QÜESTIIÓ, Vol 20, 1 pp. 29-43, 1996

UN PROCEDIMIENTO DE MDS A DOS VÍAS PARA EL ANÁLISIS MEDIANTE MÁXIMA VEROSIMILITUD DE DATOS DE DISIMILARIDAD CON ORIGEN INDETERMINADO

VERA, J. FERNANDO y GONZÁLEZ, ANDRÉS

Departamento de Estadística e Investigación Operativa
Universidad de Granada*

En este trabajo se propone un modelo estadístico para el análisis mediante Multidimensional Scaling (MDS) métrico de datos de disimilaridad asociados a un sujeto, en el que los juicios han sido obtenidos mediante escalas de tipo intervalo. La distribución propuesta para la modelización de las disimilaridades es la lognormal triparamétrica y el procedimiento empleado para la estimación de los parámetros, el de máxima verosimilitud. La métrica considerada ha sido la euclídea y se ha utilizado una transformación de potencias para el ajuste de los datos de disimilaridad corregidos a las distancias estimadas, todo ello dentro de un modelo basado en el propuesto por Ramsay.

A two-way MDS method for Maximum Likelihood Analysis of Dissimilarity Data with unknown origin

Key words: MDS métrico a dos vías, parámetro umbral, distribución lognormal triparamétrica, datos de tipo intervalo, máxima verosimilitud, variabilidad entre estímulos.

* Departamento de Estadística e Investigación Operativa. Facultad de Ciencias. Universidad de Granada. 18071 Granada.

– Article rebut el setembre de 1994.

– Acceptat el setembre de 1995.

1. INTRODUCCIÓN

Ramsay (1977,1978,1982) desarrolla un modelo confirmatorio de MDS métrico basado en la distribución lognormal biparamétrica que permite un análisis de datos de disimilaridad obtenidos en escalas de clasificación dentro de un contexto de máxima verosimilitud. El modelo de Ramsay se desarrolla bajo las hipótesis fundamentales de que los datos ajustados son *independientes*, *positivos* y poseen un *origen determinado en cero*. Aunque los estimadores de la mayoría de los parámetros del modelo lognormal sean robustos incluso frente a fuertes correlaciones (Ramsay (1981)) la hipótesis de independencia es uno de los aspectos más problemáticos y difícil de verificar experimentalmente, siendo aún más restrictiva en modelos como el de Takane (1981) o el de Takane y Carroll (1981), en los que el proceso de toma de datos puede ser más susceptible de ocasionar falta de independencia. No obstante, respecto al tipo de datos que puede analizarse mediante el modelo lognormal de Ramsay resultan particularmente restrictivas las hipótesis de positividad y origen determinado, ya que plantean problemas en diversas situaciones prácticas.

Uno de los aspectos básicos de la teoría de MDS lo constituye la hipótesis de que la disimilaridad entre dos estímulos idénticos es nula. Sin embargo, un problema general y que afecta de forma particular a la suposición de lognormalidad biparamétrica de los datos surge al comprobarse experimentalmente (Ramsay (1977)) que cuando a los individuos se les presentan dos estímulos idénticos tienden a asignarles un valor de disimilaridad no nulo. Por tanto, aunque desde el punto de vista teórico los valores diagonales de la matriz de disimilaridad deben ser nulos, desde el punto de vista práctico esta hipótesis se incumple en muchas situaciones.

Otro problema importante lo constituye el hecho de que en muchos casos de MDS métrico los datos son recogidos en escalas de disimilaridad con origen arbitrario. La modelización de este tipo de datos de disimilaridad mediante la distribución lognormal biparamétrica puede dar lugar a conclusiones inexactas debido fundamentalmente a que a priori no puede asegurarse que el dominio sea $\mathbb{R}^+$, sino en general (θ,∞) con $\theta \in \mathbb{R}$ algún valor desconocido.

Por otra parte, en algunas situaciones prácticas los datos no corresponden a valores de disimilaridad sino de similaridad. Existen diversos procedimientos para obtener datos de disimilaridad a partir de valores de similaridad dependiendo de la naturaleza del problema y en general, uno sencillo y adecuado para obtener disimilaridades que ha sido adoptado por diversos modelos de MDS como por ejemplo INDSCAL, consiste en multiplicar por (-1) los valores de similaridad (Arabie *et al.* (1987)). No obstante, los datos de disimilaridad obtenidos pueden ser negativos y poseen un origen indeterminado por lo que no pueden ser tratados directamente por un modelo lognormal biparamétrico.

