
Escalamiento Multidimensional No Métrico. Un ejemplo con *R* empleando el algoritmo SMACOF

citation and similar papers at core.ac.uk

brought

provided by Dadun

An Example with R Using SMACOF Algorithm

EMELINA LÓPEZ-GONZÁLEZ

Departamento de Métodos de Investigación e
Innovación Educativa.
Universidad de Málaga.
emelopez@uma.es

RAMÓN HIDALGO SÁNCHEZ

Servicio Central de Informática
Universidad de Málaga.
ramon@uma.es

Resumen: En este artículo se describe el desarrollo histórico y formal del escalamiento multidimensional no métrico hasta llegar a una de las últimas aportaciones de optimización: el algoritmo SMACOF. El objetivo es animar al uso de esta solución, especialmente útil en el ámbito de la investigación educativa, dado el habitual empleo de variables no métricas (ordinales) que frecuentemente se desean escalar. SMACOF resuelve los inconvenientes de soluciones previas y además se encuentra implementado en el software libre *R*. A tal fin se presenta un ejemplo, mostrando las instrucciones necesarias e interpretando la salida.

Palabras clave: Escalamiento multidimensional no métrico, regresión monótona, SMACOF, *R*.

Abstract: The historical and formal development of the non-metric multidimensional scaling is exposed in the end reaching the latest contributions in optimization: the SMACOF algorithm. Our purpose is to encourage the use of this solution, which is particularly useful in the educational research setting, as non-metric variables (ordinal scales) are often used. SMACOF solves the problems from previous solutions and, also, it is found in free software *R*. For this reason, an example is included with the necessary instructions and with the interpretation of the results.

Keywords: Non-metric multidimensional scaling, monotonic regression, SMACOF, *R*.

INTRODUCCIÓN

El Escalamiento Multidimensional (EMD) comprende una amplia familia de procedimientos multivariados, algunos de los cuales hacen posible representar las proximidades entre un conjunto de elementos como distancias en un espacio de un número reducido de dimensiones. La idea central es que las distancias que median entre los puntos se corresponden con las proximidades entre los objetos por medio de una función de ajuste resultante de un proceso iterativo de optimización, pudiéndose describir las relaciones entre los objetos sobre la base de las proximidades observadas. De este modo se establece una analogía entre el concepto “psicológico” de proximidad (semejanza o desemejanza) y el concepto geométrico de distancia (Arce, 1996), quedando reflejadas las diferencias entre ambos en una función denominada *stress* (Kruskal, 1964a y b).

Una primera división entre las modalidades de EMD se relaciona con las características de la función de ajuste que se emplee entre proximidades y distancias, dando lugar al EMD métrico y al EMD no métrico u ordinal. El Escalamiento Multidimensional Métrico (EMD-M) utiliza una función mínimo cuadrática paramétrica, mientras que el Escalamiento Multidimensional No Métrico (EMD-NM) emplea una regresión monótona de mínimos cuadrados. La elección entre uno u otro depende especialmente de la calidad métrica de los datos de partida. Cuando se consiguen proximidades con valores numéricos bien determinados, habiendo usado instrumentos precisos que calculan las desemejanzas con magnitudes muy cercanas a las distancias reales, puede emplearse el EMD-M con función paramétrica. En ámbitos como las Ciencias del Comportamiento, Ciencias Políticas, Sociales y de la Educación, en general, donde intervienen datos de preferencia y juicios de apreciación humana muy probablemente sujetos a error, las desemejanzas deben ser interpretadas en sentido ordinal, y se asume que la función que relaciona proximidades y distancias ha de ser monótona, debiendo realizar un EMD-NM.

Además de la escala de medida, en un EMD son varias las consideraciones a tener en cuenta respecto a los datos de entrada, a saber: la forma de la matriz de datos (rectangular, cuadrada, triangular), el número de vías (dos o más), el número de modos; la condicionalidad de la medida, así como los diversos formatos en los que pueden estar recogidos los datos: matrices de proximidad, tablas de contingencia, matrices de correlaciones, datos de perfil, etc¹. Igualmente, según el tipo particular de geometría en la que se busca obtener la representación, la métrica de las distancias puede ser distinta: *city-block*, euclídea, dominada.

¹ Una breve exposición de las distintas opciones de EMD según los datos de entrada se encuentra en Arce (1993) y Real (2001).

Toda vez que se opte por un escalamiento adaptado a la medida que se dispone, es preciso emplear algún algoritmo para encontrar la representación óptima consistente en un proceso iterativo de cierta envergadura que tiene como objeto minimizar la función de *stress*, lo que permitirá, en última instancia, determinar la bondad del ajuste del escalamiento. Para este cometido existen también distintas opciones. Lo habitual en el EMD-NM ha sido emplear el proceso iterativo de Kruskal-Shepard implementado en el algoritmo ALSCAL (*Alternating Least Squares Scaling*). Ahora bien, el presente trabajo muestra otro de los procedimientos para optimizar (minimizar) el *stress* en el EMD-NM: el algoritmo SMACOF (*Scaling by MAyorizing a COmplicated Function*). Su desarrollo inicial se debe, fundamentalmente, a De Leeuw (De Leeuw, 1977 y De Leeuw y Heiser, 1977).

SMACOF es un algoritmo interesante desde el momento en que resuelve algunos de los inconvenientes de soluciones previas, como ALSCAL, con el empleo del algoritmo de mayorización iterativa, resultando ser una solución más sencilla y aplicable a muy diversas variantes del EMD. En relación con la oferta informática que lo incluye, SMACOF puede trabajarse con paquetes convencionales como SPSS o SAS, aunque en éstos no se encuentra implementado como un algoritmo específico, quedando limitada una completa accesibilidad al conjunto de opciones de salida que permite². Esto no sucede con el software *R*. Sin pretender extendernos en un análisis pormenorizado de las posibilidades y funciones de unos programas u otros, una notable ventaja a favor de *R* es que su modo de trabajar facilita interaccionar SMACOF con cualquier otro análisis que se desee en el momento mismo en que se ejecuta. Bien es cierto que *R* es todavía un software poco conocido en el ámbito de la investigación educativa (no así en otros entornos científicos), sin embargo, como se apunta en Ruiz-Soler y López-González (2009), convendría considerar también sus notables ventajas para la docencia y la investigación³.

El objetivo de este trabajo es animar al empleo de SMACOF en el EMD-NM. Para ello es importante comprender algunos de los desarrollos previos a los que pretende dar respuesta. Por esta razón, partiendo de la premisa de que el EMD-NM se ajusta a las características métricas de muchas de las variables que se emplean comúnmente en investigaciones educativas, escalas e ítems ordinales, como las escalas Likert, el artículo comienza señalando cómo surgen en el proceso de

² En SPSS, por ejemplo, debe realizarse desde el programa PROXSCAL (PROXimity SCALing), iniciado a partir del trabajo Busing, Commandeur y Heiser (1997), y explicado en Meulman, Heiser y SPSS Inc. (2008).

³ El lector poco familiarizado con *R* encontrará en el trabajo de Arriaga et al. (2008) un buen texto de iniciación. Sobre la ejecución de SMACOF con *R* puede consultarse De Leeuw y Mair (2008a, 2008b).

desarrollo del escalamiento general las peculiaridades de la opción no métrica. Continuamos con una descripción sencilla de la estrategia SMACOF en respuesta a los procesos internos de ALSCAL. Advertimos que no es nuestro objetivo profundizar en sus desarrollos matemáticos y formales, y no nos detendremos a explicar la mayorización iterativa (necesitaríamos un espacio y un nivel de profundidad formal mayores de lo que aquí se pretende). Nuestra intención es ayudar al usuario a encontrar y comprender las pautas para realizar con R un EMD-NM empleando SMACOF, por lo que terminamos con un sencillo ejemplo ilustrando los pasos más elementales e interpretando sustantivamente la salida.

