2019, cuadro 3). De los 25 indicadores iniciales, 8 están disponibles a nivel de universidad y 17 a nivel de Rama de Conocimiento. Disponemos pues de una tabla inicial de 62 universidades × (8 indicadores a nivel de universidad + 5 ramas × 17 indicadores a nivel de rama) = 5,766 celdas. Sobre esta tabla **no** efectuamos ningún tratamiento de datos.

A partir de esta tabla procedemos de la siguiente forma. Para los 17 indicadores a nivel de rama obtenemos el correspondiente indicador de universidad básicamente mediante el mismo procedimiento que en el último paso seguido en el tratamiento de datos de U-Ranking, un promedio ponderado. La diferencia básica es que la agregación se realiza para el indicador de partida, sin tratamiento de datos, ya que dentro de cada rama el indicador tiene las mismas unidades de medida.

Para la dimensión de *Docencia* los pesos son el porcentaje de alumnos de grado, máster y doctorado de cada rama de la universidad. En Investigación e Innovación y Desarrollo Tecno*lógico* los pesos son el porcentaje de profesores doctores de cada rama de la universidad. En ambos casos se utiliza la media de los últimos 6 años. Esto proporciona 17 indicadores a nivel de universidad que, junto con los 8 indicadores iniciales a nivel de universidad, completan los 25 indicadores de partida para la generación de los rankings por dimensiones y global. No se ha

¹⁸ Los resultados presentados en Goerlich (2019) recogen solo el efecto del cambio en el criterio de agregación. Para ello se parte de la matriz 62 universidades × 25 indicadores una vez ha sido realizado el tratamiento de datos descrito en el apartado anterior. Este ejercicio es más general, ya que partimos de los indicadores originales, y englobamos en él otros efectos relacionados con el tratamiento de datos.

efectuado ningún tratamiento de datos, se trata de los indicadores originales con propiedades cardinales.

Todos los indicadores tienen la misma dirección, 'cuanto más mejor', ¹⁹ por lo que los 25 indicadores de partida se utilizan para construir los *rankings* individuales. Perdemos en este punto la información cardinal y nos centramos solo en la información ordinal. Disponemos pues de una matriz de 25 *rankings* individuales × 62 universidades. Dichos *rankings* pueden ser fuertes o débiles, completos o incompletos, pero son los que constituyen nuestra información de partida. ²⁰ Examinar los *rankings* individuales, a nivel de ámbito y dimensión, es un ejercicio interesante que será ofrecido en el epígrafe siguiente.

A partir de estos rankings individuales realizamos 2 procesos de agregación, sustituyendo los promedios por una agregación no compensatoria basada en la matriz de Condorcet, al nivel que estemos trabajando, y aplicando el concepto de solución de Condorcet-Kemeny-Young-Levenglick (CKYL). Primero, se agregan los rankings individuales de cada dimensión, utilizando los mismos pesos que los empleados por U-Ranking (Pérez y Aldás [dirs.] 2019, cuadro 4).21 Esto determina los rankings de universidades para cada una de las dimensiones consideradas: Docencia, Investigación e Innovación y Desarrollo Tecnológico. Segundo, puesto que el peso asignado a la dimensión Docencia es superior al 50%, 56%, la obtención del ranking global vendría determinado por el propio ranking de docencia. La solución adoptada es generar el ranking global de rendimiento de las universidades a partir de los rankings individuales, con los pesos implícitos derivados de U-Ranking —cuadro 2—. De esta forma procedemos directamente de los rankings individuales a los rankings de dimensión, en el paso anterior, pero también al ranking global de rendimiento

en este último proceso de agregación. Los pesos aplicados se ofrecen en el cuadro 2.²²

En el caso de soluciones múltiples se adopta el criterio expuesto en el epígrafe 4.2, el espacio de soluciones considera siempre *rankings* con empates (*ties*) y no solo *ranking* completos, y el algoritmo utilizado es *Fast* de Amodio, D'Ambrosio y Siciliano (2016) con 2000 iteraciones, convenientemente modificado para adaptarlo a la *outranking matrix* —apéndice—

Comparamos a continuación los resultados con este método con los que se obtienen de U-Ranking en su edición de 2019 (Pérez y Aldás [dirs.] 2019).

6.3. ANÁLISIS EXPLORATORIO DE DATOS

Como se ha expuesto en el epígrafe anterior el método de agregación utilizado nos permite obviar las cuestiones del tratamiento de *outliers* y los valores faltantes, al tiempo que efectúa una normalización implícita a través de la consideración sólo de la información ordinal contenida en los indicadores de partida. El precio a pagar por simplificar el tratamiento de datos es que perdemos la cardinalidad de un indicador sintético.

Por estas razones, antes de comparar los resultados de ambos procedimientos, conviene realizar un pequeño análisis de la información de partida. Para construir la matriz 25×62 de rankings iniciales procedemos de la siguiente forma.

Para los indicadores disponibles a nivel de *Rama de Conocimiento*, 17, partimos de una matriz de $85^{24} \times 62 = 5,270$ celdas. En dicha matriz falta información de 337 valores, un 6.4%. No todos estos valores no disponibles son faltantes puesto que, tal y como muestra el cuadro 3, algunas universidades no muestran actividad en ciertas ramas de conocimiento, y son eliminadas en U-Ranking.

Estos indicadores iniciales fueron agregados a nivel de universidad utilizando como pesos el total de alumnos equivalentes a tiempo com-

 $^{^{19}}$ Para ello la *Tasa de Abandono* (DP3d) se transforma en *Tasa de Permanencia*, es decir, 1 – tasa de abandono.

No se ha considerado la introducción de umbrales en los indicadores de partida que son variables continuas en la generación de las ordenaciones individuales, al objeto de alterar lo menos posible los datos de partida de U-Ranking, pero esta sería una cuestión a considerar en la implementación de esta metodología de agregación

²¹ Al igual que en U-Ranking el peso asignado al ámbito de *Calidad* dentro de la dimensión *Innovación y Desarrollo Tecnológico* se reparte proporcionalmente entre el resto de ámbitos de dicha dimensión.

²² Una solución alternativa hubiera sido agregar los *rankings* individuales a nivel de ámbito, dentro de cada dimensión, y construir los diversos rankings a partir de los *rankings* de ámbito.

²³ Con 62 objetos a ordenar el algoritmo exacto de *Branch-and-Bound* de Emond y Mason (2002) resultó impracticable.

 $^{^{24}}$ 85 = 17 indicadores \times 5 *Ramas de Conocimiento*.

pleto de grado, máster y doctorado por Rama de Conocimiento en el caso de la dimensión de Docencia y los profesores doctores a tiempo completo por Rama de Conocimiento en el caso de la dimensión de Investigación. Todos los indicadores de la dimensión de Innovación y Desarrollo Tecnológico lo están a nivel de universidad, y no necesitan de agregación. El proceso de agregación toma en cuenta los valores disponibles de cada indicador y universidad. Así, por ejemplo, si para un determinado indicador y universidad solo disponemos de información para 3 Ramas de Conocimiento, entonces el indicador a nivel de universidad se obtiene como una media ponderada para esas 3 Ramas de Conocimiento, re-escalando los pesos a la información disponible.

Para los indicadores disponibles a nivel de Universidad, 8, simplemente los tomamos tal cual. Disponemos en este caso de una matriz de $8 \times 62 = 496$ celdas en la que existen 9 valores faltantes.

Juntando los indicadores agregados a nivel de *Rama de Conocimiento* con los indicadores disponibles a nivel de universidad tenemos nuestra matriz de partida de indicadores elementales constituida por $25 \times 62 = 1,550$ celdas, en la que tenemos un total de 30 valores faltantes, un 1.9%, que trataremos como casos de no comparabilidad.

Estos indicadores de partida son todos en la dirección de 'cuanto más mejor', ²⁵ de forma que el mayor valor corresponde a la primera universidad en el *ranking* y el menor valor corresponde a la última universidad en el *ranking*. Los *rankings* individuales, generados a partir de estos indicadores, pueden observarse en el cuadro 4. Son estos *rankings*, junto con los pesos correspondientes a cada indicador — cuadro 2—, los que constituyen nuestro problema de agregación. Resulta interesante observar que algunas universidades alcanzan las primeras posiciones en determinados indicadores y las últimas en otros, incluso dentro de una misma dimensión.

Utilizando los *rankings* individuales del cuadro 4 y los pesos a nivel de indicador del cuadro 2 construimos la *outranking matrix*. Puesto que estos pesos están normalizados, y suman la unidad para todos los indicadores, se cumple que $c_{ij} + c_{ji} \leq 1, \ \forall i,j$, con igualdad si no hay valores faltantes en las universidades compara-

²⁵ Para ello la *Tasa de Abandono* (DP3d) se transforma en *Tasa de Permanencia*, es decir, 1 – tasa de abandono.

das y desigualdad cuando los hay. La cota superior de esta matriz, (A16), es 1,235.63. Contando el número dominancias -número de celdas por filas que son mayores o iguales a 0.5— es fácil hacerse una idea de las universidades que estarán en el grupo de cabeza y en el grupo de cola. Las 12 universidades con mayor y menor número de dominancias, así como el número de las mismas se muestran en el cuadro 5 que permite hacernos una idea de lo que debemos esperar en los extremos del ranking. La Universidad Pompeu Fabra (UPF) domina al resto, y por tanto aparecerá la primera en el ranking. En el otro extremo, la Universidad a Distancia de Madrid (UDIMA) y la Universidad Nacional de Educación a Distancia (UNED) son dominadas por todas, y por tanto aparecerán las últimas en el ranking. 26

Solo una universidad privada aparece entre las de mayor número de dominancias, la *Universidad de Navarra* (UN), mientras que el grupo de cola aparece poblado por este tipo de universidades, pues 7 de 12 son privadas. Además, en este grupo encontramos a las 3 universidades no presenciales: la *Universidad a Distancia de Madrid* (UDIMA), la *Universidad Nacional de Educación a Distancia* (UNED) y la *Universitat Oberta de Catalunya* (UOC).

restauradas a posteriori.