Una alternativa para resolver algunos de los problemas mencionados consiste en transformar los datos para su análisis y en este sentido Ramsay propone en ciertas ocasiones una transformación a priori dada por

$$g(\delta) = \mathrm{Min}\{a\delta^b - c, e\} \tag{1}$$

donde δ representa cada valor original de proximidad y a, b, c y e son constantes fijadas arbitrariamente por el investigador (Schiffman *et al.* (1981)). No obstante, el hecho de que una traslación en la escala de disimilaridad puede tener una gran influencia en la solución obtenida e incluso sobre la dimensión estimada (Heiser (1991)) hace que una transformación previa y arbitraria de ese tipo de datos pueda conducir a resultados que no se ajusten a los juicios de disimilaridad obtenidos.

Los problemas que ocasionan las traslaciones arbitrarias en la escala de disimilaridad sugieren la búsqueda de una solución alternativa. En este trabajo proponemos un modelo basado en el de Ramsay que permite un análisis de datos de disimilaridad obtenidos en escalas de tipo intervalo mediante la utilización de la distribución log-normal triparamétrica, lo que supone una solución general en MV a la estimación del origen que no depende del criterio del investigador.

2. DESCRIPCIÓN DEL MODELO

Siguiendo la notación del modelo de Ramsay, supondremos que los datos están constituidos por n estímulos y T réplicas y si no se considera simetría se dispone de T matrices cuadradas $n \times n$ de datos de disimilaridad entre cada par de estímulos, cuyos elementos denotaremos por d_{ijt}, $i, j = 1, \ldots, n$, asociados a cada pareja de estímulos (i, j) y $t = 1, \ldots, T$ a la t-ésima réplica. Denotaremos por X a la matriz de configuración y por x_i al punto correspondiente al estímulo i. El modelo de distancia empleado es el modelo euclídeo, donde por d_{ij}^* indicaremos la distancia entre cada pareja de puntos i y j asociados a los correspondientes estímulos en dimensión K.

Un diagrama de Shepard de los datos de disimilaridad frente a las distancias estimadas pone de manifiesto que una transformación monótona suele contribuir a la mejora del ajuste de los datos de disimilaridad a las distancias estimadas. Una transformación de potencias de los datos de disimilaridad corregidos en el origen ejerce este efecto, por lo que utilizaremos la transformación,

$$g(d_{ijt}) = (d_{ijt} - \theta)^{\beta} \tag{2}$$

donde θ y β son dos parámetros a estimar.

Bajo las hipótesis detalladas en la Sección 1, cada variable D_{ij} tendrá una densidad

$$f(d_{ijt} \mid \beta,\theta,X,\sigma_{ij}^2) = \frac{\beta}{\sigma_{ij}(d_{ijt}-\theta)\sqrt{2\pi}} \exp\left\{ \frac{-\log^2\left(\frac{(d_{ijt}-\theta)^{\beta}}{d_{ij}^*}\right)}{2\sigma_{ij}^2} \right\} \tag{3}$$

$$d_{ijt} > \theta; \quad i,j = 1,\ldots,n; \quad t = 1,\ldots,T.$$

donde θ es el parámetro umbral y σ_{ij} depende de cada pareja de estímulos (i,j). La log-verosimilitud asociada será

$$\log L_{\{d_{ijt}\}}(\beta,\theta,\sigma_{ij}^2,\{x_{pm}\}_{pm}) = -\frac{1}{2}\sum_t \sum_{i\neq j} \frac{1}{\sigma_{ij}^2}\log^2\left(\frac{(d_{ijt}-\theta)^{\beta}}{d_{ij}^*}\right) -$$

$$-\sum_t \sum_{i\neq j} \log(d_{ijt}-\theta) - T\sum_{i\neq j}\log(\sigma_{ij}) + M\log(\beta) - M\log(\sqrt{2\pi}), \tag{4}$$

siendo M el número de datos observados.