REFERENTES HISTÓRICOS DEL EMD-NM

No es fácil dar una definición precisa del Escalamiento Multidimensional (EMD). A juicio de De Leeuw y Heiser (1982), el EMD puede entenderse en sentido amplio, comprendiendo varias formas de análisis *cluster* y análisis multivariante lineal y, en sentido estricto, representando datos de proximidad en un espacio de pocas dimensiones, concepción también más próxima a la definición de Kruskal (1977, p. 296): “cualquier método para construir una configuración de puntos en un espacio de baja dimensionalidad, a partir de distancias medidas con error aleatorio o de la información del orden de los rangos de las mismas.”

El EMD es una técnica con diversos propósitos, desde explorar proximidades, estructuras y procesos psicológicos, hasta incluso probar hipótesis experimentales. En su modalidad no métrica, además, es aplicable a innumerables campos de las ciencias humanas en los que se juzguen preferencias de diverso tipo (legales, políticas, sociales, lingüísticas, comerciales, culturales).

Tomando el EMD en sentido estricto, la primera exposición completa de su método se encuentra en la obra de Torgerson (1952), quien generalizó las anteriores versiones unidimensionales existentes hasta el momento y describió el primer proceso métrico sistemático para obtener una representación multidimensional de puntos a partir de distancias medidas “con error”. Paralelamente, también en la década de 1950, Coombs comienza a desarrollar una teoría de datos que publica más tarde (1964) con una notable influencia. En esa publicación señala (p. 444) a sus colaboradores Bennett y Hays (1960) como precursores del EMD no métrico. El método desarrollado por éstos, a juicio de Kruskal y Wish (1978, p. 23), era no métrico en dos sentidos: a) la fuerte condición métrica del procedimiento de Torgerson quedaba sustituida por la condición de una relación ordinal entre proximidades y distancias, y b) únicamente era preciso determinar el orden del rango del objeto en cada dimensión.

Al tiempo que el trabajo de Coombs, cabe destacar en el EMD-NM la aportación de Guttman (1944, 1946), aunque el progreso más notable no se produce hasta Shepard (1962). Shepard centra por primera vez la atención en las relaciones ordinales entre proximidades y distancias, proporcionando el primer programa de EMD-NM generalizado aplicable a distancias definidas de forma no métrica. El supuesto de partida es la monotonía de la función entre las proximidades iniciales y las distancias de la configuración resultante, estableciéndose como requisitos: a) un número mínimo de dimensiones del espacio euclídeo necesario para que las distancias estén monótonamente relacionadas con las proximidades, b) un conjunto de coordenadas ortogonales para los puntos en ese espacio, y c) un diagrama que represente la verdadera función inicialmente desconocida, aunque supuestamente monótona, que relacione proximidades y distancias (el diagrama de Shepard).

El trabajo de Shepard despierta un notable interés en Kruskal, quien escribe dos imprescindibles artículos (1964a, 1964b). Kruskal reorienta la aportación previa de Shepard, considerando ahora el EMD-NM más bien como un problema de ajuste estadístico. Ambos coinciden en adoptar como objetivo principal obtener una relación monótona entre proximidades y distancias, así como en proponer una medida de desviación de la monotonía que detectara el punto final del proceso iterativo para conseguir la solución adecuada. No obstante, mientras que Shepard dio a esta medida una importancia secundaria, Kruskal propone realizar primero la regresión monótona de las distancias sobre las medidas de proximidad y usar después la varianza residual, convenientemente normalizada, como medida cuantitativa de no monotonía, el *stress*, estableciendo como solución la configuración de puntos con mejor ajuste, es decir, aquella que minimiza el *stress*. Para determinar tal configuración utilizó el “método del gradiente” o “método del descenso profundo”.

En relación con los conceptos de algoritmo y la introducción de una terminología y una notación útiles, cabe mencionar el artículo de Guttman (1968) como un ambicioso intento de recoger todos los aspectos del EMD-NM desarrollados hasta el momento.

Lo que sigue en el tiempo son numerosos trabajos de simulación sobre las propiedades estadísticas del EMD-NM que mayormente giran en torno a tres preocupaciones (pueden verse en Martín-Casado y Galindo, 1994): el grado de definición métrica que proporciona el EMD-NM, el valor esperado de la medida de bondad del ajuste de la función monótona, y el número de dimensiones adecuado. Muchos de estos desarrollos terminan consolidándose en el enfoque unificado que presentan Takane, Young y De Leeuw (1977), quienes aportan el primer algoritmo capaz de realizar análisis métrico y no métrico, según el modelo de EMD euclídeo ponderado o simple, y con una o varias matrices de entrada. El procedimiento,

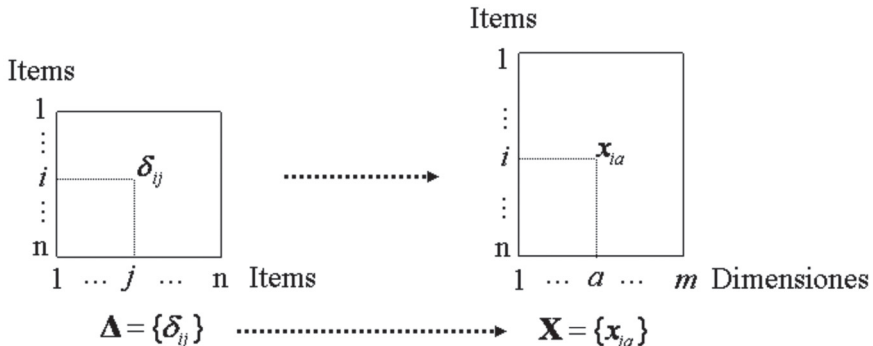
ALSCAL, perfeccionó el ajuste del modelo a los datos mediante una variante del método de ajuste de mínimos cuadrados convergente, rápida y relativamente libre del problema de los mínimos locales (Martín-Casado y Galindo, 1994).

DESARROLLOS FORMALES EN EL ESCALAMIENTO MULTIDIMENSIONAL NO MÉTRICO

Aparentemente el objetivo del EMD es sencillo y, sin embargo, encierra un proceso laborioso que puede ser interesante revisar en lo que respecta a la modalidad no métrica u ordinal. Hacemos sólo una modesta aproximación a la lógica interna que subyace en él. Sugerimos al lector que desee profundizar en su desarrollo matemático, los excelentes textos de Cox y Cox (2001, cap. 3) y Borg y Groenen (2005, cap. 9).

Supongamos que los datos de partida constituyen un conjunto Ω de n elementos⁴ (sujetos, objetos, variables o ítems) y una matriz de entrada Δ , cuadrada y de dos vías (ítems x ítems), cuyos elementos son el resultado de generar las proximidades⁵ para cualquier par de ítems (i, j) : $\Delta = (\delta_{ij})_n$. Esta matriz, además, es simétrica, luego $\delta_{ij} = \delta_{ji}$.

Figura 1. Ejemplo del objetivo del EMD



⁴ Dado que el ejemplo que se presenta al final del trabajo se aplica a los elementos de un cuestionario, en adelante nos referimos a los elementos sobre los que se crean las proximidades como ítems.

⁵ La proximidad entre dos ítems puede producirse como semejanza (similaridad) o como desemejanza (disimilaridades). Aquí se toman las proximidades como desemejanzas. Distintas medidas de proximidad pueden consultarse en Cox y Cox (2001).

El objetivo del EMD es, en general, conseguir una matriz de salida de coordenadas entre los ítems $X = (x_{ia})_m$ en el espacio R^m , matriz rectangular para el caso que seguimos, donde las filas representan los ítems ($i = 1, \dots, n$) y las columnas están referidas a las dimensiones del espacio ($a = 1, \dots, m$). Se tratará, por tanto, de estimar los valores x_{ia} (Figura 1).