33

²⁶ Este hecho podría utilizarse para reducir la dimensión del problema de optimización, (4), puesto que, dada la estructura de la *outranking matrix*, estas universidades podrían excluirse del problema de optimización, y ser

	áhita	Indicador							Ur	nive	rsida	1						
Dimensiór	Ámbito	Código	COMILLAS	UA	UAB	UAH	UAL	UAM	UANE	UB	UBU	исзм	UCA	UCEU	UCH	UCLM	υсм	UCO
		DR1c	57	49	43	30	51	15	19	37	14	21	27	5	52	23	33	47
Docencia	Recursos	DR2c	10	37	22	29	42	28	16	32	47	12	44	19	17	48	33	34
		DR3b	51	41	11	24	7	8	58	22	52	18	50	47	60	45	9	13
		DP1d	2	47	20	29	32	12	1	17	44	19	48	35	16	40	26	43
		DP2d	15	21	17	29	52	10	55	23	20	22	39	7	16	36	31	46
		DP3d	6	35	30	37	50	18	22	34	16	10	51	45	38	19	25	39
	Calidad	DC2d	10	53	32	16	29	24	2	22	57	14	50	3	23	59	33	30
		DC3R	62	12	19	17	33	5	50	8	47	1	31	50	49	41	6	24
	Internacionalización	DI1d	20	22	18	12	30	23	1	11	56	17	57	41	4	53	16	49
		DI2b	2	48	47	14	10	18	38	43	16	7	41	5	59	56	34	12
	Recursos	IR1a	58	42	8	24	31	6	50	9	49	7	34	55	62	13	25	21
		IR2a	56	50	13	22	24	5	1	11	38	8	21	53	NA	40	34	17
	Producción	IP1a	58	41	2	36	35	4	60	5	44	10	52	54	59	31	23	19
'É		IP2b	NA	16	3	32	26	17	NA	38	41	2	34	NA	NA	21	18	1
gació		IP3b	45	36	2	20	56	4	13	9	49	42	41	5	1	55	7	24
Investigación	Calidad	IC1ar	58	53	10	12	54	6	61	1	25	55	44	13	4	41	24	14
드		IC2ar	55	53	11	24	49	5	62	10	17	23	50	48	46	39	31	12
		IC3ar	41	49	10	15	51	7	61	8	30	57	52	31	22	44	38	27
	Internacionalización	II2g	45	32	13	35	22	11	59	37	29	3	30	52	NA	41	51	16
	Internacionalizacion	113ar	59	47	5	20	43	9	56	8	19	6	35	55	50	51	18	28
- L		TR1b	51	23	16	44	24	22	51	7	14	6	36	26	51	33	43	20
Desa Ógico	Recursos	TR2d	58	35	8	7	41	17	38	13	49	10	42	51	NA	45	32	19
ón y crool		TR3a	11	23	10	22	37	24	3	17	60	6	62	40	58	29	26	39
ᅙᆖ	Producción	TP1	55	7	50	14	19	23	51	46	2	4	8	52	56	34	35	28
	Internacionalización	TI1b	52	3	12	23	47	27	52	18	34	5	15	46	NA	45	33	25

Fuente: Elaboración propia a partir de los datos de Pérez y Aldás (dirs.) (2019). En rojo la primera universidad en el ranking, en azul la última universidad en el ranking para un indicador dado. NA significa valor faltante.

	Ámbito	Indicador							U	nivers	idad							
Dimensión		Código	ucv	UDC	UDE	UDG	UDIMA	UDL	UFV	UGR	UHU	UIB	UIC	UJAEN	UJI	ULL	ULPGC	UM
	Recursos	DR1c	55	9	56	42	60	22	1	36	35	29	32	46	48	24	28	50
		DR2c	58	53	11	26	61	20	8	43	46	54	5	51	36	52	45	57
		DR3b	55	46	27	34	54	39	62	4	33	37	56	19	35	14	30	5
Docen	Producción	DP1d	13	57	7	27	3	23	15	37	52	42	5	51	34	33	54	25
		DP2d	13	56	2	9	59	19	8	41	54	38	3	42	43	58	50	51
		DP3d	4	28	2	36	61	17	8	41	56	58	3	43	46	57	54	26
	Calidad	DC2d	9	38	7	43	1	27	4	40	41	28	34	37	21	61	62	42
		DC3R	50	25	50	39	50	35	50	16	46	45	50	44	21	30	43	10
	Internacionalización	DI1d	43	61	29	7	8	21	25	35	51	26	3	52	19	31	37	39
		DI2b	53	33	1	49	61	24	58	22	13	52	54	36	40	57	19	46
	Recursos	IR1a	60	33	56	20	61	2	11	29	44	12	57	45	37	46	47	35
		IR2a	NA	30	36	27	NA	29	NA	28	41	9	NA	48	25	37	54	33
	Producción	IP1a	56	49	57	15	61	8	48	21	47	7	30	34	24	26	53	37
,Ę		IP2b	NA	36	NA	12	NA	NA	NA	7	45	15	NA	23	9	43	46	28
Investigación		IP3b	12	48	11	44	62	17	50	29	52	53	6	59	34	51	43	21
nvesti	Calidad	IC1ar	47	51	60	28	62	22	19	29	48	20	30	42	38	7	37	43
=		IC2ar	59	42	60	35	61	8	52	32	43	14	36	38	37	3	44	51
		IC3ar	59	37	39	25	62	19	60	14	46	13	32	28	26	6	40	55
	Internacionalización	II2g	53	19	5	12	56	21	NA	34	58	28	57	55	40	42	38	17
		II3ar	61	38	36	22	62	44	60	11	26	10	57	42	29	4	40	46
rollo		TR1b	51	27	51	28	51	19	51	40	50	12	51	39	35	51	51	25
Jesar gico		TR2d	58	23	NA	28	NA	33	57	48	53	54	20	43	36	55	50	46
ación y Desa Tecnológico		TR3a	7	54	8	61	14	32	20	52	46	59	9	47	49	57	53	36
ě	Producción	TP1	56	20	56	49	56	42	56	32	12	31	47	13	44	48	39	37
	Internacionalización	TI1b	NA	38	NA	22	NA	17	52	16	14	20	52	29	28	49	50	35

Fuente: Elaboración propia a partir de los datos de Pérez y Aldás (dirs.) (2019). En rojo la primera universidad en el ranking, en azul la última universidad en el ranking para un indicador dado. NA significa valor faltante.

	Ámbito	Indicado	1							Universidad									
imensión		Código	UMA	имн	имом	UNU	NED	UNEX	UNICAN	UNILEON	UNIOVI	UNIRIOJA	UNIZAR	uoc	UPC	UPCT	UP		
Docencia	Recursos	DR1c	53	34	10	2	61	20	3	45	11	7	6	62	12	18	5		
		DR2c	49	25	1	23	62	56	9	50	31	14	24	60	2	4	7		
		DR3b	3	32	61	6	12	49	29	15	1	25	17	31	36	40	2		
		DP1d	45	50	9	6	56	38	53	41	55	59	28	11	60	61			
		DP2d	53	45	4	1	62	33	37	25	48	11	35	60	12	61			
		DP3d	44	27	1	7	62	20	32	9	53	49	23	60	52	55	4		
	Calidad	DC2d	52	11	12	5	58	46	31	49	44	54	55	8	13	51			
		DC3R	23	14	50	50	50	29	15	26	28	37	22	50	32	42			
	Internacionalización	DI1d	32	34	36	2	40	62	47	58	60	38	44	46	14	33			
		DI2b	37	51	3	23	60	44	26	28	25	39	21	62	9	30			
	Recursos	IR1a	41	18	59	52	51	36	5	39	28	32	30	53	4	3			
		IR2a	45	49	NA	32	46	47	6	39	18	14	26	51	4	10			
	Producción	IP1a	45	11	62	14	55	43	13	39	27	12	25	46	6	20			
		IP2b	39	31	NA	NA	40	33	19	25	20	29	11	NA	8	6			
		IP3b	28	10	54	18	27	33	40	30	37	35	38	60	14	57			
		IC1ar	31	5	56	3	57	35	2	49	11	52	23	59	21	36			
<u>-</u>	Calidad	IC2ar	34	7	58	18	57	40	1	47	13	27	25	56	2	4			
		IC3ar	36	21	42	16	58	11	1	45	2	48	29	56	23	47			
	Internacionalización	II2g	31	26	4	33	44	54	15	20	36	23	18	14	2	24			
		II3ar	39	37	54	17	58	34	7	33	23	53	16	45	3	52			
rollo		TR1b	45	4	51	18	46	42	15	49	31	47	5	30	9	32			
Desar gico	Recursos	TR2d	24	29	1	15	47	31	4	26	27	44	14	52	6	22			
9		TR3a	35	12	5	27	4	42	30	44	48	38	31	1	33	45			
	Producción	TP1	21	5	25	53	40	17	10	22	36	1	33	56	3	6	4		
	Internacionalización	TI1b	8	11	44	37	36	42	10	51	41	52	24	39	1	48			

Fuente: Elaboración propia a partir de los datos de Pérez y Aldás (dirs.) (2019). En rojo la primera universidad en el ranking, en azul la última universidad en el ranking para un indicador dado. NA significa valor faltante.

		Indicador	dor Universidad														
Dimensiór	Ámbito	Código	UPM	UPNA	UPO	UPV	UPV-EHU	URJC	URLL	URV	US	USAL	usc	υv	UVA	UVIC	uvigo
Docencia	Recursos	DR1c	13	31	44	8	4	59	54	38	40	25	17	39	16	41	26
		DR2c	15	21	40	3	13	59	6	27	55	38	18	30	39	41	35
		DR3b	43	38	10	26	42	44	57	53	48	16	2	21	28	59	20
	Producción	DP1d	62	46	18	22	31	24	8	21	49	30	39	14	36	10	58
		DP2d	44	28	32	14	49	27	18	24	57	26	40	30	34	5	47
		DP3d	42	48	15	59	31	11	13	47	33	5	21	14	12	29	24
	Calidad	DC2d	36	35	15	19	48	25	6	26	56	47	39	20	60	45	17
		DC3R	11	36	18	7	13	20	50	38	9	34	4	3	27	48	40
	Internacionalización	DI1d	27	45	55	10	42	13	5	15	50	24	48	28	54	9	59
		DI2b	20	27	6	11	42	55	29	31	50	32	45	8	17	35	15
ón	Recursos	IR1a	19	27	17	14	22	40	48	15	26	43	10	23	38	54	16
		IR2a	7	20	31	12	42	52	55	23	35	43	16	19	44	3	15
	Producción	IP1a	22	29	32	16	28	42	38	3	33	40	18	9	51	50	17
		IP2b	47	13	14	30	44	10	NA	35	5	37	24	4	42	NA	27
igaci		IP3b	25	58	39	16	46	22	23	8	26	15	31	19	47	61	32
	Calidad	IC1ar	34	46	50	27	15	45	26	8	32	16	17	18	39	40	33
		IC2ar	15	20	54	9	21	33	22	16	26	29	28	30	41	45	19
		IC3ar	33	34	50	20	18	43	3	5	35	9	17	12	53	54	24
	Internacionalización	II2g	6	46	47	8	10	43	59	7	49	48	9	25	39	50	27
		II3ar	27	41	48	24	12	30	1	14	31	25	13	21	32	49	15
Innovación y Desarrollo Tecnológico	Recursos	TR1b	3	11	37	2	34	48	38	21	29	8	13	10	17	51	41
		TR2d	2	30	40	12	37	21	25	5	11	39	16	34	18	56	9
		TR3a	34	43	51	18	50	13	2	15	25	21	41	19	56	28	55
	Producción	TP1	9	29	27	16	41	26	54	24	15	43	18	38	30	56	11
	Internacionalización	TI1b	13	31	19	9	32	40	52	2	6	43	7	26	30	52	21

Fuente: Elaboración propia a partir de los datos de Pérez y Aldás (dirs.) (2019). En rojo la primera universidad en el ranking, en azul la última universidad en el ranking para un indicador dado. NA significa valor faltante.