Un aspecto importante consiste en estudiar la contribución de cada estímulo al valor total de la varianza mediante la descomposición de la variabilidad en componentes específicas de cada estímulo que denotaremos por α_i (Ramsay (1978)). Merece especial interés la forma de relacionar dichas componentes ya que aunque el modelo aditivo propuesto por Ramsay ofrece una solución al problema, una relación de tipo multiplicativo explica de forma satisfactoria dicha relación cuando es utilizada en un modelo lognormal.

El modelo considerado es

$$\sigma_{ij}^2 = (\alpha_i^2\alpha_j^2)^{1/2}. \tag{5}$$

Si por el contrario se desea estudiar la variabilidad debida exclusivamente al individuo se considerará $\sigma_{ij} = \sigma$, que puede interpretarse como el error típico asociado al individuo y que es el aspecto de mayor interés en la mayoría de los análisis.

3. ESTIMACIÓN DE LOS PARÁMETROS

La estimación de los parámetros del modelo se realiza mediante el criterio de máxima verosimilitud. Por simplicidad en el cálculo para la estimación de los parámetros consideramos la variabilidad, σ^2, dependiente sólo del individuo encuestado.

Debido a la naturaleza del modelo la estimación de los parámetros se realiza numéricamente y sujeta a restricciones siguiendo un procedimiento similar en líneas generales al descrito por Ramsay aunque con las características propias que la consideración de un parámetro umbral origina en la distribución lognormal. Distinguimos tres bloques principales:

3.1. Estimación del parámetro umbral

La introducción del parámetro umbral en la distribución lognormal crea diversas complicaciones en cuanto a la estimación de los parámetros respecto del caso biparamétrico. Heyde (1963) demuestra que la distribución lognormal triparamétrica no está determinada de forma única por sus momentos lo que imposibilita utilizar el método de estimación basado en los mismos. Hill (1963) prueba que los estimadores globales por máxima verosimilitud, es decir, aquellos que ofrecen un máximo global de la función de verosimilitud, conducen a estimaciones inadmisibles bajo ciertas situaciones extremas. Por tanto, varios investigadores han utilizado procesos de estimación alternativos que puedan estar exentos de los problemas que aparecen en los métodos de los momentos y máxima verosimilitud.

Cohen (1951) y Harter & Moore (1966) han propuesto estimadores locales por máxima verosimilitud, Hill (1963) propuso estimadores bayesianos, Box & Cox (1964) utilizaron técnicas de estimación gráfica, etc... Hay que destacar los trabajos de Cohen & Whitten (1980) y de Cohen *et al.* (1985) en los que se proponen diversas modificaciones en los estimadores por máxima verosimilitud y momentos de Cohen (1951).

Para la estimación del parámetro umbral hemos utilizado estimadores locales por máxima verosimilitud teniendo en cuenta los problemas de convergencia que la estimación global plantea en algunos casos extremos, tal y como ponen de manifiesto Wilson & Worcester (1945) y Lambert (1964). Usamos una técnica similar a la descrita por Cohen (1951) la cual, como pone de manifiesto Calitz (1973), no plantea excesivos problemas de convergencia.

La ecuación implícita que da lugar a la obtención del parámetro umbral viene dada por:

$$\left[\sum_t \sum_{i \neq j} \frac{1}{d_{ijt} - \theta}\right] \left[\frac{1}{2M} \sum_{i<j} \left(\sum_t \left\{\log(d_{ijt} - \theta) + \log(d_{jit} - \theta)\right\}\right)^2 - \right.$$

$$\frac{1}{n(n-1)}\sum_t\sum_{i\neq j}\log^2(d_{ijt}-\theta)\Bigg] - R\sum_t\sum_{i\neq j}\frac{\log(d_{ijt}-\theta)}{d_{ijt}-\theta}+ \tag{6}$$

$$\frac{1}{2}\sum_{i<j}\left\{\left(\sum_t\left\{\frac{1}{d_{ijt}-\theta}+\frac{1}{d_{jit}-\theta}\right\}\right)\left(\sum_t\left\{\log(d_{ijt}-\theta)+\log(d_{jit}-\theta)\right\}\right)\right\}=0.$$

Aunque la resolución de esta ecuación mediante procedimientos numéricos ofrece un estimador por máxima verosimilitud del parámetro umbral independiente de los demás parámetros, nuestra experiencia indica que en aquellas ocasiones en las que no se dispone de un número suficiente de réplicas o los datos muestran mucha variabilidad resulta adecuado estimar dicho parámetro mediante un procedimiento condicional alternativo.