Conocidas las coordenadas entre dos ítems (i, j) es posible calcular su distancia. Una expresión general es la distancia de Minkowsky que viene dada por:

$$d_{ij} = \left[\sum_{a=1}^m |x_{ia} - x_{ja}|^\lambda \right]^{\frac{1}{\lambda}},$$

donde $\lambda > 0$. Dando valores a λ se obtienen casos particulares en distintas geometrías. Si $\lambda = 1$, por ejemplo, se trabaja con la métrica de *city-bloc*; si $\lambda = \infty$, con la distancia dominada. Si hacemos $\lambda = 2$ obtenemos la distancia euclídea⁶ entre los elementos (i, j):

$$d_{ij} = \left[\sum_{a=1}^m (x_{ia} - x_{ja})^2 \right]^{\frac{1}{2}}.$$

La distancia euclídea cumple, entre otras, las siguientes propiedades:

- a. Simetría: $d_{ij} = d_{ji}$.
- b. No negatividad: $d_{ij} = d_{jj} = 0 \leq d_{ij}$.
- c. $d_{ij} = 0$.

Como ha quedado dicho, en el EMD-NM las proximidades de entrada suelen ser ordinales o bien provenir de juicios personales sujetos a error. Otra importante distinción es que en el EMD-M la función que relaciona las desemejanzas δ_{ij} entre ítems y las distancias d_{ij} entre puntos, $\delta_{ij} = f(d_{ij})$, es lineal y con pendiente positiva. Igualmente, en la opción métrica es necesario que las distancias euclídeas cumplan otros dos axiomas más:

⁶ En un espacio cartesiano, la distancia entre dos puntos i y j es el segmento que conecta ambos puntos y que viene determinado por la fórmula: $d_{ij}(X) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2}$. Así, $d_{ij}(X)$ es la raíz cuadrada de la suma de la diferencias entre las coordenadas para cada dimensión del espacio, o simplemente la aplicación del teorema de Pitágoras para la hipotenusa del ángulo recto que se forma con los puntos.

d. $d_{ij} = 0$ sólo si $i = j$.

e. Desigualdad triangular: $d_{ij} \leq d_{ik} + d_{kj}$.

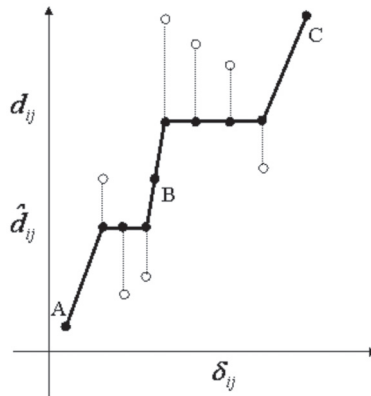
La solución para el EMD-M que dio Torgerson (1952) buscaba cumplir las propiedades anteriores, siendo una de las dificultades más notables garantizar la desigualdad triangular, lo que se conoce como el “problema de la constante aditiva” (ver, por ejemplo, Cailliez, 1983 y Cox y Cox, 2001). Otro de los aspectos de difícil solución es saber cuál es el número m de dimensiones adecuado.

Así como ambos escalamientos, métrico y no métrico, coinciden en un mismo objetivo (conocida la matriz de proximidades Δ estimar la matriz de coordenadas X —Figura 1—), el procedimiento en el caso del EMD-NM es diferente, ya que el problema con la desigualdad triangular no se produce. En el EMD-NM el interés está en obtener unas distancias que se relacionen monótonamente con las desemejanzas δ_{ij} por medio de la función: $d_{ij} = f(\delta_{ij})$ (Shepard, 1962). El método contempla una serie de momentos que culminan con el proceso iterativo de Kruskal-Shepard (Krusal, 1964b):

- a. Se convierten las proximidades en rangos, asignando a la desemejanza menor el rango 1; a la siguiente el 2, hasta llegar a la mayor que tendrá el rango $n(n-1)/2$.
- b. Se sitúan los ítems en un espacio de m dimensiones, lo que supone determinar la *matriz de configuración inicial*, X ($n \times m$). Éste es uno de los momentos más importantes en el EMD-NM, y no de fácil solución. Hay diversos procedimientos. Algunos autores hablan de colocar los ítems en el espacio de forma aleatoria, o sea, determinar la matriz al azar (Arce, 1993 y Rivas, 2001). También puede obtenerse como resultado de un proceso de Poisson de m dimensiones (Cox y Cox, 2001), o con la transformación de Guttman (1968, cit. en De Leeuw y Heiser, 1980), o bien emplear la denominada “configuración racional” derivada de algunas variantes del escalamiento métrico (pueden consultarse en Borg y Groenen, 2005).
- c. Definida una primera matriz de configuración inicial por un procedimiento u otro, se calculan las distancias euclídeas d_{ij} entre los puntos de manera tal que el rango de las distancias entre los estímulos (i, j) coincida con el rango asociado a la proximidad δ_{ij} correspondiente. Es decir: si $\delta_{ij} < \delta_{kl}$, entonces $d_{ij} \leq d_{kl}$. Esto no siempre sucede, por lo que se precisa obtener otras distancias que representen a las primeras y que aseguren la similitud en la

ordenación de rangos. Las nuevas distancias son las *disparidades* \hat{d}_{ij} —pseudo-distancias (Kruskal, 1977; Heiser, 1989), distancias objeto (*target distances*) o *dbats* (Borg y Groenen, 2005, p. 43)—, obtenidas como resultado de una *regresión monótona* (también llamada regresión isótona), cumpliendo, ahora sí, tener el mismo orden de rangos que las desemejanzas correspondientes: si $\delta_{ij} < \delta_{kl}$, entonces $\hat{d}_{ij} \leq \hat{d}_{kl}$.

Figura 2. Diagrama de Shepard en el que puede apreciarse la regresión monótona



En la Figura 2 vemos la regresión monótona⁷ de d_{ij} sobre δ_{ij} reflejada en el diagrama de Shepard: $d_{ij} = f(\delta_{ij})$. Los puntos huecos son los pares de proximidades y distancias (δ_{ij}, d_{ij}) y los puntos llenos los pares de proximidades y disparidades $(\delta_{ij}, \hat{d}_{ij})$. En cierto sentido puede decirse que las disparidades son una versión “suavizada” de las desemejanzas (Heiser, 1989). En la Figura 2 hay también algunos puntos en los que la monotonía se cumple, no siendo preciso su suavizado; así sucede con A, B y C. Finalmente, las líneas verticales señalan los “errores de representación” entre distancias y disparidades: $e_{ij} = [f(\delta_{ij}) - \hat{d}_{ij}]$; $e_{ij} = (d_{ij} - \hat{d}_{ij})$, siendo el error de representación cuadrado: $e^2_{ij} = [f(\delta_{ij}) - \hat{d}_{ij}]^2$. Haciendo la suma cuadrática de estos errores para todos los pares (i, j) se obtiene una medida de “maldad” del ajuste o de no-mono-

⁷ Al trabajar con desemejanzas como medida de proximidad, la línea resultante de la regresión es monótona porque no decrece. Si empleáramos semejanzas, la regresión monótona resultaría una línea que siempre decrece o se mantiene.

tonía (la varianza residual de la regresión efectuada) denominada “*stress* bruto” (*raw stress*):

$$\sigma_r^2 = \sum_{i < j} [f(\delta_{ij}) - \hat{d}_{ij}]^2 = \sum_{i < j} [d_{ij} - \hat{d}_{ij}]^2$$

Esta función de *stress* bruto suele verse afectada por la variabilidad que provocan los cambios de escala de las desemejanzas. Para evitar la dependencia de escala se normaliza, e incluso pueden hacerse algunos otros cambios que dan lugar a diversas versiones del *stress*: *sstress* (Takane et al., 1977); *stress-1* y *stress-2*. Aquí baste indicar la expresión del *stress* normalizado como sigue (Kruskal, 1964 a y b; Cox y Cox, 2001, p. 64):