Cuadro 5. Ur	Cuadro 5. Universidades con mayor y menor número de dominancias en la outranking matrix.						
Universidades en grupo de cabeza							
Orden	Código	Universidad	Dominancias				
1	UPF	Universitat Pompeu Fabra	61				
2	UPC	Universitat Politècnica de Catalunya	59				
3	UAM	Universidad Autónoma de Madrid	56				
4	UC3M	Universidad Carlos III de Madrid	56				
5	UN	Universidad de Navarra	56				
6	UPV	Universitat Politècnica de València	56				
7	UAB	Universitat Autònoma de Barcelona	52				
8	UB	Universitat de Barcelona	51				
9	UNICAN	Universidad de Cantabria	51				
10	UV	Universitat de València	51				
11	URV	Universitat Rovira i Virgili	47				
12	UDL	Universitat de Lleida	47				
51	UJAEN	Universidad de Jaén	9				
52	UCA	Universidad de Cádiz	9				
53	UANE	Universidad Nebrija	9				
54	UVIC	Vic-Universidad Central de Catalunya	8				
55	ULPGC	Universidad de Las Palmas de Gran Canaria	8				
56	COMILLAS	Universidad Pontificia Comillas	6				
57	ини	Universidad de Huelva	5				
58	UCH	Universidad Cardenal Herrera-CEU	3				
59	UOC	Universitat Oberta de Catalunya	2				
60	UCV	Universidad Católica de Valencia San Vicente Mártir	1				
61	UNED	Universidad Nacional de Educación a Distancia	0				
62	UDIMA	Universidad A Distancia de Madrid	0				

Fuente: Elaboración propia a partir de los datos de Pérez y Aldás (dirs.) (2019).

6.4. RESULTADOS Y COMPARACIÓN CON U-RANKING DE RENDIMIENTO

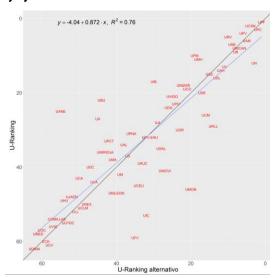
Comparamos ahora los resultados de U-Ranking de rendimiento (Pérez y Aldás [dirs.] 2019) con los del ranking derivado de cambiar el método de agregación y, simultáneamente, todo lo que ello implica en términos del tratamiento de datos. U-Ranking se ofrece en los grandes grupos derivados de redondear el indicador sintético a 1 decimal (Pérez y Aldás [dirs.] 2019, p. 39). Esto es así, tanto en el indicador de rendimiento global, que agrega las universidades en 11 grupos, como en los indicadores de las dimensiones. La Docencia agrega las universidades en 8 grupos, la Investigación en 17 y la Innovación y Desarrollo Tecnológico en 24. Nosotros efectuamos la comparación a partir de la ordenación que sale directamente del indicador sintético, sin ningún redondeo.

En el proceso de agregación de U-Ranking, en el último paso, cuando se agregan las dimensiones la Docencia tiene un peso del 56% cuadro 2-. En una agregación no compensatoria esto significaría darle la mayoría absoluta a la *Docencia*, de forma que el *ranking* global vendría dado por el ranking de Docencia. Para evitar este efecto la solución adoptada es proceder siempre a la agregación desde los rankings individuales. Para mantener la importancia relativa de las dimensiones, y la de los ámbitos dentro de cada dimensión, se recalcularon los pesos implícitos de cada indicador individual en el ranking global de rendimiento. Dichos pesos aparecen en la última columna del cuadro 2. Así podemos ver que los indicadores que más pesan son los de la calidad docente (6.7%) y a continuación los de producción e internacionalización docente (5.7%). En el otro extremo, los indicadores que menos pesan son los de recursos de la dimensión Innovación y Desarrollo Tecnológico (1.4%), pero obsérvese que el único indicador de producción de esta dimensión tiene un peso similar a todos los indicadores de la dimensión de Investigación (3.3%). Estos son los pesos utilizados en la generación del ranking de rendimiento, directamente a partir de la *outranking matrix* de los 25 rankings individuales.

La correlación entre U-Ranking y el *ranking* generado en este trabajo, U-Ranking alternativo, es relativamente elevada, 0.87, sobre todo si tenemos en cuenta los cambios introduci-

dos.²⁷ El gráfico 1 permite apreciar visualmente la correlación entre ambos *rankings*. El número medio de re-*rankings* es de 6, pero el gráfico 1 permite comprobar que esta cifra se debe mayoritariamente a unas pocas universidades, fundamentalmente en el centro de la distribución, que a pesar de todo muestra muy pocos cambios bruscos. En los extremos las alteraciones son escasas y de poca magnitud.

Gráfico 1. Rendimiento de las universidades



Fuente: Elaboración propia a partir de Pérez y Aldás (dirs.) (2019).

Cuatro universidades mantienen exactamente su posición en el ranking, de forma que parece claro qué universidades están en los extremos de la distribución: la Universitat Pompeu Fabra (UPF) en el primer puesto, y la Universidad a Distancia de Madrid (UDIMA) en el último. Otras 36 no ven alterado su orden en más de 5 posiciones. Y solo 10 universidades ven alterada su posición en el ranking más de 10 puestos: 4 de ellas empeoran, Universidad de Nebrija (UANE), Universidad de Burgos (UBU), Universidad de Alicante, (UA) y Universitat de les Illes Balears (IB), mientras que 6 universidades mejoran, Mondragón Unibertsitatea (UMON), Universidad Francisco de Vitoria (UFV), Universitat Internacional de Catalunya (UIC), Universitat Ramon Llull (URLL), Universidad de Oviedo (UNIOVI) y Universidad San Pablo-CEU (UCEU). Resulta interesante observar que un componente importante de la volatilidad se debe a las universidades privadas.

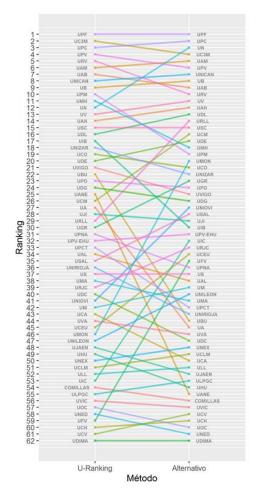
Un examen de la evolución del algoritmo en el proceso de cómputo del *ranking* mostró que la función objetivo, (4), es extremadamente plana

²⁷ En Goerlich (2019), donde sólo se cambia el método de agregación la correlación es de 0.91, lo que muestra que alterar el tratamiento de datos no introduce cambios sustanciales en el ranking finalmente resultante.

en el entorno del óptimo. En la práctica ello significa que pequeñas alteraciones en el orden de algunas universidades generan variaciones muy pequeñas, o prácticamente despreciables, en el valor de la función objetivo, o lo que es lo mismo, varios *rankings* con pequeñas alteraciones en el orden de algunas universidades obtienen prácticamente el mismo soporte en nuestro conjunto de información. Ello muestra lo difícil que es, en la práctica, ordenar universidades que son diversas en varios sentidos y apunta a ofrecer resultados por grupos de universidades.

El gráfico 2 permite examinar visualmente los re-rankings debidos al cambio de metodología que, salvo las excepciones mencionadas, no son excesivos. Naturalmente el grado de re-rankings es mucho menor si consideramos los 11 grupos de universidades en los que se publican finalmente los resultados de U-Ranking para las 62 universidades.

Gráfico 2. *Re-rankin*g de universidades según la metodología empleada



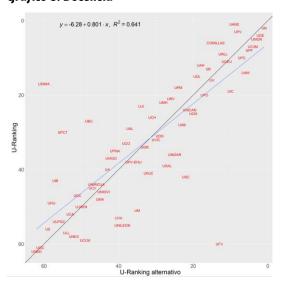
Fuente: Elaboración propia a partir de Pérez y Aldás (dirs.) (2019).

6.5. RESULTADOS Y COMPARACIÓN CON U-RANKING PARA LAS DIMENSIONES

Conviene examinar los resultados para las distintas dimensiones ya que, aunque no son radicalmente diferentes de lo que sucede en el *ranking* global, existen algunas discrepancias dignas de mención.

Docencia resulta ser la dimensión más volátil. La correlación entre ambos rankings es solo de 0.80. Esta mayor volatilidad es perceptible en el gráfico 3, sobre todo en el centro de la distribución, donde se producen algunos saltos bruscos de considerable magnitud. Así, llama la atención la mejora, respecto a U-Ranking, de la Universidad Francisco de Vitoria (UFV), que pasa de la posición 60 en U-Ranking a la posición 13 en el ranking alternativo, y el empeoramiento de la Universidad a Distancia de Madrid (UDIMA), que pasa de la posición 17 en U-Ranking a la posición 60 en el ranking alternativo. De nuevo llama la atención que se trate de dos universidades privadas. Un examen de los datos originales muestra que se trata de dos universidades que muestran valores muy extremos en algunos indicadores, situándose en las primeras posiciones en unos y en las últimas en otros. La utilización de medias geométricas como criterio de agregación por parte de U-Ranking, muy sensible a valores extremos, podría explicar estas diferencias.

Gráfico 3. Docencia



Fuente: Elaboración propia a partir de Pérez y Aldás (dirs.) (2019).

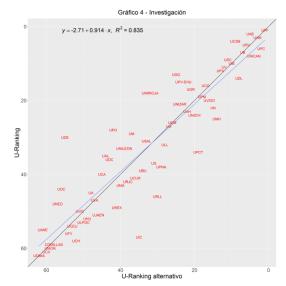
El número medio de re-rankings es 7. De nuevo, cuatro universidades mantienen exactamente su posición en el ranking, incluidas las dos últimas: Universitat Oberta de Catalunya (UOC) y Universidad Nacional de Educación a

Distancia (UNED), ambas universidades no presenciales. La universidad líder en esta dimensión si se altera, ya que en U-Ranking es la Universidad Nebrija (UANE) mientras que en el ranking alternativo es la Universidad de Navarra (UN), que en U-Ranking ocupa la segunda posición, ambas privadas. Otras 33 universidades no ven alterado su orden en más de 5 posiciones. Y 12 universidades ven alterada su posición en el *ranking* más de 10 puestos: 5 de ellas empeoran, además de la Universidad a Distancia de Madrid (UDIMA), también lo hacen la Universidad Politécnica de Cartagena (UPCT), Universidad de Burgos (UBU), la Universitat de les Illes Balears (UIB) y la Universitat Jaume I de Castellón (UJI); mientras que 7 universidades mejoran, además de la Universidad Francisco de Vitoria (UFV), también lo hacen la Universidad de Santiago de Compostela (USC), la *Universidad de Murcia* (UM), la *Universidad* de León (UNILEON), la Universidad de Valladolid (UVA), la Universidad de Salamanca (USAL) y la Universidad de Zaragoza (UNIZAR).

Por el contrario, la dimensión de *Investigación* resultó ser la más estable, con un número medio de re-*rankings* de 5 y un coeficiente de correlación entre ambos *rankings* de 0.91.²⁸ Esta mayor estabilidad es también perceptible en el gráfico 4 que ya permite apreciar un menor número de alteraciones, y sobre todo mucho menos bruscas que en la dimensión de docencia.

En este caso son 8 universidades las que mantienen exactamente su posición en ambos rankings, incluidas de nuevo las universidades en los extremos, que coinciden con las universidades en primera y última posición del ranking global: la Universitat Pompeu Fabra (UPF) en el primer puesto y la Universidad a Distancia de Madrid (UDIMA) en el último. Un total de 32 universidades no ven alterado su orden en más de 5 posiciones. Y solo 9 universidades ven alterada su posición en el ranking más de 10 puestos: 5 de ellas empeoran, la Universidad de Deusto (UDE), la Universidad de La Rioja (UNIRIOJA), la Universidad Pablo de Olavide (UPO), la Universitat de Girona (UDG) y la Universitat Oberta de Catalunya (UOC), mientras que 4 universidades mejoran, la Universitat Internacional de Catalunya (UIC), la Universitat Ramon Llull (URLL), la Universidad Politécnica de Cartagena (UPCT) y la Universidad Miguel Hernández de Elche (UMH).