El procedimiento utilizado está basado en el descrito por Ramsay para la estimación del caso biparamétrico y consiste en la actualización del parámetro umbral en un ciclo iterativo que resuelve numéricamente la ecuación de verosimilitud mediante un procedimiento de gradiente. Dicha ecuación adopta la expresión

$$\sum_t\sum_{i\neq j}\frac{1}{d_{ijt}-\theta}\log\left(\frac{(d_{ijt}-\theta)^{\beta}}{d^*_{ij}}\right)+\frac{\sigma^2}{\beta}\sum_t\sum_{i\neq j}\frac{1}{d_{ijt}-\theta}=0 \tag{7}$$

donde los demás parámetros se mantienen fijos dentro del ciclo iterativo que actualiza θ.

3.2. Estimación de las componentes de la varianza

Una estructura de la variabilidad que explica simultáneamente los dos factores principales del modelo viene dada por

$$\sigma^2_{ij}=\sigma^2(\alpha_i^2\alpha_j^2)^{1/2}. \tag{8}$$

La estimación se realiza bajo restricciones que resuelven el problema de la indeterminación de forma que si el interés se centra exclusivamente en la componente relativa al individuo, bastará con considerar $\alpha_i = 1$, $i = 1,\ldots,n$. En ese caso, el estimador condicional de σ^2 viene dado por

(9) $$\sigma^2 = \frac{1}{M}\sum_t \sum_{i \neq j} \log^2 \left(\frac{(d_{ijt} - \theta)^{\beta}}{d^*_{ij}} \right).$$

Cuando el interés se centra en el estudio de la componente de variabilidad asociada a cada estímulo dicha componente es estimada en función de la expresión,

(10) $$\alpha_p^2 = \left[\frac{1}{2Tn} \sum_j \frac{S_{pj} + S_{jp}}{\mid \alpha_j \mid} \right]$$

donde para cada $i, j = 1, \ldots, n; i \neq j$,

(11) $$S_{ij} = \sum_t \log^2 \left(\frac{(d_{ijt} - \theta)^{\beta}}{d^*_{ij}} \right)$$

estando su estimación sujeta a la restricción dada por,

(12) $$\sum_i \alpha_i^2 = n.$$

Esta restricción se lleva a cabo mediante el empleo de una función de penalización $Q(X, \alpha) = q(X) + q(\alpha)$ (Fiacco & McCormick, (1968)). Al igual que considera Ramsay para el caso biparamétrico, el término $q(\alpha)$ asociado a la restricción sobre los errores típicos de los sujetos viene dado por

(13) $$q(\alpha) = -\frac{1}{2} \left(\sum_i \alpha_i^2 - n \right)^2.$$

Cuando el interés se centra en el estudio del factor de variabilidad asociado al individuo, pueden obtenerse estimaciones post hoc de las componentes de variabilidad de los estímulos una vez estimados los restantes parámetros del modelo.

3.3. Estimación de la matriz de configuración y del parámetro de la transformación

La estimación de la matriz de configuración constituye el aspecto central de MDS. El modelo euclídeo a dos vías posee invarianza traslacional y rotacional, por lo que para eliminar las indeterminaciones se introducen restricciones sobre las coordenadas de los estímulos. Las restricciones impuestas vienen dadas por

(14) $$\sum_i x_{im} = 0; \quad m = 1, \ldots, K$$

(15) $$\sum_i x_{im}x_{il} = 0; \quad m, l = 1, \ldots, K; m \neq l.$$

La primera restricción resuelve la invarianza traslacional situando el origen de coordenadas en el centroide de la configuración y la segunda resuelve la invarianza rotacional ya que fija la orientación en la dirección de los ejes principales. Las restricciones se imponen añadiendo una nueva componente a la función de penalización. Así, la componente de $Q(X, \alpha)$ respecto a X viene dada por

(16) $$q(X) = -\frac{1}{2}\sum_m \left(\sum_i x_{im}\right)^2 - \frac{1}{2}\sum_{m \neq l} \left(\sum_i x_{im}x_{il}\right)^2.$$

El estimador asociado a cada coordenada, $x_{pq}, p = 1, \ldots n, q = 1, \ldots K$ de la matriz de configuración, X, se obtiene mediante un procedimiento que utiliza el método del gradiente para minimizar en X la suma de los residuos al cuadrado

(17) $$SRC = \sum_{i \neq j}\sum_t e_{ijt}^2$$

donde los errores vienen dados por

$$e_{ijt} = \beta \log(d_{ijt} - \theta) - \log(d_{ij}^*)).$$

Puesto que minimizar $-\log L$ respecto a X equivale a minimizar SRC respecto a X, se utiliza ésta para obtener el estimador de la configuración.