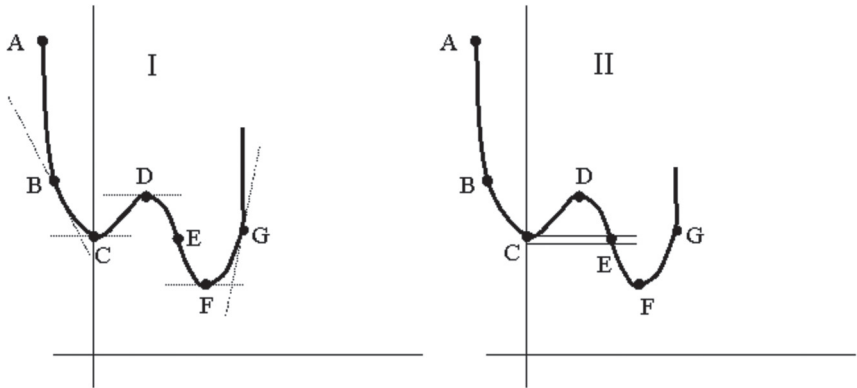
$$S = \sqrt{\frac{\sum_{i < j} (d_{ij} - \hat{d}_{ij})^2}{\sum_{i < j} d_{ij}^2}} = \sqrt{\frac{S^*}{T^*}} \quad \text{dónde} \quad S^* = \sum_{i < j} (d_{ij} - \hat{d}_{ij})^2 \quad \text{y} \quad T^* = \sum_{i < j} d_{ij}^2$$

- d. De los anteriores argumentos es fácil deducir el interés porque entre distancias y disparidades exista la menor discrepancia posible, de modo que el siguiente paso será minimizar la función de *stress* respecto a y (las coordenadas de la configuración), y también respecto a \hat{d}_{ij} .

No olvidemos que hasta ahora nos movemos en dos planos (Arce, 1996): uno empírico, con los ítems y las desemejanzas entre ítems, y otro teórico, con los puntos (coordenadas de X) y las distancias euclídeas entre ellos. Lo que se pretende es que el modelo de EMD-NM represente lo mejor posible los datos empíricos. Ahora bien, cabe esperar que el ajuste entre modelo y datos no sea bueno inicialmente (para empezar, las coordenadas de los puntos han podido obtenerse incluso aleatoriamente). Igualmente, el ajuste no se hace directamente con los datos originales comparando distancias y desemejanzas, sino que se comparan distancias con disparidades por medio, precisamente, de la función de *stress*. Las disparidades son, por tanto, un conjunto de valores que han de ser monótonos respecto de las proximidades y al mismo tiempo minimizar la suma de cuadrados $\sum_{i < j} (d_{ij} - \hat{d}_{ij})^2$.

Esta búsqueda de la mejor representatividad del modelo minimizando el *stress* se lleva a cabo por diversos procesos iterativos de aproximaciones sucesivas. Uno de ellos es el llamado “método de los gradientes” o “método del descenso profundo” que se emplea en ALSICAL (explicado en Cox y Cox, 2001).

Figura 3. Representación del método del descenso profundo y método del tunelado



En realidad se trata de ir variando ligeramente los valores de las coordenadas de la matriz X de modo que la diferencia entre las distancias y las disparidades vaya disminuyendo de un paso a otro, repitiéndose el proceso hasta que ninguna pequeña modificación suponga mejoría alguna. Formalmente implica sumar a la matriz de configuración inicial X una matriz de valores añadidos donde cada elemento de la misma contiene los valores adecuados para ser sumados a la coordenada del ítem i en la dimensión a . Estos valores se determinan mediante la derivada del *stress* con respecto a cada coordenada a ($\partial S / \partial x_{ia}$), proceso donde va surgiendo el problema de los mínimos local y global (Borg y Groenen, 2005, pp. 172, 276).

Observemos la Figura 3 (I y II). Si representamos la función de *stress* sobre una dimensión, ésta puede tener varios “valles” o mínimos donde la derivada de la función valga cero. En la Figura 3 (I) se aprecia, por ejemplo, que los puntos extremos C y F son unos mínimos de la función representada, siendo D un máximo. En estos puntos las rectas tangentes a la función son horizontales, sus pendientes valen cero y la derivada en estos puntos también, lo que no sucede en B, E y G, donde las tangentes presentan inclinación, siendo las derivadas en esos puntos distintas a cero. Aunque en alguna iteración aparezca un mínimo local, lo que interesa es encontrar el valor más bajo posible, es decir, un mínimo global, por lo que habitualmente es preciso intentarlo repetidas veces con distintas matrices de configuración inicial X para encontrar dicho mínimo. Cuando se consigue un valor de *stress* más bajo a igual dimensionalidad, entonces probablemente se haya alcanzado el mínimo global.

El “método del tunelado” es una variante del método del descenso profundo propuesta por Groenen y Heiser (1996). Miremos la Figura 3 (II). Se trata de “trasladarse” del punto C (que supone estar en un mínimo local, es decir, haber encontrado una configuración de la matriz X) al punto E trazando una especie de “túnel imaginario”. Una vez situados en el punto E será posible moverse y llegar al punto deseado, el punto F, que es un mínimo global. Esta operación se realiza por medio de una “función de tunelado” y el resultado supone haber encontrado con el mismo *stress* otra configuración de X que haga posible el mínimo global (Cox y Cox, 2001).

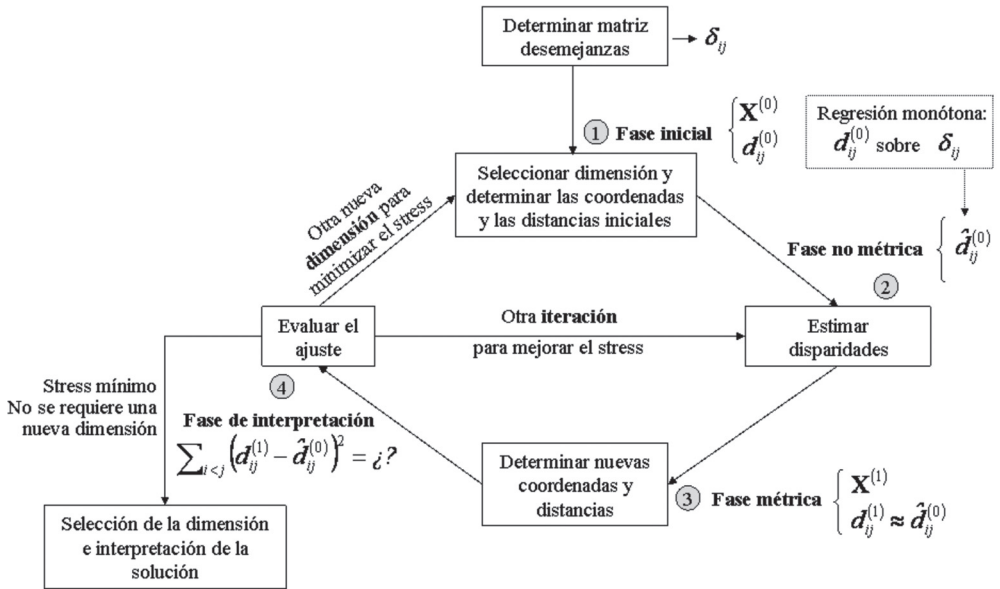
- e. Un problema de difícil solución es saber cuál es el número de dimensiones m adecuado para valorar la minimización del *stress*, y cómo combinar esta elección con la configuración inicial de la matriz X . Como regla general, se sabe que el *stress* disminuye a medida que m aumenta, aunque aplicando el método del tunelado los mínimos locales son más probables con pocas dimensiones y menos, o incluso ausentes, con un número alto de dimensiones.
- f. Superados los procesos descritos y llegando a este punto diremos que hemos dado una vuelta al procedimiento y, sin embargo, habrá que proceder una y otra vez hasta encontrar la solución adecuada, lo que constituye el algoritmo iterativo de Shepard-Kruskal (Figura 4).