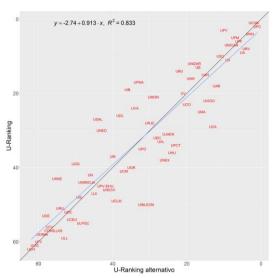
Gráfico 4. Investigación



Fuente: Elaboración propia a partir de Pérez y Aldás (dirs.) (2019).

Finalmente la dimensión de *Innovación y Desa-rrollo Tecnológico* ofrece un número medio de re-*rankings* de 6 y un coeficiente de correlación entre ambos *rankings* de 0.91. Al igual que sucede con la dimensión de *Investigación* esta estabilidad es perceptible en el gráfico 5 que permite apreciar, de nuevo, un comportamiento más estable que en el caso de la dimensión de Docencia.

Gráfico 5. Innovación y desarrollo tecnológico



Fuente: Elaboración propia a partir de Pérez y Aldás (dirs.) (2019).

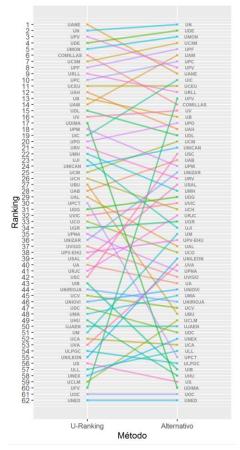
En este caso 6 universidades mantienen exactamente su posición en ambos *rankings*, incluidas las 3 universidades en la cola de la distribución: *Universidad Cardenal Herrera-CEU* (UCH), *Vic-Universidad Central de Catalunya* (UVIC) y *Universidad Francisco de Vitoria* (UFV), mientras que las dos primeras universidades se alternan en cada uno de los *rankings*,

²⁸ Para esta dimensión el algoritmo devolvió siete *ran-kings* óptimos con diferencias mínimas.

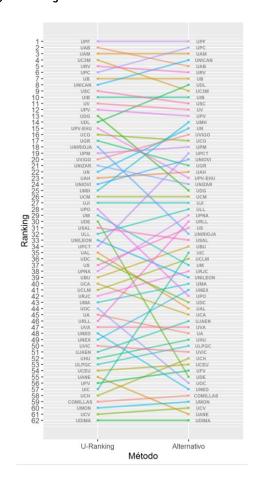
en U-Ranking la primera es la Universidad Carlos III de Madrid (UC3M) y la segunda la Universitat Politècnica de Catalunya (UPC), mientras que en el ranking alternativo ambas intercambian sus posiciones. Un total de 31 universidades no ven alterado su orden en más de 5 posiciones. Y 12 universidades ven alterada su posición en el ranking más de 10 puestos: 7 de ellas empeoran, la Universitat de les Illes Balears (UIB), la Universidad de Salamanca (USAL), la Universidad Pública de Navarra (UPNA), la Universidad Nacional de Educación a Distancia (UNED), la Universidad Nebrija (UA-NE), la Universitat de Lleida (UDL) y la Universitat de Girona (UDG), mientras que 5 universidades mejoran, la Universidad de León (UNI-LEON), la Universidad de Cádiz (UCA), la Universidad de Huelva (UHU), la Universidad de Extremadura (UNEX) y la Universidad Politécnica de Cartagena (UPCT).

Gráfico 6. Re-ranking de universidades por dimensiones según la metodología empleada)

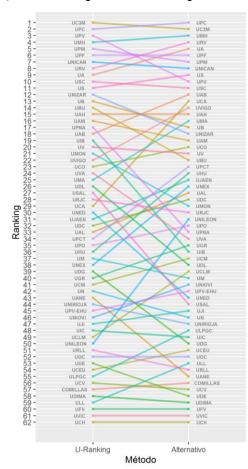
a) Docencia



b) Investigación



c) Innovación y desarrollo tecnológico



Fuente: Elaboración propia a partir de Pérez y Aldás (dirs.) (2019).

7. CONCLUSIONES

Este trabajo ha expuesto una metodología alternativa para la generación de rankings óptimos o de consenso a partir de rankings individuales. Dichos métodos presentan dos ventajas fundamentales. Por una parte, se trata de métodos de agregación no compensatorios, que contrastan frente a los métodos tradicionales de agregación basados en promedios aritméticos o geométricos, que siempre implican un cierto grado de compensación entre los indicadores de partida. Por otra parte, permiten evitar muchas de las decisiones sobre el tratamiento de datos previo que es necesario afrontar en la construcción de indicadores compuestos, tales como el tratamiento de outliers, la normalización, o la imputación de valores faltantes (OCDE/Unión Europea/JRC 2008). El precio que hay que pagar por ello es que tratan solo información ordinal, los rankings individuales de los indicadores de partida, de forma que perdemos cualquier información de tipo cardinal.

Los métodos expuestos tienen una base teórica sólida basada en la teoría de la elección social y la evaluación multicriterio (Munda 2008), al mismo tiempo que tratan un problema ampliamente conocido en la literatura matemática relacionada con la Investigación Operativa (Martí y Reinelt 2011). Frente a los métodos tradicionales más sencillos presentan el inconveniente de su mayor complejidad computacional y la posibilidad de aparición de soluciones no únicas, lo que obliga a buscar un criterio de selección del ranking óptimo en el caso de que se dé esta situación. Desde el punto de vista computacional los métodos exactos solo son actualmente operativos para problemas de dimensión relativamente reducidos, del orden de unos 20 o 30 objetos a ordenar, en función de lo fuerte que sea la información en los rankings de partida. Se dispone, sin embargo, de algoritmos heurísticos que resuelven el problema en un tiempo razonable para un número de objetos a ordenar mucho mayor.²⁹

El trabajo realiza una aplicación de los métodos expuestos a la determinación de un ranking de universidades españolas a partir de los datos originales de U-Ranking (Pérez y Aldás [dirs.] 2019). El ejercicio aplica la metodología desde los datos originales de partida, y por tanto los cambios entre nuestros resultados y los ofrecidos por U-Ranking 2019 se deben no solo al método de agregación, sino también al tratamiento previo de datos. En particular nuestro ejercicio permite evitar realizar imputaciones por valores faltantes, no tratar los valores atípicos - outliers y no efectuar normalización alguna, ya que los datos de partida son los rankings derivados de los indicadores originales, lo que constituye una normalización implíci-

A pesar de todos estos cambios, los resultados no muestran alteraciones sustanciales respecto a los ofrecidos por U-Ranking 2019, sobre todo en los extremos de la distribución. Para el ranking de rendimiento global la correlación entre ambos rankings es del 0.87. Las mayores alteraciones se producen en universidades situadas mayoritariamente en el centro de la distribución, lo que muestra lo difícil que es ordenar situaciones intermedias. Por otra parte, un volumen importante de cambios de un orden de magnitud apreciable se produce entre universidades privadas, que tienden a mostrar una mayor volatilidad que las universidades públicas. Una parte de esta volatilidad podría deberse a su mayor especialización en determinadas dimensiones, notablemente docencia, y al hecho de que diversos criterios de agregación sean sensibles al grado de especialización de las universidades.

Desde el punto de vista de los *rankings* de las diferentes dimensiones la *Docencia* es, sin duda, la dimensión más volátil, o donde apreciamos más diferencias entre ambas metodologías, con un coeficiente de correlación entre *rankings* de 0.80. Es de destacar, sin embargo, que las dos primeras universidades en el *ranking*, según ambas metodologías, son privadas. Por el contrario las dimensiones de *Investigación* e *Innovación y Desarrollo Tecnológico* son mucho más estables, sobre todo la primera, con una correlación entre los *rankings* de ambas metodologías de 0.91, y muchas menos

²⁹ Por otra parte, la experimentación con algoritmos más sofisticados mostró una correlación extremadamente alta con la regla de Copeland (Moulin 1989), del orden de 0.99, cuyo cálculo es inmediato, y por tanto constituye una alternativa más que razonable a la solución exacta del problema de optimización (4).

alteraciones sustanciales entre los *rankings* de las universidades.

Naturalmente, estas discrepancias serían mucho menores si la comparación se efectuara al nivel de agregación al que se presentan los resultados de U-Ranking 2019, que mapean las 62 universidades en 11 grupos en el *ranking* global de rendimiento, 8 grupos en el caso del *ranking* de *Docencia*, 17 grupos en el caso del *ranking* de *Investigación* y 24 grupos en el caso del *ranking* de *Innovación y Desarrollo Tecnológico*.

En conjunto, y aunque el presente trabajo no constituye un análisis de sensibilidad de U-Ranking, los resultados parecen bastante robustos a los cambios introducidos para la gran mayoría de universidades, sobre todo si tenemos en cuenta que no solo se altera el método de agregación, sino también una gran parte del tratamiento previo de datos. No obstante, siempre es posible encontrar alguna excepción a esta estabilidad, normalmente entre las universidades privadas.

APÉNDICE

CONVENCIONES DE NOTACIÓN

Un *ranking* no es más que una ordenación (completa) de objetos de acuerdo con un determinado criterio. Supongamos que los objetos a ordenar son 5: (A, B, C, D, E), y que el *ranking* viene determinado por $\langle D$, A, B, E, $C \rangle$. Esta representación es útil en los informes, pero no es operativa desde el punto de vista computacional, para lo cual necesitamos una representación numérica del *ranking*. Con n objetos una forma natural de representar un *ranking* utiliza los primeros n números naturales, de n a n.

Sin embargo en la práctica existen muchas formas equivalentes de representar numéricamente un *ranking*. La primera de ellas simplemente asigna los valores del *ranking* a la lista de objetos, empezando con el *ranking* del primer objeto. En este caso el *ranking* sería (2, 3, 5, 1, 4), que indica que *A* ocupa la segunda posición en el *ranking*, *B* la tercera posición, y así sucesivamente.

Otra forma representa el *ranking* a partir del orden que ocupan los objetos en la lista, que mantenemos fija, utilizando números como etiquetas. En este caso el *ranking* sería $\langle 4, 1, 2, 5, 3 \rangle$, que indica que el primer objeto del *ranking* es el cuarto en la lista de objetos, D, el segundo objeto del *ranking* es el primero en la lista de objetos, A, y así sucesivamente.

Obsérvese que en la primera representación asignamos el valor numérico del *ranking* a cada uno de los objetos de una lista dada, mientras que en la segunda asignamos el orden que ocupan dichos objetos en la lista al *ranking* propiamente dicho. En este caso el *ranking* viene representado por el orden de los objetos en la lista.³⁰

Una representación alternativa, poco compacta, pero muy útil desde el punto de vista computacional, es una representación matricial. Para n objetos, podemos representar un ranking por medio de una matriz $n \times n$, S, tal que

$$s_{ij} = \begin{cases} 1 & \text{si el objeto } i \text{ precede en el ranking al objeto } j \\ 0 & \text{si el objeto } i \text{ sucede en el ranking al objeto } j \end{cases}$$
(A1)

Y ceros en la diagonal principal, $s_{ii}=0$, $\forall i$. En este caso el $\emph{ranking}$ vendría dado por

que indica que A precede en el ranking a B, a C y a E, pero no a D; B precede en el ranking a C y a D, pero no a A ni a D; y así sucesivamente. Naturalmente esta representación verifica $s_{ii} + s_{ii} = 1, \ \forall i \neq j$.