El valor de X que minimiza (17) es la solución de la ecuación,

(18) $$x_{pq} \sum_{j \neq p} u_{pj} = \sum_{j \neq p} u_{pj}x_{jq}, \quad p = 1, \ldots n, \quad q = 1, \ldots K$$

donde

(19) $$u_{pj} = \frac{\beta}{(d_{pj}^*)^2} \sum_t \left\{\log(d_{pjt} - \theta) + \log(d_{jpt} - \theta) - 2T\log(d_{pj}^*)\right\}.$$

Para la estimación de la matriz de configuración puede utilizarse el método *Scoring* tal y como se describe en Ramsay (1982).

El otro aspecto que trata esta sección consiste en la estimación del parámetro de la transformación. Estimar la transformación de los datos simultáneamente al análisis de los mismos constituye uno de los aspectos centrales del modelo presentado para el ajuste de disimilaridades a distancias euclídeas, ya que la hipótesis básica del modelo es la lognormalidad de los datos transformados por (2).

La estimación de éste se realiza con la misma metodología que en el resto de los parámetros, considerando un bloque dentro del ciclo iterativo general en el que el parámetro es actualizado.

La expresión utilizada para actualizar el parámetro viene dada por,

(20)
$$\beta = \frac{T \sum_{i \neq j} \log^2(d^*_{ij})}{\sum_t \sum_{i<j} \left\{ \log(d_{ijt} - \theta) + \log(d_{jit} - \theta) \right\} \log(d^*_{ij})}.$$

La estimación de β está sujeta a la restricción de positividad.

4. TRATAMIENTO COMPUTACIONAL

Los estimadores de los parámetros que maximizan la logverosimilitud se obtienen mediante procedimientos numéricos. Para ello hemos utilizado un método similar al propuesto por Takane *et al.* (1977) y Ramsay (1982) que consiste en un proceso iterativo compuesto formado por un ciclo de iteraciones principales dentro del cual cada parámetro es actualizado en un proceso iterativo que se denomina secundario y que considera fijos el resto de parámetros.

Para la actualización del parámetro umbral se ha utilizado un procedimiento iterativo de minimización que emplea el vector gradiente para determinar la dirección de búsqueda junto a un algoritmo que utiliza interpolación y extrapolación cúbica (Fletcher (1980)) para determinar el tamaño de paso óptimo.

Se han utilizado dos criterios de convergencia: por un lado uno puramente geométrico mediante el cual se controla el módulo del vector gradiente y por otro lado un procedimiento basado en el estadístico χ^2.

Para ilustrar el método presentamos un ejemplo analizado por Ramsay (1986) en el que los datos corresponden a 105 valores de disimilaridad emitidos por un sujeto sobre 15 actividades recreativas. Los estímulos corresponden a las actividades: Concierto, museo, teatro, cine, televisión, conferencia, lectura, hockey, ballet, debate, moda, cine documental, exhibición, compras y restaurante.

Dado que el valor de disimilaridad mínimo es 2 (entre los estímulos *televisión* y *cine*) se ha elegido 1.99 como valor inicial de θ. El resumen del proceso de actualización de θ y de los valores de logverosimilitud asociados en cada iteración principal se presenta en la Tabla 1.