El proceso iterativo descrito es el que se sigue en buena medida en ALSCAL y en SMACOF. Las diferencias entre ambos radican en el modo en que se minimiza el *stress*. ALSCAL emplea los procedimientos descritos (descenso profundo, tunelado), mientras que SMACOF implica la generalización del algoritmo de mayorización iterativa a espacios multidimensionales. No es objeto de este trabajo describir la mayorización iterativa. Diremos solamente que su característica fundamental es generar secuencias no crecientes de valores de la función de *stress* que terminan fácilmente en un mínimo local. El lector interesado puede consultar De Leeuw y Heiser (1977 y 1980) y Borg y Groenen (2005). Sí señalamos, no obstante, en las siguientes líneas, cuáles son las claves de la mejora de SMACOF respecto a ALSCAL.

SMACOF Y ALSCAL

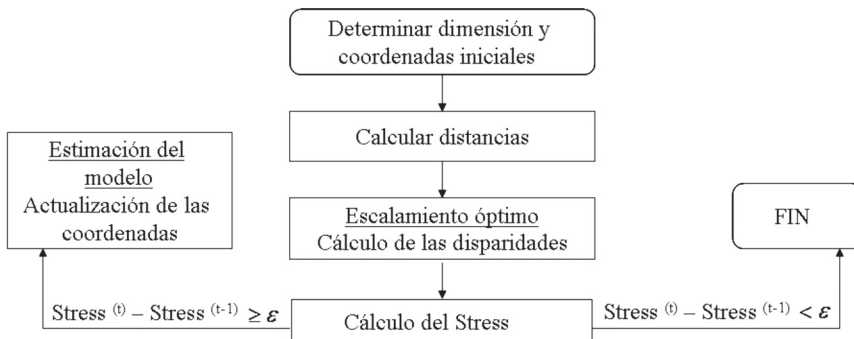
Por sus múltiples posibilidades, ALSCAL ha sido calificado por Cox y Cox (2001, p. 217) como una auténtica “caja de Pandora”. Permite analizar datos mediante un

Figura 4. Algoritmo Shepard-Kruskal para el EMD-NM (siguiendo a Jobson, 1992, p. 586)



modelo euclídeo simple o ponderado que puede ser: a) nominal, ordinal, de intervalo o de razón, b) simétrico y asimétrico, c) condicionado o incondicionado, d) viable con observaciones ausentes, e) con o sin replicaciones, etc. Es considerado incluso más eficaz con datos no métricos que con métricos dada la proximidad que se consigue en la obtención de mínimos globales.

Figura 5. Procedimiento general seguido por ALSICAL



Tiene dos fases que se van alternando: representación óptima y estimación del modelo (Figura 5). La fase de representación óptima (que se corresponde con las fases 1 y 2 señaladas en la Figura 4) transforma las observaciones según lo antes expuesto. La fase de estimación del modelo (la actualización de las coordenadas de los ítems) tiene una especial peculiaridad. Cada coordenada se estima por mínimos cuadrados de forma independiente mientras se mantienen fijos los restantes puntos y, sin embargo, la actualización de cada coordenada depende de los valores de todos los demás puntos. Por esta razón es necesario iterar una y otra vez hasta que se estabiliza la localización de todas las coordenadas del punto en cuestión en el espacio R^m . Cuando esto sucede, lo que en realidad ha ocurrido es que se ha movido dicho punto a la mejor localización posible pero manteniendo fijos el resto. El proceso continua sucesivamente para todos los demás puntos. Podemos hacernos una idea de la dificultad de convergencia del proceso, así como de su lentitud.

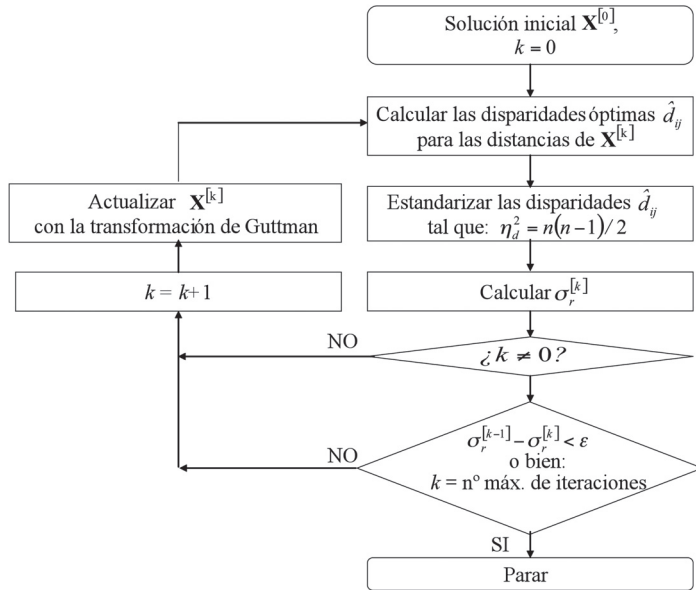
Por otro lado, ALSCAL, al igual que otros algoritmos del EMD-M, propone generar configuraciones en diferentes dimensiones y estudiar el comportamiento de cada configuración estimada. El problema está en que obtiene resultados de forma independiente, por lo que la configuración en la dimensión m no puede relacionarse con la obtenida en la dimensión $m+1$. Esto plantea la necesidad de un procedimiento que proporcione resultados en varias dimensiones de forma simultánea, manteniendo alguna relación entre las configuraciones estimadas y pudiendo determinar la aportación de la nueva dimensión a la configuración que se está estimando. Siendo así, la dimensión adecuada sería aquella que, globalmente, explicara una mayor variabilidad de los datos.

La clave está, por tanto, en conseguir en el mínimo número de iteraciones: a) la estimación de las coordenadas de todos los puntos *a la vez*, b) el análisis *conjunto* de varias configuraciones conforme aumenta la dimensionalidad y, al mismo tiempo, c) una medida de bondad del ajuste de cada configuración en función del número de dimensiones consideradas.

Respecto al primer aspecto, SMACOF se presenta claramente ventajoso porque no trabaja con cada punto de manera independiente considerando fijos los demás, sino que lo hace modificando las posiciones de todos los puntos *a la vez*. Respecto a la cuestión b), una de las soluciones sería trabajar con modelos anidados en ALSCAL (ver, por ejemplo, González-Carmona, Román y Vera, 1999); podemos decir, no obstante, que SMACOF también puede simultanear los resultados en varias dimensiones. Y en relación a la minimización de la función de pérdida, mientras que ALSCAL y otros algoritmos quedan atrapados en los mínimos locales (especialmente por la dificultad de converger con todos los puntos al mismo tiempo), SMACOF lo resuelve eficazmente aplicando la mayorización iterativa, que, dicho brevemente, permite alcanzar mucho antes los mínimos buscados.

Las anteriores consideraciones se llevan a cabo siguiendo los pasos que se esquematan en la Figura 6 (Borg y Groenen, 2005, p. 205). Como observará el lector pueden establecerse paralelismos entre las Figuras 4, 5 y 6.

Figura 6. Esquema del algoritmo SMACOF



EJEMPLO DESARROLLADO CON EL PAQUETE SMACOF DE R

Método

En el siguiente ejemplo hacemos un EMD-NM con una sola matriz. Los datos de entrada no son proximidades sino que proceden de una matriz rectangular de dos vías (filas x columnas) y dos modos (sujetos x ítems). Se trata de datos multivariados (también denominados “de perfil”) que contienen las respuestas dadas por los sujetos encuestados en distintos ítems. Por ello, primero hemos de construir la matriz de proximidades calculando las distancias euclídeas entre los distintos elementos y teniendo en cuenta la métrica ordinal de las respuestas emitidas por los sujetos (una escala Likert).