A partir de dicha matriz siempre es posible seguir el camino inverso y recuperar el *ranking* en su forma original. Sumando por filas la matriz anterior obtenemos el número de dominancias de cada objeto sobre el resto

Que es otra forma de expresar el *ranking*, esta vez con los números en orden decreciente, y acabando en cero. Es decir, D tiene por detrás a 4 objetos en el *ranking*, A a 3 objetos, B a 2 objetos, E a 1 objeto, y C es el último objeto del *ranking*, por lo que no tiene a nadie por detrás: $\langle D, A, B, E, C \rangle$.

Obsérvese que la suma por filas de la matriz anterior no es más que la cuenta de Borda (1781) con las puntuaciones originales (Emerson 2013).

³⁰ Por esta razón, en *R*, la primera representación se obtiene mediante la instrucción rank y la segunda mediante la instrucción order.

³¹ La función condorcetscore, en *R*, implementa esta representación, que no es más que la matriz de Condorcet o *outranking matrix* para un único *ranking*.

Este caso, en el que todos los objetos son comparados y ordenados, se conoce con el nombre de **orden lineal** (Cook, Kress y Seiford 1986), y es el único tratado por una gran parte de la literatura matemática sobre el tema (Martí y Reinelt 2011).

Empates (ties)

En la práctica, cuando comparamos dos objetos debemos acomodar la posibilidad de que no podamos ordenarlos porque no exista una relación clara de preferencia entre ellos. Es decir, debemos acomodar la posibilidad de **empates** (*ties*).

Por ejemplo, supongamos que somos **indiferentes** en el orden que ocupan los objetos B y E. El ranking vendría dado ahora por $\langle D, A, B - E, C \rangle$, donde '-' indica el empate entre B y E. Podríamos escribir igualmente $\langle D, A, E - B, C \rangle$. Alternativamente podríamos decir que consideramos que los objetos B y E son **equivalentes**. Aunque como veremos a continuación los términos indiferencia y equivalencia pueden tener connotaciones distintas.

La primera de las representaciones numéricas podría escribirse ahora de varias formas alternativas. Por ejemplo, (2, 3, 5, 1, 3), que indica que *B* y *E* ocupan, ambos, la tercera posición en el *ranking*. Pero también sería posible escribir, (2, 4, 5, 1, 4), o (2, 3.5, 5, 1, 3.5), abandonando la representación numérica mediante números naturales, o incluso (2, 3, 4, 1, 3), que no deja *gaps*, y que es lo que la literatura llama *rankings* densos. Desde el punto de vista práctico es irrelevante, puesto que todo lo que importa es el orden entre los objetos, aunque a efectos computacionales conviene mantener la homogeneidad en las representaciones.

La segunda de las representaciones numéricas puede escribirse ahora como $\langle 4, 1, 2-5, 3 \rangle$, que indica que los objetos segundo, B, y quinto, E, de la lista ocupan la tercera posición en el *ranking*. Empates de cualquier orden pueden ser representados numéricamente de esta forma.

La representación matricial anterior puede ser acomodada ahora de esta forma,

$$s_{ij} = \begin{cases} 1 & \text{si el objeto } i \text{ precede en el ranking al objeto } j \\ 0.5 & \text{si el objeto } i \text{ está empatado con el objeto } j \\ 0 & \text{si el objeto } i \text{ sucede en el ranking al objeto } j \end{cases}$$

(A2)

Y ceros en la diagonal principal, $s_{ii}=0$, $\forall i$. En este caso el $\emph{ranking}$ vendría dado por

que indica que B precede en el ranking a C, está empatado con E, y va precedido por A y D. Obsérvese que esta representación sigue verificando que $s_{ij} + s_{ji} = 1, \ \forall i \neq j$.

Al igual que antes, sumando por filas obtenemos las puntuaciones originales de Borda (1781) en el caso de empates (Emerson 2013),

que simplemente promedia la dominancia para los objetos que resultan indiferentes.

Cuando consideramos *rankings* que pueden incluir empates hablamos de *rankings* **débiles** (Cook, Kress y Seiford 1986).

Aunque esta acomodación de los empates en la representación numérica de los *rankings* parece sencilla y directa, esto no es más que una ilusión. Desde el punto de vista práctico y conceptual la presencia de empates es fuente de innumerables problemas, tanto prácticos como computacionales. Muchos algoritmos que agregan *rankings* de forma equivalente en el caso de órdenes lineales no lo hacen en presencia de empates.

La existencia de un empate entre dos objetos puede tener en la práctica diferentes interpretaciones, que giran entorno a la distinción conceptual entre indiferencia y equivalencia. Por eiemplo, si el rankina responde a una opinión -o al resultado de 2 indicadores-puede interpretarse como que los objetos empatados resultan indiferentes, en el sentido de que se desea que sean las opiniones de otras personas -o el valor de otros indicadores- las que determinen el orden final entre estos dos objetos. Alternativamente, puede pensarse que más que una indiferencia se trata de que son realmente equivalentes en un sentido más fuerte, y ninquno de esos objetos debe ser realmente preferido al otro. En este caso, claramente, los rankings con empates deben formar parte del espacio de soluciones posibles entre las que buscar a la hora de agregar *rankings* individuales.

Los empates, encontrados con frecuencia en la práctica, ofrecen algunas cuestiones interesantes sobre las que reflexionar cuando pensamos en que debemos considerar como una solución al problema de la agregación de rankings individuales. Por ejemplo, supongamos que solo tenemos 2 objetos a ordenar, (A, B), y disponemos de 3 *rankings* individuales: (i) $\langle A-B \rangle$, (ii) $\langle A-B \rangle$ y (iii) $\langle A, B \rangle$. En este ejemplo, los 2 primeros rankings muestran indiferencia entre A y B, mientras que el tercero ordena a A por delante de B. Si la solución que buscamos es un ranking lo más similar posible a los rankings individuales, entonces el *ranking* con empates, $\langle A-B \rangle$, será la elección correcta. Sin embargo, si queremos buscar un ganador, entonces observamos que A es indiferente a B dos veces, pero ordenado por delante de él otra vez, por lo que la elección correcta desde esta perspectiva sería el ranking $\langle A, B \rangle$. Lo que este ejemplo trata de ilustrar es que pueden haber diferentes soluciones al mismo problema, todas ellas correctas, dependiendo del enfoque, y que, normalmente, rankings con empates deben formar parte del espacio de soluciones.

Una reflexión final en relación a los empates y al espacio de soluciones posibles. Incluso cuando el objeto del ejercicio requiere la determinación de un ganador, puede haber situaciones en las que un empate sea la solución más razonable. Este es el caso, por ejemplo, si los *rankings* individuales muestran simetría en la preferencia por todos los objetos. Como ejemplo trivial, consideremos el caso en el que, con dos objetos a ordenar, (A, B), y disponemos de 2 *rankings* individuales simétricos: (i) $\langle A, B \rangle$ y (ii) $\langle B, A \rangle$. En este caso la única solución razonable es el *ranking* con empate: $\langle A-B \rangle$, incluso aunque no haya presente ningún empate en los *ranking* individuales.³²

El mensaje de estas reflexiones es que parece razonable que en el espacio de soluciones se incluyan *rankings* con empates (*ties*), es decir *rankings* débiles.

No comparabilidad

ordenar están empatados.

Otra cuestión que aparece con frecuencia en la práctica es el caso de *rankings* parciales, es decir, situaciones en las que 2 objetos no pue-

decir, situaciones en las que 2 objetos no pue
32 Más generalmente, cuando la *outranking matrix* es constante, en el sentido de que el triángulo superior de la matriz es igual al triángulo inferior, entonces el ran-

king óptimo es aquel en el que todos los objetos a

den ser comparados. Por ejemplo, a pesar de que los objetos a ordenar son 5: (*A, B, C, D, E*), podemos encontrar una situación en la que no dispongamos de información sobre uno (o más) de los objetos en cuestión, por ejemplo *B, (A, NA, C, D, E)*.

Este tipo de situaciones pueden darse por muchos motivos, y en cierta forma son específicas a cada problema, como también lo es la solución a aplicar. Por ejemplo, en la teoría de las votaciones es fácil tener muchos candidatos, pero los votantes solo expresan sus preferencias sobre unos pocos. En este caso, parece natural pensar que aquel candidato que no ha sido votado por un individuo va a la cola del ranking, y el resto de candidatos -objetos- le preceden. Manteniendo, en nuestro ejemplo, la misma ordenación que antes para el resto de objetos, el *ranking* sería pues $\langle D, A, E, C, B \rangle$, que puede ser representado numéricamente por cualquiera de las formas anteriores. Obsérvese que si tenemos varios objetos para los que no disponemos de información, y adoptamos esta solución, generamos empates (ties) entre todos ellos, que son precedidos por todos los objetos para los que sí disponemos de información. Este es el enfoque adoptado por la mayor parte de la literatura teórica, mucha de ella relacionada con los sistemas de búsqueda de información web o el meta-análisis (Fagin, Kumar y Sivakumar 2003; Fagin et al. 2006; Ailon 2010).

Puede darse la situación de que la falta de información sobre un objeto no signifique que debamos ponerlo a la cola, sino que por el contrario esta sea una solución poco razonable. ³³ Por ejemplo, si los *rankings* proceden de indicadores es posible que no dispongamos información sobre algunos objetos (*missing values*). En este caso, y si no queremos efectuar imputación alguna, parece más honesto tratar la falta de información como no comparabilidad, que poner el objeto en cuestión a la cola del *ranking*. La representación del *ranking* simplemente omite este objeto, aunque lo mantengamos en la lista de objetos a ordenar. El *ranking* sería pues (*D*, *A*, *E*, *C*).

La representación numérica en estos casos es algo más sutil. En la primera de las representaciones eliminar B del ranking, (2, 4, 1, 3), no es adecuado, puesto que no podemos rastrear

³³ Puede darse el caso, incluso, que estos objetos deban ponerse a la cabeza del ranking, si la votación lo que desea es excluir a los candidatos –objetos– menos deseados, o las búsquedas *web* hagan referencia a los objetos menos valorados.

que sucede con *B*, que aparece en la lista de objetos a ordenar, pero no en el *ranking*. Observa que tenemos 5 objetos en la lista de objetos a ordenar, pero solo 4 en el *ranking*. Lo más adecuado en este caso en mantener a *B* en el *ranking* con un carácter especial, que indique expresamente que para ese objeto no disponemos de información, por ejemplo (2, *NA*, 4, 1, 3).

En la segunda de las representaciones, que muestra el *ranking* en términos del orden de los objetos en la lista, eliminando B si permite identificar que este objeto no es comparable con los demás, $\langle 4, 1, 5, 3 \rangle$. Puesto que el objeto 2, B, no aparece en la representación del *ranking*, concluimos que este no es comparable con los demás.

Ninguna de estas representaciones es práctica desde el punto de vista computacional. Sin embargo, la representación matricial si lo es. En este caso, definimos la matriz **S**, tal que,

$$s_{ij} = \begin{cases} 1 & \text{si el objeto } i \text{ precede en el ranking al objeto } j \\ 0.5 & \text{si el objeto } i \text{ está empatado con el objeto } j \\ 0 & \text{en caso contrario, incluido la no-comparabilidad} \end{cases}$$

(A3)

Y ceros en la diagonal principal, $s_{ii}=0$, $\forall i$. En este caso, cuando para un objeto no disponemos de información, la correspondiente fila y columna de ${\bf S}$ tiene todos sus elementos cero. El ranking vendría dado por

que indica que B no es comparable con el resto de objetos —no disponemos de información para ello—, razón por la cual su fila y su columna es idénticamente igual a cero. Obsérvese que en este caso ya no se cumple $s_{ij}+s_{ji}=1, \ \forall i\neq j$. Esta relación se sigue manteniendo cuando los objetos, i y j, son comparables, mientras que si no lo son tenemos, $s_{ij}=s_{ji}=0, \ \forall i\neq j$. Este enfoque en el tratamiento de objetos no comparables, que no necesariamente deben ponerse a la cola, no ha sido prácticamente tratado por la literatura, y solo se menciona en Emond y Mason (2000, 2002).