Tabla 1

Resumen de las iteraciones principales en el parámetro umbral

ITERACIÓN	VALOR DE θ	LOGVEROSIMILITUD
1	1.990000	−422.2884
2	1.939591	−376.8369
3	1.762395	−373.9680
4	1.682636	−372.1657
5	1.639950	−369.6468
6	1.575465	−367.5696
7	1.439258	−367.1207
8	1.307071	−366.3680
9	1.270560	−354.6981
10	1.203587	−353.6509
11	1.090658	−351.7810
12	0.720304	−351.5068
13	0.496138	−351.1981
14	0.427442	−351.0564
15	0.382136	−351.0506
16	0.259492	−350.7776
17	0.232850	−350.7397

El criterio de convergencia empleado ha sido que la diferencia entre dos valores consecutivos de la logverosimilitud no exceda de 0.05. Aunque dicha condición se verifica entre las iteraciones 14 y 15, el procedimiento continúa ya que en dichas ocasiones no se obtuvo convergencia en el ciclo de actualización de la matriz de configuración. El valor final estimado para θ ha sido 0.23285. La configuración obtenida se presenta en la Figura 1 junto a la solución ofrecida por MULTISCALE.

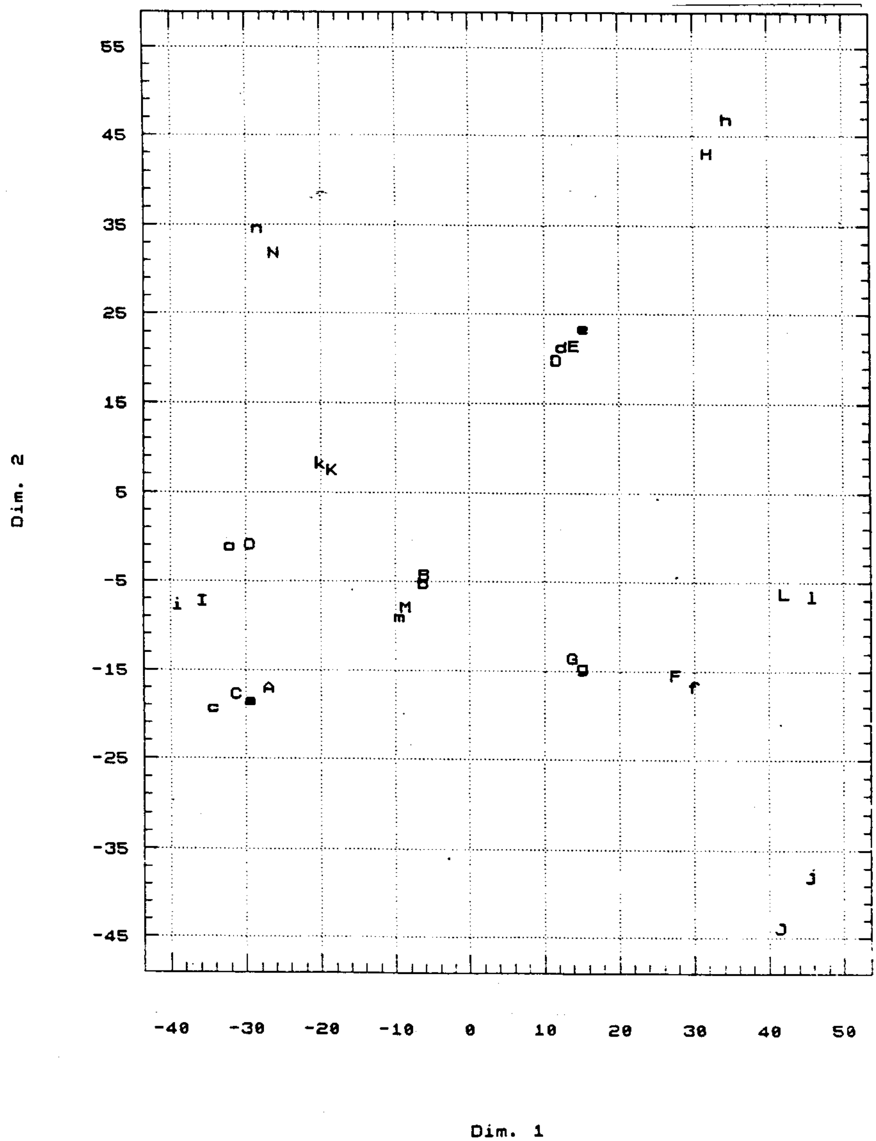

Figura 1

Representación de las configuraciones para $\theta = 0.23285$ (mayúsculas) y $\theta = 0$ (minúsculas), en dimensión dos.