El algoritmo SMACOF está implementado en R en un paquete que lleva el mismo nombre (Leeuw y Mair, 2008b). Aunque el paquete Smacof permite varias técni-

cas (*unfolding models*, *idioscal*, *indscal*, etc.), en esta primera aproximación desarrollamos un SMACOF simple que trabaja con una matriz simétrica de desemejanzas.

Muestra, instrumento y variables

El fichero de datos procede de una muestra de 266 estudiantes de Medicina a los que se les pidió que valoraran el grado de acuerdo con cada uno de los ítems de la Tabla 1 sobre la “Finalidad de la educación médica”. Las respuestas se establecen con una escala Likert de 4 puntos (desde 1: “totalmente en desacuerdo”, hasta 4: “totalmente de acuerdo”). Son datos recogidos de la investigación que se desarrolla en Dam (2006).

Tabla 1. Ítems empleados

Finalidad de la educación médica	
ítem 1	Transmitir la herencia cultural
ítem 2	Desarrollar la inteligencia
ítem 3	Trabajar la dimensión socializadora
ítem 4	Concienciar de los problemas sociales
ítem 5	Enseñar a ser cooperativos
ítem 6	Desarrollar integralmente al individuo
ítem 7	Integrar a los individuos a la sociedad
ítem 8	Desarrollar la creatividad
ítem 9	Permitir manejar la tolerancia como un valor
ítem 10	Permitir desarrollar la autoestima
ítem 11	Trabajar por un bien común

Procedimiento

Aunque *R* es capaz de leer ficheros de datos procedentes de otros programas (SPSS, *Excel*, bases de datos, etc.), partimos de un fichero ASCII que podemos crear usando, por ejemplo, el *Notepad* de *Windows*. En dicho fichero tenemos una columna por cada ítem y una fila por cada registro. En la primera fila situamos los nombres de los ítems. Así, las primeras filas serán las siguientes:

i1	i2	i3	i4	i5	i6	i7	i8	i9	i10	i11
2	3	3	3	4	3	3	3	3	4	3
3	4	4	4	4	4	3	4	4	3	3
3	4	3	4	4	4	3	4	4	4	3
3	4	4	4	4	4	3	4	4	4	4
4	4	4	4	4	4	4	4	4	4	1
2	3	3	3	3	3	4	3	3	3	3
2	4	4	3	1	3	4	4	3	3	3
2	3	4	3	4	4	2	4	3	3	3
3	3	4	4	2	4	4	0	3	0	4
1	2	3	3	3	3	2	1	3	2	2
4	4	4	4	4	4	4	4	4	4	4
4	3	3	4	4	4	3	4	3	4	4

Al trabajar con R^8 , debemos primero leer los datos que vamos considerar, cargar el paquete *Smacof* y efectuar después el EMD-NM, terminando con un diagnóstico gráfico que facilite la interpretación de la salida.

- a. Para leer los datos, atendiendo a las características del fichero que acabamos de configurar, ejecutamos el siguiente comando:

```
> datos <- read.table(file="datos.dat", header=TRUE)
```

donde con “file” indicamos el nombre del fichero de datos, y con “header=TRUE”, que en la primera línea se encuentran los nombres de los ítems. Así obtenemos el objeto “datos” que usaremos más adelante.

- b. La siguiente tarea es instalar el paquete *Smacof* de *R*. Suponiendo que el equipo está conectado a Internet, ejecutamos:

```
> install.packages(smacof) [1]
```

Elegimos también algún CRAN *mirror* y pulsamos *OK*. Obtenemos un mensaje indicándonos que *Smacof* ha sido instalado satisfactoriamente. Es preciso también instalar la librería *Smacof*, tecleando:

```
> library(smacof) [2]
```

⁸ Recomendamos la lectura de Ruiz-Soler y López-González (2009) y Arriaga et al. (2008). Además, independientemente de la plataforma en la que trabajemos, *R* puede instalarse siguiendo las instrucciones de las páginas <http://www.r-project.org/>

c. Para poder realizar el EMD-NM, lo primero será conseguir la matriz de desemejanzas de los ítems:

```
> dis <- dist(t(datos)) [3]
> dis
      i1      i2      i3      i4      i5      i6      i7      i8      i9      i10
i2 20.68816
i3 24.69818 18.76166
i4 23.74868 17.32051 13.92839
i5 25.61250 19.39072 17.37815 14.49138
i6 28.46050 20.59126 16.49242 14.96663 13.49074
i7 20.42058 18.02776 20.51828 19.97498 21.33073 23.00000
i8 22.97825 19.54482 20.29778 18.54724 18.38478 19.23538 19.72308
i9 25.25866 20.73644 17.20465 16.37071 16.67333 16.00000 19.72308 17.88854
i10 23.00000 20.85665 19.10497 19.00000 18.73499 19.51922 17.88854 17.46425 15.71623
i11 26.53300 20.44505 17.14643 16.43168 16.61325 14.62874 20.56696 20.00000 15.42725 16.82260
```

La orden $dist(x)$ ⁹ proporciona la matriz de distancias entre las filas de X ; para ello usamos $t(datos)$, lo que supone transponer el fichero de datos y permitir obtener dicha matriz. Por defecto, $dist()$ trabaja con la distancia euclídea. El resultado es una matriz de entrada cuadrada, simétrica, de dos vías (filas x columnas) y un modo (ítems x ítems): cada elemento δ_{ij} muestra la proximidad entre los ítems i y j , y se cumple que $\delta_{ij} = \delta_{ji}$.

Resultados

La salida del EMD-NM se consigue por medio de la función $smacofSym()$, asignando el resultado al objeto “res”. En el prompt de R ejecutamos la siguiente instrucción:

```
> res <- smacofSym(dis, metric=FALSE) [4]
```

y para ver los resultados obtenidos:

```
> res
Model: Symmetric SMACOF
Number of objects: 11
Nonmetric stress: 0.006490167
Number of iterations: 35
```

⁹ En la instrucción $dist(x, method=euclidean)$, ‘method’ puede ser algunos de los siguientes términos: *euclidean*, *maximum*, *manhattan*, *canberra*, *binary* o *minkowski*.

La información muestra el modelo trabajado (Smacof Simétrico), el número de ítems analizados (11), el valor del *stress* obtenido (0.006490167) y el número de iteraciones necesarias para llegar a la solución (35). Si queremos ver la configuración obtenida, escribimos:

```
> summary(res)
Configurations:
      D1      D2
i1  0.1971  0.0296
i2  0.0455  0.0698
i3 -0.0407  0.0417
i4 -0.0309  0.0295
i5 -0.0715  0.0186
i6 -0.0851  0.0099
i7  0.0905 -0.0174
i8  0.0226 -0.0524
i9 -0.0366 -0.0316
i10 -0.0188 -0.0736
i11 -0.0720 -0.0241
```

Hasta ahora se ha usado el comando `smacofSym()` en su forma más simple, es decir, tomando los valores por defecto para los argumentos de la función. No obstante, dichos valores operativos pueden cambiarse. En la Tabla 2 señalamos algunos de los argumentos que nos parecen más interesantes.