Al igual que antes, sumando por filas obtenemos las puntuaciones originales de Borda (1781) cuando existen casos no comparables, una situación que ha sido escasamente tratada por la literatura (Emerson 2013),

Cuando consideramos *rankings* con falta de información para algunos objetos hablamos de *rankings* parciales o incompletos (Cook, Kress y Seiford 1986).

Ponderaciones

Un último aspecto práctico que debemos acomodar es la incorporación de pesos. Es posible que estos sean innecesarios, de forma que todos los *rankings* tengan el mismo peso a la hora de buscar un *ranking* agregado o representativo, por ejemplo en el caso de las votaciones el principio de 'un hombre — un voto' implica que todos los votos debe pesar lo mismo, es lo que se conoce como el **principio de anonimato**. Sin embargo, en la construcción de *rankings* agregados a partir de indicadores individuales es posible, o incluso deseable, que unos indicadores tengan más importancia que otra, y lleven asociados pesos diferentes.

Solo la representación matricial de un *ranking* acomoda pesos con facilidad. Si, en nuestro ejemplo, el *ranking* individual es $\langle D, A, B, E, C \rangle$, y el peso asociado a dicho *ranking* 8, su correspondiente representación matricial, incorporando el peso, es simplemente

Es decir, se trata simplemente de multiplicar el peso por la representación matricial del *ranking* sin el peso. Naturalmente los pesos pueden estar normalizados de forma que sumen la unidad. Y de forma idéntica para el caso en el que aparezcan empates (*ties*) o casos de no comparabilidad.

LA APROXIMACIÓN VÍA COEFICIENTE DE CORRELACIÓN τ DE KENDALL

Las sección 3 han definido la matriz de Condorcet o *outranking matrix* como la matriz de dominancias en comparaciones bilaterales entre todos los pares posibles de objetos a ordenar y, a partir de ella, ha definido el *ranking* óptimo o de consenso como aquel recibe el mayor apoyo por parte de la *outranking matrix*, es decir el *ranking* ganador entre todas las comparaciones bilaterales. La sección 4 formaliza este problema de optimización, (4), que tiene una justificación estadística en el principio de máxima verosimilitud. Finalmente, la sección 5 engarza dicha solución con la teoría de la elección social y la teoría de la evaluación multicriterio.

Emond y Mason (2000, 2002) han propuesto un concepto de solución similar, en realidad idéntico cuando no existen empates (ties), al ya mencionado de buscar el máximo apoyo en la *outranking matrix*. Su punto de partida es el coeficiente de correlación de rangos, τ , de Kendall (1948), aunque como veremos a continuación no en su formulación original.

El concepto de solución de Emond y Mason (2000, 2002) se basa en encontrar el *ranking* **que maximiza la correlación media ponderada con los** *m rankings* **individuales, utilizando una variante del coeficiente de correlación de rangos, \tau_b, de Kendall (1948), en el caso más general en que podemos tener empates y casos no comparables, y al que ellos denominan \tau_v.**

Kendal (1948) definió el coeficiente de correlación entre dos *rankings* como el número mínimo de intercambios entre pares de objetos adyacentes requeridos para ir de un *ranking* a otro, convenientemente estandarizado. Esta es una medida inequívoca en el caso de *rankings* completos y sin empates, pero algo más difusa cuando esto no es así. En este caso más general, Kendal (1948) parte de una formulación matricial de un *ranking* diferente de la representación matricial introducida al principio de este apéndice. En concreto, para n objetos, Kendal (1948) representa un *ranking* por medio de una matriz $n \times n$, A, tal que³⁴

$$a_{ij} = \begin{cases} 1 & \text{si el objeto } i \text{ precede en el ranking al objeto } j \\ -1 & \text{si el objeto } i \text{ sucede en el ranking al objeto } j \\ 0 & \text{en caso contrario: empate, no-comparabilidad o } i = j \end{cases}$$
(A4)

En este caso el $ranking \langle D, A, B, E, C \rangle$ vendría dado por

El ranking (D, A, B-E, C) vendría dado por

Y el $ranking \langle D, A, E, C \rangle$ vendría dado por

Y a partir de aquí Kendall (1948) define el coeficiente de correlación entre 2 *rankings*, **A** y **B**, como

$$\tau_b = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij}}{\sqrt{\sum_{i=1}^{n} \sum_{i=1}^{n} a_{ij}^{2} \cdot \sum_{i=1}^{n} \sum_{i=1}^{n} b_{ij}^{2}}}$$
(A5)

Si disponemos de m rankings individuales encontrar el ranking, r, con representación matricial \mathbf{S} , $\{s_{ij}\}$, que maximiza la correlación media ponderada, medida por τ_b , con estos m rankings individuales, consiste en identificar \mathbf{S} que maximiza la siguiente expresión

$$\underset{S \in \Omega}{Max} \frac{\sum_{k=1}^{m} \omega_k \tau_b(\boldsymbol{S}, \boldsymbol{D}^k)}{\sum_{k=1}^{m} \omega_k}$$
 (A6)

donde Ω es el espacio de *rankings* bajo consideración y \boldsymbol{D}^k es la representación matricial del *ranking* individual k definida de acuerdo con (A4).

³⁴ Kendall (1948) no trató los casos de *rankings* parciales o incompletos, pero 0 es la asignación natural como ya se ha mencionado anteriormente. Asignar valores faltantes, como hace la función kemenyscore, de la librería ConsRank de *R*, impide operar con esta formulación matricial en el caso de rankings parciales o incompletos.

Este es el concepto de solución propuesto por Emond y Mason (2000, 2002), pero sustituyendo τ_b por una modificación del mismo definido a partir de una representación matricial de un ranking alternativa a (A4). Esta modificación solo es relevante en el caso de que existan rankings débiles.

Cuando los rankings son completos y sin empates el denominador de τ_b viene dado por n.(n-1), de forma que podemos simplificar τ_b como

$$\tau_b = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij}}{n.(n-1)}$$
 (A7)

Y la función objetivo a maximizar puede ser escrita como

$$\underset{S \in \Omega}{Max} \frac{\sum_{k=1}^{m} \omega_{k} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij} d_{ij}^{k} \right\}}{n.(n-1)}$$
(A8)

El denominador es irrelevante ahora, e intercambiando los sumatorios podemos simplificar (A8) como

$$\underset{S \in \Omega}{Max} \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij} \sum_{k=1}^{m} \omega_{k} d_{ij}^{k} \iff \underset{S \in \Omega}{Max} \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij} e_{ij}$$
(A9)

donde $e_{ij} = \sum_{k=1}^m \omega_k d_{ij}^k$. Este problema tiene exactamente la misma estructura que (4), solo que las matrices que representan los *rankings* individuales han sido construidas de distinta forma.

Es posible demostrar, sin embargo, que en el caso de rankings sin empates existe una relación uno-a-uno entre la representación (A1) y la representación (A4). Si a_{ij} son los elementos de la representación (A1) y e_{ij} son los elementos de la representación (A4) es fácil ver que $e_{ij}+\mid e_{ij}\mid=2.a_{ij}$, $\forall i,j$, puesto que los elementos de e_{ij} fuera de la diagonal principal son todos 1 o -1, y los valores 1 se corresponden con el mismo valor en a_{ij} , mientras que los valores -1 se corresponden con el valor 0 en a_{ij} .

La conclusión es que, en el caso de rankings completos y sin empates, maximizar la correlación media ponderada, medida por τ_b , con estos m rankings individuales es equivalente a encontrar el ranking óptimo o de consenso definido como aquel ranking que tiene el mayor soporte en todas las comparaciones bilaterales. Es decir, es indiferente resolver (A9) o resolver (4). Esta relación uno-a-uno entre ambas representaciones matriciales no se cumple en el caso de *rankings* débiles, es decir entre (A2) y (A4). Aunque si se verifica en el caso de *rankings* parciales, siempre que no haya empates, es decir entre (A3) y (A4) cuando no hay empates. En estos casos el problema de optimización (A6) no se puede simplificar en términos de (A8), porque (A5) no se simplifica en (A7), y el problema de optimización (A6) debería arrastrar la normalización de τ_b . Lo que lo complica de forma sustancial.

Cuando tenemos empates, o *rankings* incompletos, el denominador será una cantidad inferior a n.(n-1), dado que la representación (A4) utiliza ceros para representar los empates, y también los *rankings* incompletos. En estos casos la normalización que aparece en τ_b es necesaria para asegurar la propiedad obvia de que la correlación de un *ranking* consigo mismo sea 1, y no un valor inferior.

A partir de esta observación, y de algunos problemas en la definición de τ_b cuando hay empates (Emond y Mason 2002, sec.-3.3), Emond y Mason (2000, p. 10) se preguntan cómo representar matricialmente un *ranking* con empates para que el factor de normalización de un coeficiente de correlación de rangos tipo Kendall (1948) sea n.(n-1). En este caso estos autores muestran que el *ranking* **A** debe representarse como (Emond y Mason 2002, p. 21)³⁵

$$a_{ij} = \begin{cases} 1 & \text{si el objeto } i \text{ precede en el ranking al objeto } j, \\ \text{o está empatado con él} \\ -1 & \text{si el objeto } i \text{ sucede en el ranking al objeto } j \\ 0 & \text{si } i = j \text{ o los objetos no son comparables} \end{cases}$$
(A10)

En este caso $ranking \langle D, A, B-E, C \rangle$ vendría dado por

³⁵ Emond y Mason (2002, p. 21) no consideran en la representación (A10) el caso de rankings incompletos, que son añadidos de forma intuitiva posteriormente (Emond y Mason 2002, sec.-8.1). Esto genera ciertos problemas con su definición es este caso, como se pone de manifiesto en el texto.

Y el coeficiente de correlación propuesto por Emond y Mason (2002, p. 21) entre 2 *rankings,* **A** y **B**, es

$$\tau_{x} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij}}{n.(n-1)}$$
 (A11)

Obsérvese que $\tau_{_{x}}=\tau_{_{b}}$ en el caso de rankings completos sin empates, pero difieren en caso contrario, por tanto la extensión de Emond y Mason (2002) del coeficiente de correlación de Kendall (1948) solo aplica en el caso de empa-correlación de un ranking consigo mismo es 1 aunque haya empates. Sin embargo, τ_{ν} es menor que uno para la correlación de un ranking consigo mismo en el caso de rankings incompletos (Emond y Mason 2002, sec.-8.1), un hecho que parece haber sido pasado por alto por parte de Emond y Mason (2000, 2002). Así pues, todo indica que arrastrar la normalización de $\tau_{\scriptscriptstyle b}$ en la implementación del concepto de solución en la línea propuesta por Emond y Mason (2000, 2002), de maximizar la correlación media ponderada, parece inevitable.