El valor final de la función de logverosimilitud, -350.7397, es menor que el que se obtiene para $\theta = 0$, cuyo valor es -350.8692, siendo la diferencia entre ambos significativa según el criterio establecido. El máximo de la función de logverosimilitud se alcanza en $\theta = 0.23$.

A pesar de que los datos analizados en este ejemplo han sido obtenidos por Ramsay en una escala de razón, en el gráfico se aprecian diferencias significativas entre ambas configuraciones, lo cual pone de manifiesto el hecho de que en general el individuo encuentra un valor de disimilaridad distinto de cero entre dos estímulos idénticos y en este caso equivalente a 0.23 en la escala utilizada.

Por tanto, puede concluirse que cuando los datos provienen de escalas de razón el proceso descrito permite estimar el origen de la escala utilizada y en el caso en que los datos son de tipo intervalo la estimación del parámetro umbral constituye un aspecto que debe ser considerado en el estudio.

REFERENCIAS

[1] **Arabie, P., Carroll, J. D.** y **DeSarbo, W. S.** (1987). "Three-Way Scaling And Clustering". *Sage Publications, Inc.*

[2] **Box, G. E. P.** y **Cox, D. R.** (1964). "An analysis of transformations". *J. R. Statist. Soc. B*, **26**, 211–251.

[3] **Calitz, F.** (1973). "Maximum likelihood estimation of the parameters of the three-parameter lognormal distribution-a reconsideration". *Austral. J. Statist.*, **3**, 185–190.

[4] **Cohen, A. C.** (1951). "Estimating parameters of logarithmic-normal distributions by maximum likelihood". *J. Amer. Statist. Assoc.*, **46**, 206–212.

[5] **Cohen, A. C.** y **Whitten, B. J.** (1980). "Estimation in the three-parameter lognormal distribution". *J. Amer. Statist. Assoc.*, **75**, 399–404.

[6] **Cohen, A. C., Whitten, B. J.** y **Ding, Y.** (1985). "Modified moment estimation for the three-parameter lognormal distribution". *J. Qual. Tech.*, **17**, 92–99.

[7] **Fiacco, A. V.** y **McCormick, G. P.** (1968). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques.* John Wiley and Sons, Inc.

[8] **Fletcher, R.** (1980). *Practical Methods of Optimization Vol.1. Unconstrained Optimization.* John Wiley and Sons.

[9] **Harter, H. L.** y **Moore, A. H.** (1966). "Local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples". *J. Amer. Statist. Assoc.*, **61**, 842–851.

[10] **Heiser, W. J.** (1991). "A Generalized Majorization Method for Least Squares MDS of Pseudodistances that May Be Negative". *Psychometrika*, **56**, 7–27.

[11] **Heyde, C. C.** (1963). "On a property of the lognormal distribution". *J. R. Statist. Soc. B*, **25**, 392–393.

[12] **Hill, B. M.** (1963). "The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic". *J. Amer. Statist. Assoc.*, **58**, 72–84.

[13] **Lambert, J. A.** (1964). "Estimation of parameters in the three-parameter lognormal distribution". *Austral J. Statist.*, **6**, 29–32.

[14] **Ramsay, J. O.** (1977). "Maximum Likelihood Estimation in Multidimensional Scaling". *Psychometrika*, **42**, 241–266.

[15] **Ramsay, J. O.** (1978). "Confidence Regions for MDS Analysis". *Psychometrika*, **43**, 145–160.

[16] **Ramsay, J. O.** (1981). *Multiscale-Introduction to Multidimensional Scaling ed. by S. S. Schiffman, M. L. Reynolds and F. W. Young.* Academic Press, Inc.

[17] **Ramsay, J. O.** (1982). "Some Statistical Approaches to MDS Data". *J. R. Statist. Soc. A*, **145**, 285–312.

[18] **Ramsay, J. O.** (1986). *MULTISCALE II Manual.* Department of Psychology. McGill University.

[19] **Schiffman, S. S., Reynolds, M. L.** y **Young, F. W.** (1981). *Introduction to Multidimensional Scaling. Theory, Methods and Applications.* Academic Press, Inc.

[20] **Takane, Y., Young, F. W.** y **de Leew, J.** (1977). "Nonmetric Individual Differences MDS: An alternating Least Squares Method with Optimal Scaling Features". *Psychometrika*, **42**, 7–67.