Tabla 2. Algunos argumentos del paquete Smacof de R

Argumento	Función
<code>ndim = n</code>	Permite indicar el número de dimensiones de la solución. Por defecto, 2.
<code>metric = TRUE o FALSE</code>	El valor por defecto es TRUE, indicando que realizará un EMD-M. Si elegimos FALSE, obtendremos un EMD-NM.
<code>verbose = TRUE o FALSE</code>	Permite mostrar (TRUE) o no los valores intermedios de stress. El valor por defecto es FALSE.
<code>modulus = n</code>	Número de iteraciones por llamada a la regresión monótona.
<code>ltmax = n</code>	Máximo número de iteraciones

Si queremos obtener una solución con tres dimensiones, ejecutamos:

```
> res3d <- smacofSym(dis, ndim=3, metric=FALSE)      [5]
> res3d
Model: Symmetric SMACOF
Number of objects: 11
Nonmetric stress: 0.000350064
Number of iterations: 67
> summary(res3d)
Configurations:
```

	D1	D2	D3
i1	0.1820	0.0236	0.0395
i2	0.0523	0.0626	-0.0112
i3	-0.0445	0.0464	-0.0435
i4	-0.0341	0.0297	0.0005
i5	-0.0565	0.0106	0.0369
i6	-0.0825	0.0159	0.0073
i7	0.0764	-0.0248	-0.0475
i8	0.0218	-0.0302	0.0645
i9	-0.0423	-0.0430	-0.0020
i10	-0.0036	-0.0696	-0.0168
i11	-0.0689	-0.0212	-0.0276

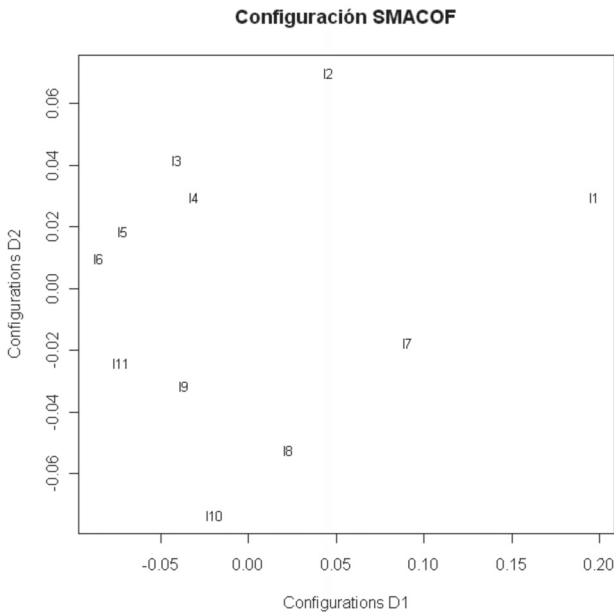
Es posible también calcular un EMD-NM que muestre los valores intermedios del *stress* simplemente efectuando el comando:

```
> smacofSym(datos, metric=FALSE, verbose=TRUE)
```

Ahora bien, *R* es un excelente programa de representación gráfica. Con el objeto “res”, creado en [4], y el comando “plot” podemos obtener distintas representaciones gráficas. Por ejemplo, la configuración SMACOF:

```
> plot(res, main="Configuración SMACOF")      [6]
```

Figura 7. Configuración de ítems por SMACOF



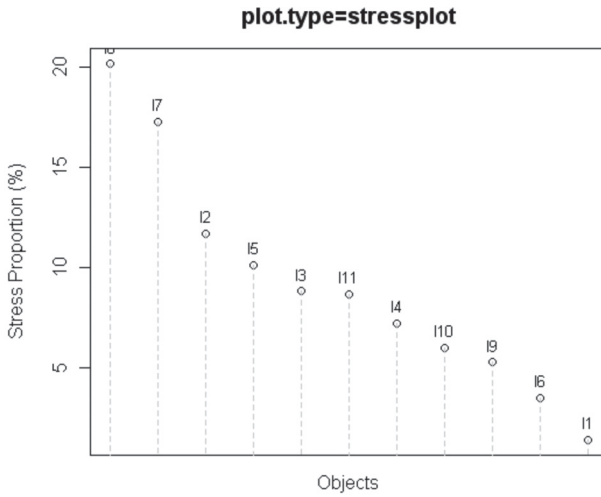
La Figura 7 es el mismo gráfico que se obtiene por defecto al ejecutar el comando `plot()` sobre un objeto `smacof`. Para conseguir otros gráficos se cambia el argumento que proporciona el tipo de gráfico que queremos obtener, “`plot.type`”, por uno de los siguientes:

- “`conjplot`”: Configuración del EMD (por defecto).
- “`stressplot`”: Descomposición del *stress* entre los ítems.
- “`shepard`”: Diagrama de Shepard.
- “`resplot`”: Configuración de las distancias frente los residuos.

Por ejemplo, la descomposición del *stress* entre los ítems (Figura 8) resulta de la orden:

```
> plot(resf2, plot.type = "stressplot", main="Stressplot") [7]
```

Figura 8. Descomposición del stress entre los ítems

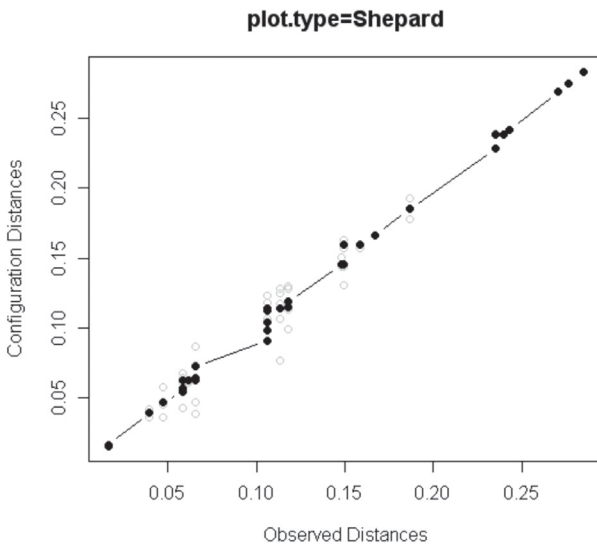


El diagrama de Shepard (Figura 9):

> plot(resf2, plot.type = "Shepard", main="Shepard")

[8]

Figura 9. Diagrama de Shepard: regresión monótona



Y la representación de los residuos (Figura 10):

```
> plot(resf2, plot.type = "resplot", main="Residuos") [9]
```

Figura 10. Gráfico de residuos



Discusión

La interpretación de la configuración en dos dimensiones es sencilla. Atendiendo a la dimensión 1 (los ítems situados en la derecha son aquellos en los que el grado de acuerdo con la Finalidad de la educación médica es menor), el ítem 1 (Transmitir la herencia cultural) parece menos acorde con la Finalidad de la educación médica, mientras que el i6 (Desarrollar integralmente al individuo) es el que muestra mayor acuerdo. Dichas observaciones se confirman también considerando los valores medios que proporciona la función *mean()*:

```
> mean(datos)
```

i1	i2	i3	i4	i5	i6
2.266917	2.996241	3.259398	3.342105	3.379699	3.597744
i7	i8	i9	i10	i11	
2.714286	3.033835	3.327068	3.060150	3.372180	

Respecto a las agrupaciones de ítems, el i6 queda asociado con el i5 (Enseñar a ser cooperativos) y el i11 (Los docentes médicos conllevan a trabajar por un bien común), con valores medios 3.597, 3.379 y 3.372, respectivamente.

Si seguimos observando esta primera dimensión podemos apreciar otro grupo de ítems con buena concordancia con la Finalidad de la educación médica, aunque algo menor que los tres anteriores, formado por el i3 (Trabajar la dimensión socializadora, 3.259), el i4 (Concienciar de los problemas sociales, 3.342) y el i9 (Permitir manejar la tolerancia como un valor, 3.327).