Naturalmente, con esta definición matricial de los *rankings*, (A10), y del coeficiente de correlación de rangos, (A11), el problema de optimización

$$\underset{S \in \Omega}{Max} \frac{\sum_{k=1}^{m} \omega_{k} \tau_{x}(\boldsymbol{S}, \boldsymbol{G}^{k})}{\sum_{k=1}^{m} \omega_{k}}$$
 (A12)

donde G^k es la representación matricial del ranking individual k definida de acuerdo con (A10), se simplifica como

$$\max_{S \in \Omega} \frac{\sum_{k=1}^{m} \omega_k \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij} g_{ij}^k \right\}}{n.(n-1)}$$
 (A13)

El denominador es irrelevante, e intercambiando los sumatorios podemos simplificar (A13) como

$$\begin{array}{ll}
\operatorname{Max} \sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij} \sum_{k=1}^{m} \omega_{k} g_{ij}^{k} \\
\Leftrightarrow \operatorname{Max} \sum_{s=0}^{n} \sum_{j=1}^{n} S_{ij} h_{ij}
\end{array} \tag{A14}$$

donde $h_{ij} = \sum_{k=1}^{m} \omega_k g_{ij}^k$. Este problema tiene de nuevo la misma estructura que (4), solo que las matrices que representan los *rankings* individuales han sido construidas de distinta forma.

Sin embargo en este caso no existe una relación uno-a-uno entre la representación (A2) y la representación (A10). La **conclusión** es que, **en el caso general, maximizar la correla-** ción media ponderada, medida por τ_x , con estos m rankings individuales no es equivalente a encontrar el ranking óptimo o de consenso definido como aquel ranking que tiene el mayor soporte en todas las comparaciones bilaterales. Es decir, es no indiferente resolver (A14) o resolver (4).

Para que se diera dicha equivalencia el *ranking* **A** debería representarse como

$$a_{ij} = \begin{cases} 1 & \text{si el objeto } i \text{ precede en el ranking al objeto } j \\ -1 & \text{si el objeto } i \text{ sucede en el ranking al objeto } j \\ 0.5 & \text{si los objetos } i \text{ y } j \text{ están empatados} \\ 0 & \text{si } i = j \text{ o los objetos no son comparables} \end{cases}$$

(A15)

y utilizar τ_x como coeficiente de correlación. Sin embargo en este caso, incluso si el ranking es completo, el denominador de τ_b no es n.(n-1), y τ_x será menor que uno para la correlación de un ranking consigo mismo en el caso de empates o rankings incompletos. Es estos casos deberíamos utilizar τ_b y arrastrar el factor de normalización en (A5) para resolver un problema de optimización del tipo (A6), que no se simplifica en un problema de optimización del tipo (A9)/(A14).

Así pues, en el caso general no es equivalente el concepto de solución de Emond y Mason (2000, 2002), que se basa en encontrar el ranking que maximiza la correlación media ponderada con los m rankings individuales, ya sea utilizando τ_x o τ_b , y cualquiera de las representaciones matriciales de los rankings propuestas, (A4), (A10) o (A15), con el expuesto en el apartado 4 que define el ranking óptimo o de consenso como aquel que tiene el mayor soporte en todas las comparaciones bilaterales a partir de la matriz de Condorcet o outranking matrix. 36

Resulta interesante recordar que, en el caso de *rankings* completos y sin empates, todas las aproximaciones son equivalentes.

Desde el **punto de vista práctico** encontramos mucho **más intuitivo trabajar con la matriz de Condorcet o** *outranking matrix*, tal y como ha sido introducida en los apartados 3 y 4, que con el coeficiente de correlación de

³⁶ No obstante, para un *ranking* individual, una equivalencia uno-a-uno entre la representación en términos de la *outranking matrix* y la representación en términos de Kendall (1948) solo se da cuando dichas representaciones se establecen mediante (A4) en el primer caso y (A15) en el segundo.

Kendall (1948) en cualquiera de sus versiones. Esta es la aproximación seguida por la teoría de la elección social (Arrow y Raynaud 1986), por la literatura de la evaluación multicriterio (Munda 2008) y también por la literatura matemática de los órdenes lineales (Martí y Reinelt 2011), que sin embargo solo parece considerar rankings completos y sin empates. De esta forma no dependemos de ninguna normalización para la simplificación del problema.

Desde el punto de vista computacional ello implica que, en el caso general, no podemos trabajar con la librería ConsRank (D'Ambrosio, Amodio y Mazzeo 2019) de R, basada en la aproximación de Emond y Mason (2002) y que incluye métodos exactos -Branch-and-Boundy heurísticos (D'Ambrosio 2007; D'Ambrosio, Amodio e Iorio 2015; Amodio, D'Ambrosio y Siciliano 2016; D'Ambrosio et al. 2017; D'Ambrosio, Amodio y Mazzeo 2019). Para que ambos métodos fueran equivalentes sería necesario representar los rankings individuales como (A15), en lugar de como (A10), lo que además estaría sujeto a la cualificación de que el denominador del coeficiente de correlación no es n.(n-1), tanto cuando existen empates como casos no comparables, y por tanto el concepto de solución implementado no se correspondería con el de maximizar la correlación media ponderada con los m rankings individua-

ALGORITMO DE *BRANCH-AND-BOUND* (RAMA Y PODA)

Describimos a continuación el algoritmo de *Branch-and-Bound* (Rama y Poda), y que constituye una variante del desarrollado por Emond y Mason (2000, 2002) e implementado en *R*, librería *ConsRank*, por D'Ambrosio, Amodio y Mazzeo (2019).

El problema viene definido por (4). El primer paso consiste en determinar una cota superior al producto escalar de las matrices \mathbf{S} y \mathbf{A} . Para una matriz $n \times n$, el producto escalar máximo de la representación matricial de un *ranking* débil de n objetos, y que consiste solo en valores 0, 0.5 y 1, \mathbf{S} , y la *outranking matrix*, \mathbf{A} , viene dado por la siguiente suma, a la que llamamos $\mathbf{P}^{\mathbf{x}}$,

$$P^* = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \max \left\{ c_{ij}, c_{ji} \right\}$$
 (A16)

Supongamos que disponemos de un *ranking* débil de *n* objetos que constituye nuestra solución inicial, para el que calculamos su represen-

tación matricial, $\{s_{ij}\}$. Dada esta **solución inicial**, evaluamos la expresión (4)

$$p = \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij} c_{ij} = \sum_{k=1}^{m} \omega_k \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij} a_{ij}^k \right\}$$
 (A17)

Y definimos la **penalización inicial**, P, substrayendo este valor, p, de P*.

$$P = P^* - p \tag{A18}$$

Este es el punto de inicio de nuestro algoritmo de búsqueda.

Obsérvese que la **penalización** *P* está restringida a ser **no-negativa** y que una **penalización de cero significa que el** *ranking* **en cuestión es una solución**. Así pues, si nuestra solución inicial arroja una penalización inicial de cero significa que hemos encontrado una solución —no necesariamente única—. En general, sin embargo, el valor de *P* no será cero. **El problema es buscar el conjunto de** *rankings* débiles de *n* objetos para encontrar la penalización mínima.

Algoritmo de búsqueda.

Sea la penalización corriente mínima igual a la penalización inicial. Considera los dos primeros objetos representados en el *ranking* débil de la solución inicial. Hay solo tres posiciones relativas posibles entre estos dos objetos: ambos están empatados o uno de ellos precede al otro. Dividimos el conjunto de todos los *rankings* débiles posibles de *n* objetos en **tres** subconjuntos mutuamente excluyentes o **ramas** basándonos en la posición relativa de estos dos primeros objetos. Supongamos que los dos objetos en cuestión son etiquetados como i y j. Consideramos los correspondientes elementos c_{ij} y c_{ji} de la *outranking matrix*, $c_{ii} = \sum_{k=1}^{m} \omega_k a_{ii}^k$.

Para cada una de estas ramas calculamos la **penalización incremental** de la siguiente forma.

- Rama 1: Objeto / precede al objeto /:
 (/ ⑤ /).
 - o Si $c_{ij} \ge c_{ji}$ entonces la penalización incremental es cero.
 - O Si $c_{ij} < c_{ji}$ entonces la penalización incremental es $c_{ji} c_{jj}$.
- Rama 2: Objeto j está empatado con el objeto j: (j ~ j).³⁷

³⁷ Esta rama no es necesaria si solo buscamos en el espacio de rankings completos (D'Ambrosio, Amodio e Iorio 2015).

- Si $c_{ij} > c_{ji}$ entonces la penalización incremental es $0.5 \times (c_{ij} c_{ji})$.
- o Si $c_{ij} = c_{ji}$ entonces la penalización incremental es cero.
- o Si $c_{ij} < c_{ji}$ entonces la penalización incremental es $0.5 \times (c_{ij} c_{ij})$.
- Rama 3: Objeto *i* sucede al objeto *j*: (*j* ⑤ *i*).
 - o Si $c_{ij} > c_{ji}$ entonces la penalización incremental es $c_{ij} c_{ji}$.
 - o Si $c_{ij} \leq c_{jir}$ entonces la penalización incremental es cero.

Para cada rama tenemos una penalización incremental. Si la penalización incremental para una rama es mayor que la penalización inicial, entonces podemos dejar de considerarla porque todos los *rankings* débiles de esa rama tendrán una penalización total más grande que la penalización inicial, y en consecuencia no pueden ser soluciones.

Por el contrario, supongamos que la penalización incremental de una rama es menor o igual a la penalización inicial. Entonces fijamos la penalización corriente acumulada para esa rama igual a la penalización incremental.

Ahora consideramos el siguiente objeto en la solución inicial. Generamos nuevas ramas a partir de la rama padre situando el siguiente objeto del ranking en todas las posiciones relativas posibles respecto a los objetos ya existentes, incluyendo todos los posibles empates. Para cada nueva rama generada de esta forma, calculamos las penalizaciones incrementales entre el nuevo objeto y todos los objetos ya existentes de la misma forma que anteriormente, a partir de los correspondientes elementos de la outranking matrix. Las nuevas penalizaciones son sumadas y añadidas a la penalización acumulada existente para la rama padre, obteniendo de esta forma la penalización acumulada actual para la nueva rama.

De nuevo, comprobamos la penalización acumulada actual para cada rama frente a la penalización actual mínima, abandonando cualquier rama cuya penalización acumulada sea mayor que el mínimo actual.

Continuamos de esta forma hasta que todas las ramas se extinguen o hasta que una rama se sigue hasta el final, y en consecuencia es una solución. En este punto habremos encontrado un *ranking* débil de *n* objetos cuya penalización total es menor o igual que la penalización actual mínima. La mejor penalización actual es

entonces actualizada y continuamos de esta forma hasta que el espacio entero de soluciones ha sido explorado, garantizando de esta forma que todas las soluciones con penalización mínima han sido encontradas.

La descripción del algoritmo permite intuir que la velocidad de convergencia depende en gran medida de la solución inicial, (A17), a partir de la cual generamos la penalización inicial, (A18). Cuanto menor sea esta penalización inicial más rápidamente progresa el algoritmo a través del espacio de soluciones posibles (Emond y Mason 2000, p. 23). En la práctica las soluciones iniciales propuestas por Amodio, D'Ambrosio y Siciliano (2016, p. 670) o la proporcionada por la regla de Copeland generaron excelentes soluciones iniciales.

CONTRAEJEMPLO: INDICADOR CON UN PESO SUPERIOR AL 50% QUE NO RESULTA SER EL INDICADOR ÓPTIMO

Al introducir los pesos en el apartado 3.2 se indica que, en general, es esencial que ninguno de los indicadores individuales tenga un peso asignado superior al 50%, ya que de otra forma el *ranking* óptimo coincidirá con dicho indicador individual (Munda y Nardo 2009, p. 1516, nota al pie 5).