[21] **Takane, Y.** (1981). "Multidimensional Successive Categories Scaling: A maximum Likelihood Method". *Psychometrika*, **46**, 9–28.

[22] **Takane, Y.** y **Carroll, J.D.** (1981). "Nonmetric Maximum Likelihood Multidimensional Scaling from Directional Rankings of Similarities". *Psychometrika*, **46**, 389–405.

[23] **Wilson, E. B.** y **Worcester, J.** (1945). "The normal logarithmic transform". *Rev. Econ. statist.*, **27**, 17–22.

ENGLISH SUMMARY:

A TWO-WAY MDS METHOD FOR MAXIMUM LIKELIHOOD ANALYSIS OF DISSIMILARITY DATA WITH UNKNOWN ORIGIN

Vera, J. Fernando and González, Andrés

Data Analysis, and particularly Multidimensional Scaling, is typically separated into exploratory or confirmatory analyses. Ramsay's MDS model belongs to the latter group. This model, given in Ramsay[1977], [1980] and in Ramsay[1982], was formulated for data analysis of direct classification judges whith continuous responses scales (Schiffman *et al.* [1981]). Ramsay's model is taken here and accomodated for interval scales. Previously, the data are transformed by using the power transformation and the metric used is the Euclidean. The distribution of dissimilarities proposed here is the three-parameters lognormal distribution, and maximum likelihood estimators are used.

More precise, the data are responses d_{ijt} resulting from the presentation to one *subject* a pair of objects, concepts or other *stimuli*, i and j, $(i,j = 1,...,n)$ on trial or *replication* t, $(t = 1,..,T)$. So we have the random variable D_{ij} with observations d_{ijt}, $(t = 1,...,T)$, assumed to be all independent and identically distributed. Let d^*_{ij} be the theoretical distance between the points i and j. We suppose the dissimilarities may have unknown origin and therefore their thistribution contains a threshold parameter.

The power transformation is

$$g(d_{ijt}) = (d_{ijt} - \theta)^{\beta}$$

and the associated density for D_{ij} is

$$f(d_{ijt} \mid \beta,\theta,X,\sigma^2_{ij}) = \frac{\beta}{\sigma_{ij}(d_{ijt}-\theta)\sqrt{2\pi}} \exp\left\{ \frac{-\log^2\left(\dfrac{(d_{ijt}-\theta)^{\beta}}{d^*_{ij}}\right)}{2\sigma^2_{ij}} \right\}$$

$$d_{ijt} > \theta; \quad i,j = 1,\dots,n; \quad t = 1,\dots,T.,$$

We have used local maximum likelihood estimators for the threshold parameter estimation, using a technique similar to the one described by Cohen (1951), or a

conjugate gradient procedure, depending on the data structure. The new likelihood equations for the threshold parameter estimation are

$$\left[\sum_t \sum_{i \neq j} \frac{1}{d_{ijt} - \theta}\right] \left[\frac{1}{2M} \sum_{i<j} \left(\sum_t \{\log(d_{ijt} - \theta) + \log(d_{jit} - \theta)\}\right)^2 - \right.$$

$$\left. \frac{1}{n(n-1)} \sum_t \sum_{i \neq j} \log^2(d_{ijt} - \theta)\right] - R \sum_t \sum_{i \neq j} \frac{\log(d_{ijt} - \theta)}{d_{ijt} - \theta} +$$

$$\frac{1}{2} \sum_{i<j} \left\{ \left(\sum_t \left\{\frac{1}{d_{ijt} - \theta} + \frac{1}{d_{jit} - \theta}\right\}\right) \left(\sum_t \{\log(d_{ijt} - \theta) + \log(d_{jit} - \theta)\}\right) \right\} = 0$$

or

$$\sum_t \sum_{i \neq j} \frac{1}{d_{ijt} - \theta} \log\left(\frac{(d_{ijt} - \theta)^\beta}{d^*_{ij}}\right) + \frac{\sigma^2}{\beta} \sum_t \sum_{i \neq j} \frac{1}{d_{ijt} - \theta} = 0$$

Maximum Likelihood Estimation under restrictions are considered, and the likelihood equations are solved by a similar procedure to that proposed by Takane *et al.* (1977) and Ramsay (1982).