Al considerar la dimensión 2, que interpretamos como pertinencia del ítem al factor Finalidad de la educación médica, vemos el ítem 2 (Desarrollar la inteligencia) alejado del resto de ítems, pudiéndose afirmar que sería el de menor relación con la Finalidad de la educación médica. En esta dimensión encontramos dos grupos de ítems: el primero formado por los ítems i3 (Trabajar la dimensión socializadora), i4 (Concienciar de los problemas sociales), i1 (Transmitir la herencia cultural), i5 (Enseñar a ser cooperativos) e i6 (Desarrollar integralmente al individuo), y el segundo grupo por el i7 (Integrar a los individuos a la sociedad), el i11 (Los docentes médicos conllevan a trabajar por un bien común) y el i9 (Permitir manejar la tolerancia como un valor), quedando el resto de ítems separados de estos dos grupos.

El diagrama de Shepard (Figura 9) muestra la regresión monótona, donde aparecen las distancias observadas (proximidades) frente a las distancias transformadas (disparidades). El ajuste ha sido suavizado, sobre todo, en aquellas proximidades más bajas, por eso los residuos son mayores en las proximidades menores (Figura 10). Observamos también que los errores se producen por encima y por debajo de la línea de regresión isótoma (Figura 9), es decir, son positivos y negativos, lo que sugiere ausencia de heteroscedasticidad.

En la Figura 8 se aprecia la contribución de cada ítem en el *stress*. Los ítems que más contribuyen (entre un 15 y un 20%) son el i8 (Desarrollar la creatividad) y el i7 (Integrar a los individuos a la sociedad) –pueden considerarse los de más difícil ajuste por el modelo–, mientras que la contribución en el *stress* es mínima por parte del i1 (Transmitir la herencia cultural) y el i6 (Desarrollar integralmente al individuo).

CONCLUSIÓN

En los apartados anteriores hemos visto que SMACOF responde a los objetivos de un EMD-NM, siendo su proceso iterativo similar a la solución ALSICAL habitualmente empleada. Ahora bien, SMACOF tiene notables ventajas: supera los inconvenientes de ALSICAL, es un procedimiento apropiado para variables ordinales y se encuentra accesible en el software libre *R*. Por tales motivos consideramos que sería muy interesante un mayor uso en el ámbito educativo.

Recepción del original: 6 de febrero de 2009

Recepción de la versión definitiva: 2 de febrero de 2010

REFERENCIAS

- Arce, C. (1993). *Escalamiento multidimensional. Una técnica multivariante para el análisis de datos de proximidad y preferencia*. Barcelona: PPU.
- Arce, C. (1996). Escalamiento multidimensional. En J. Arnau, *Métodos y técnicas avanzadas de análisis de datos en Ciencias del Comportamiento* (pp. 23-46). Barcelona: Universitat de Barcelona.
- Arriaga, A., Fernández-Palacín, F., López-Sánchez, M. A., Muñoz-Márquez, M., Pérez-Plaza, S. y Sánchez-Navas, A. (2008). *Estadística básica con R y R-Commander*. Cádiz: Universidad de Cádiz.
- Bennett, J. F. y Hays, W. L. (1960). Multidimensional unfolding: Determining the dimensionality of ranked preferente data. *Psychometrika*, 25, 27-43.
- Borg, I. y Groenen, P. (2005). *Modern multidimensional scaling. Theory and applications* (2ª ed.). New York: Springer.
- Busing, F. M. T. A., Commandeur, J. J. F. y Heiser, W. J. (1997). PROXSCAL: A multidimensional scaling program for individual differences scaling with constraints. En W. Bandilla y F. Faulbaum (Eds.). *Advances in statistical software* (vol. 6, pp. 67-73). Stuttgart: Lucins & Lucins.
- Cailliez, F. (1983). The analytical solution of the additive constant problem. *Psychometrika*, 48, 305-308.
- Coombs, C. H. (1964). *A theory of data*. New York: Wiley.
- Cox, T. F. y Cox, M. A. A. (2001). *Multidimensional scaling* (2ª ed.). Boca Raton: Chapman & Hall/CRC.

- Dam, B. (2006). *Valores de docentes y alumnos sobre la educación en el Decanato de Medicina de la Universidad Centro Occidental Lisandro Alvarado*. Tesis doctoral no publicada, Universidad de Málaga, Málaga, España.
- De Leeuw, J. (1977). Applications of convex analysis to multidimensional scaling. En J. R. Barra, F. Brodeau, G. Romier y B. van Cutsem (Eds.), *Recent developments in statistics* (pp. 133-145). Amsterdam: North-Holland.
- De Leeuw, J. y Heiser, W. J. (1977). Convergence of correction-matrix algorithms for multidimensional scaling. En J. C. Lingoes, E. E. Roskam e I. Borg (Eds.). *Geometric representations of relational data* (pp. 735-752). Ann Arbor, MI: Mathesis Press.
- De Leeuw, J. y Heiser, W. J. (1980). Multidimensional scaling with restrictions on the configuration. En P. Krishnaiah (Ed.), *Multivariate analysis* (vol. V, pp. 501-522). Amsterdam: North Holland Publishing.
- De Leeuw, J. y Heiser, W. J. (1982). Theory of multidimensional scaling. En P. Krishnaiah y Kanal (Eds.), *Handbook of statistics* (Vol 2: Classification, pattern recognition and reduction of dimensionality, pp. 285-316). Amsterdam: North Holland Publishing.
- De Leeuw, J. y Mair, P. (2008a). Multidimensional scaling using majorization: SMACOF in R. *Statistics Preprint Series*, 537. Extraído el 30 de enero de 2009, de <http://preprints.stat.ucla.edu/537/smacof.pdf>
- De Leeuw, J. y Mair, P. (2008b, mayo, 24). *The SMACOF package*. Extraído el 30 de enero de 2009, de: <http://cran.r-project.org/web/packages/smakof/smakof.pdf>
- González-Carmona, A., Román, Y. y Vera, J. F. (1999). Anidamiento en MDS mínimo cuadrático no métrico. *Estadística Española*, 41(144), 129-143.
- Groenen, P. J. F. y Heiser, W. J. (1996). The tunneling method for global optimization in multidimensional scaling. *Psychometrika*, 61, 529-550.
- Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, 9, 139-150.
- Guttman, L. (1946). An approach for quantifying paired comparisons and rank order. *Annals of Mathematical Statistics*, 17, 144-163.
- Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, 33, 469-509.
- Heiser, W. J. (1989). Order invariant unfolding analysis under smoothness restrictions. En G. De Soete, H. Feger, y K. C. Klauer (Eds.), *New developments in psychological choice modelling* (pp. 3-31). Amsterdam: Elsevier Science.

- Jobson, J. D. (1992). *Applied multivariate data analysis. Vol. II: Categorical and multivariate methods*. New York: Springer.
- Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, 1-27.
- Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29, 115-129.
- Kruskal, J. B. (1977). Multidimensional scaling and other methods for discovering structure. En K. Enslein, A. Ralston, y H.S. Wilf (Eds.). *Statistical methods for digital computers. Vol. 2* (pp. 296-339). New York: Wiley.
- Kruskal, J. B. y Wish, M. (1978). *Multidimensional scaling*. Newbury Park, CA: Sage.
- Martín-Casado, A. M. y Galindo, M. P. (1994). Análisis de proximidades: desarrollo, progresos y aplicaciones. *Estadística Española*, 36(135), 99-144.
- Meulman, J. J., Heiser, W. J. y SPSS Inc. (2008). *Categories 17.0*. Chicago: SPSS Inc.
- Real, J. E. (2001). *Escalamiento multidimensional*. Madrid: La Muralla.
- Ruiz-Soler, M. y López-González, E. (2009). El entorno estadístico R: ventajas de su uso en la docencia y la investigación. *Revista Española de Pedagogía*, 243, 255-274.
- Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, 27(2), 125-140.
- Takane, Y., Young, F. W. y De Leeuw, J. (1977). Nonmetric individual differences in multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika*, 42, 7-67.
- Torgerson, W. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika*, 17, 401-419.