Aunque este argumento es intuitivo en el caso de la agregación no compensatoria, puesto que dar a un indicador individual un peso superior al 50% es equivalente a otorgarle la mayoría, ofrecemos a continuación un ejemplo en el que esto no sucede, y un indicador con ²/₃ del peso, no resulta ser el indicador óptimo.

Considera que tenemos 2 objetos a ordenar, (A, B), y disponemos de 3 *rankings* individuales: (i) $\langle A - B \rangle$, (ii) $\langle A - B \rangle$ y (iii) $\langle A, B \rangle$, o alternativamente el *ranking* $\langle A - B \rangle$ tiene un peso de $^{2}/_{3}$ y el *ranking* $\langle A, B \rangle$ tiene un peso de $^{1}/_{3}$. La matriz de Condorcet para este caso es

$$\begin{array}{ccc}
A & B \\
A & \begin{bmatrix} 0 & 2 \\
1 & 0 \end{bmatrix}
\end{array}$$

Con 2 objetos a ordenar existen 3 *rankings* posibles: $\langle A, B \rangle$, $\langle B, A \rangle$ y $\langle A-B \rangle$.

El $ranking \langle A, B \rangle$ recibe un apoyo por parte de nuestra información de 2.

El $ranking \langle B, A \rangle$ recibe un apoyo por parte de nuestra información de 1.

El $ranking \langle A-B \rangle$ recibe un apoyo por parte de nuestra información de 1.5.

En consecuencia el *ranking* óptimo o de consenso, de acuerdo con el criterio de CKYL es el *ranking* $\langle A, B \rangle$, que solo tiene un peso de 1/3.

Los valores de este ejemplo están tomados de Emond y Mason (2000, p. 4), aunque estos autores no hacen la observación de que el *ranking* individual con un peso superior al 50% no resulte ser el *ranking* óptimo.

BIBLIOGRAFÍA

AILON, N. (2010). «Aggregation of Partial Rankings, *p*-Ratings and Top-*m* Lists». *Algorithmica* 57: 28-300.

ALI, A. y M. MEILĂ (2012). «Experiments with Kemeny ranking: What works when?». *Mathematical Social Sciences* 64, n.º 1 (julio): 28-40.

AMODIO, S., A. D'AMBROSIO y R. SICILIANO (2016). «Accurate algorithms for identifying the median ranking when dealing with weak and partial rankings under the Kemeny axiomatic approach». *European Journal of Operational Research* 249, n.º 2 (marzo): 667-676.

ARROW, K. J. (1963). *Social choice and individual values*. 2^a edición. Nueva York: Wiley.

ARROW, K. J. y H. RAYNAUD (1986). *Social choice* and multicriterion decision making. Cambridge (EE. UU.): MIT Press.

ATHANASSOGLOU, S. (2015). «Multidimensional welfare rankings under weight imprecision: a social choice perspective». *Social Choice and Welfare* 44: 719-744.

AZZINI, I. y G. MUNDA (2020). «A new approach for identifying the Kemeny median ranking». *European Journal of Operational Research* 281, n.º 2 (marzo): 38-401.

Bartholdi, J., C. A. Tovey y M. A. Trick (1989). «Voting schemes for which it can be difficult to tell who won the election». *Social Choice and Welfare* 6, n.º 2: 157-165.

BORDA, J. C. (1781). *Mémoire sur les élections au scrutiny*. París: Histoire de l'Academie Royale des Sciences.

CHENERY, H. B. y T. WATANABE (1958). «International comparisons of the structure of production». *Econometrica* 26, 4 (octubre): 487-521.

CONDORCET, J. A. N. de Caritat (1785). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la probabilité des voix. París: Imprimerie Royale.

COOK, W. D., M. KRESS y L. M. SEIFORD (1986). «An axiomatic approach to distance on partial orderings». *RAIRO - Operational Research* 20, n.° 2: 115-122.

D'AMBROSIO, A. (2007). *Tree Based Methods for Data Editing and Preference Rankings* (Tesis doctoral en estadística). Nápoles (Italia): Uni-

versità degli Studi di Napoli Federico II, Dipartimento di Matematica e Statistica.

D'AMBROSIO, A., G. MAZZEO, C. IORIO y R. SICILIANO (2017). «A differential evolution algorithm for finding the median ranking under the Kemeny axiomatic approach». *Computers and Operations Research* 82 (junio): 126-138.

D'AMBROSIO, A., S. AMODIO y C. IORIO (2015). «Two algorithms for finding optimal solutions of the Kemeny rank aggregation problem for full rankings». *Electronic Journal of Applied Statistical Analysis* 8, n.º 2: 198-213.

D'AMBROSIO, A., S. AMODIO y G. MAZZEO (2019). «Package 'ConsRank': Compute the Median Ranking(s) According to the Kemeny's Axiomatic Approach». Versión 2.1.0. 2019-12-05. Disponible en:

https://cran.r-

project.org/web/packages/ConsRank/index.html

DECANCQ, K. y M. A. Lugo (2013). «Weights in multidimensional indices of wellbeing: An overview». *Econometric Reviews* 32, n.º 1: 7-34.

EMERSON, P. (2013). «The original Borda count and partial voting». *Social Choice and Welfare* 40, n.º 2: 353-358.

EMOND, E. J. (1997). «Maximum rank correlation as a solution concept in the *M* rankings problem with applications to multi criteria decision analysis». *DOR (CAM) Research Note RN* 9705. Ottawa (Canadá): Department of National Defence.

EMOND, E. J. y D. W. MASON (2000). *A new technique for high level decision support. Operational Research Division.* Ottawa (Canadá): Department of National Defence Canada, Operational Research Division, Directorate of Operational Research (Corporate, Air & Maritime) (ORD Project Report PR2000/13).

EMOND, E. J. y D. W. MASON (2002) «A new rank correlation coefficient with application to the continuous ranking problem». *Journal of Multi-Criteria Decision Analysis* 11, n.º 1: 17-28.

FAGIN, R., R. KUMAR y D. SIVAKUMAR (2003). «Comparing top *k* lists». *SIAM Journal of Discrete. Mathemathics* 17, n.º 1: 134-160.

Fagin, R., R. Kumar, M. Mahdian, D. Sivakumar y E. Vee (2006). «Comparing partial rankings».

- SIAM Journal of Discrete Mathematics 20, n.º 3: 628-648.
- FISHBURN, P. C. (1977). «Condorcet social choice functions». *SIAM Journal of Applied Mathemathics* 33, n.º 3: 469-489.
- FISHBURN, P. C., W. V. GEHRLEIN y E. MASKIN (1979). «Condorcet's proportions and Kelly's conjecture». *Discrete Applied Mathematics* 1: 229-252.
- GOERLICH, F. J. (2019). «Rankings ¿Podemos aprender algo de la teoría de la elección social?». Ponencia presentada en *Seminario: Sistemas de información universitaria: Rankings e indicadores.* València: Ivie, 4 de junio de 2019. Disponible en: https://www.ivie.es/wpcontent/uploads/2019/05/Francisco-Goerlich.pdf.
- GOERLICH, F. J. y A. VILLAR (2009). *Desigualdad y Bienestar Social. De la teoría a la práctica*. Bilbao: Fundación BBVA.
- GROSS, O. A. (1962). «Preferential Arrangements». *The American Mathematical Monthly* 69, n.º 1: 4-8.
- HERRERO, C. (2016). «An Endogeneous scoring rule». Working Papers Serie AD n.º 2016-05. València: Ivie (Instituto Valenciano de Investigaciones Económicas).
- HERRERO, C. y A. VILLAR (2019). «The Borda-Condorcet Social Evaluation Function». Working Papers n.º 19.02. Sevilla: Universidad Pablo de Olavide, Departamento de Economía.
- Kemeny, J. G. (1959). «Mathematics without numbers». *Daedalus* 88, n.º 4: 571-591.
- KENDALL, M. G (1948). *Rank correlation methods*. Londres: Charles Griffin and Company Limited.
- MAGGINO, F. (ed.) (2017). *Complexity in Society: From Indicators Construction to their Synthesis.* Berlín: Springer-Verlag (Social Indicators Research Series n.º 70).
- MARTÍ, R. y G. REINELT (2011). *The Linear Ordering Problem. Exact and Heuristic Methods in Combinatorial Optimization.* Heidelberg (Alemania): Nueva York: Springer (Applied Mathematical Sciences n.º 175).
- MAZZIOTTA, M. y A. PARETO (2016a). «On a generalized non-compensatory composite index for measuring socio-economic phenomena». *Social Indicators Research* 127, n.º 3: 983-1003.
- __ (2016b). «Methods for constructing non-compensatory composite indices: A comparative study». *Forum for Social Economics* 45, n.º 2-3: 213-229.

- __ (2018). «Measuring Well-Being Over Time: The Adjusted Mazziotta-Pareto Index Versus Other Non-compensatory Indices». *Social Indicators Research* 136, n.º 3: 967-976.
- MOULIN, H. (1989). *Axioms of cooperative decision making*. Econometric Society Monographs. Cambridge: Cambridge University Press.
- Munda, G. (2008). *Social Multi-Criteria Evaluation for a Sustainable Economy*. Nueva York: Springer.
- __ (2012). «Choosing Aggregation Rules for Composite Indicators». *Social Indicators Research* 109, n.º 3: 337-354.
- __ (2015). «Beyond GDP: An overview of measurement issues in redefining 'wealth'». *Journal of Economic Surveys* 29, n.º 3: 403-422.
- MUNDA, G. y M. NARDO (2009). «Noncompensatory/nonlinear composite indicators for ranking countries: A defensible setting». *Applied Economics* 41, n.º 12: 1513-1523.
- MUNDA, G. y M. SAISANA (2011). «Methodological considerations on regional sustainability assessment based on multicriteria and sensitivity analysis». *Regional Studies* 45, n.º 2: 261-276.
- OCDE/UNIÓN EUROPEA/JRC (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. París: OECD Publishing.
- PARUOLO, P., M. SAISANA y A. SALTELLI (2013). «Ratings and rankings: Voodoo or science?». *Journal of the Royal Statistical Society. Series A: Statistics in Society* 176, n.º 3 (junio): 609-634.
- PÉREZ, F., J. ALDÁS-MANZANO (dirs.), R. ARAGÓN E I. ZAERA (2019). *U-Ranking 2019: Indicadores sintéticos de las universidades españolas.* 7ª edición. Bilbao: Fundación BBVA; València: Ivie
- SAISANA M., A. SALTELLI y S. TARANTOLA (2005). «Uncertainty and Sensitivity analysis techniques as tools for the quality assessment of composite indicators». *Journal of the Royal Statistical Society. Series A: Statistics in Society* 168, n.º 2 (marzo): 307-323.
- STIGLITZ, J.E., A. SEN y J. P. FITOUSSI (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. París: Commission on the Measurement of Economic Performance and Social Progress.
- VANSNICK, J. C. (1990). «Measurement theory and decision aid». En Carlos A. Bana e Costa (ed.). *Readings in multiple criteria decision aid*. Berlín: Nueva York: Springer-Verlag: 81-100.
- YOUNG, H. P. y A. LEVENGLICK (1978). «A consistent extension of Condorcet's election principle». *SIAM Journal of Applied Mathematics* 35, n.º 2: 285-300.

