

---

# INTRODUCCION AL ANALISIS MULTIDIMENSIONAL NO-METRICO

Juan Javier Sánchez Carrión

---

## 1. *Introducción*

El modelo de las escalas multidimensionales que vamos a considerar es una extensión de las ideas subyacentes a los modelos de escalas unidimensionales, aplicable a aquellos casos en que la variabilidad de los estímulos analizados (por ejemplo, los coches, detergentes, naciones, etc., que juzgan los entrevistados) se produce respecto de varias dimensiones.

Dado un número de estímulos que difieren respecto de una serie de propiedades o dimensiones se trata de ver cuál es el número mínimo de estas dimensiones que permite explicar la variabilidad de los estímulos y cuáles son sus coordenadas (parámetros) en esas mismas dimensiones. A diferencia de las escalas unidimensionales, que permiten observar las propiedades de los estímulos sólo en relación a un valor numérico, susceptible de variar a lo largo de un continuo, el análisis multidimensional permite analizar la complejidad de los estímulos mediante su representación en un espacio de dos, tres o más dimensiones.

En el espacio multidimensional los estímulos están representados por puntos, correspondiendo su posición al grado o cantidad de «atributo complejo»

---

que aquéllos posean <sup>1</sup>; mientras que la distancia entre dos estímulos (entre dos puntos en el espacio) está en función de su grado de (di)similaridad: cuanto más semejantes sean, más próximos estarán en el espacio.

El interés del análisis multidimensional no-métrico, frente al métrico, radica en el hecho de que en el primero sólo es necesario hacer supuestos no-métricos sobre la naturaleza de los datos (se asume que son de nivel ordinal: el orden de (di)similaridad entre los estímulos), mientras que en la solución del análisis (las distancias entre los estímulos en un espacio  $r$  dimensional) se recupera la información métrica subyacente a los mismos. Esto es posible, tal como señalan Coxon y Jones (1984), debido al avance técnico que supuso la demostración hecha por Shepard de que «las constricciones meramente no métricas (esto es, información ordinal en los datos), si se imponen en número suficiente ponen límites muy estrechos a las posibles soluciones del análisis (en este contexto se entiende por solución la proyección de los estímulos en un número de ejes de referencia: por ejemplo, las latitudes y longitudes de los puntos). Tan rígidos son los límites impuestos en la solución por las constricciones ordinales que es posible identificar la mejor solución *métrica* (las proyecciones de los estímulos en los ejes de referencia definidos a un nivel de intervalos iguales de medida) para datos meramente ordinales». Dada la dificultad de encontrar en las ciencias sociales datos de tipo interval o métrico y el interés de las soluciones de esta naturaleza, la bondad del método queda justificada.

Según los diferentes modelos de escalas multidimensionales no-métricas el proceso de medición difiere. En el *modelo básico*, que es el que vamos a explicar en este artículo, la variabilidad que se observa en las respuestas de los entrevistados a la hora de evaluar los estímulos se adscribe a la propia variabilidad de los estímulos <sup>2</sup>. Sólo se miden los estímulos, tratando de lograr que sea en el menor número posible de dimensiones. Se desprecia la singularidad de las respuestas de cada uno de los entrevistados y se calcula una evaluación media para el conjunto.

En el análisis de las *diferencias individuales* y en el análisis de las *preferencias individuales* las diferencias que hay en las respuestas de los entrevistados también se tienen en cuenta, considerando que no sólo hay variabilidad en los estímulos sino también en su percepción por parte de los sujetos, por lo

<sup>1</sup> Algunos atributos o propiedades de los estímulos son unidimensionales, tal es el caso de la longitud, la temperatura o el peso, mientras que otros son multidimensionales, valga como ejemplo la posición espacial o el color. La posición espacial viene determinada por la longitud y la latitud, o la abscisa y la ordenada, suponiendo un espacio de dos dimensiones; el color viene definido por el brillo y la saturación. A estos segundos atributos son a los que nos referimos al hablar de atributos complejos. Para una definición más exhaustiva de los conceptos estímulo, atributo, dimensión, etc., véase TORGENSON (1958, pp. 247 y ss.).

<sup>2</sup> Los entrevistados podrían decir que España y México son dos países más semejantes que España y Japón; es decir, que respecto de alguna(s) propiedad(es) o característica(s) de los países los tres aparecen como diferentes a los ojos de los entrevistados.

que ambos (estímulos y sujetos) son objetos de medición<sup>3</sup>. En el primer caso (las diferencias individuales) vemos cómo perciben los mismos estímulos diferentes sujetos; y en el segundo (análisis de las preferencias) se estudia cuáles son las preferencias, frente a un grupo de estímulos, de cada uno de los entrevistados (véase una ilustración de ambos modelos en Coxon y Jones, 1984).

Al margen de las diferencias mencionadas, tal como señala Shepard (1972, página 1), las diversas técnicas de escalas multidimensionales coinciden en que

a) pretenden mostrar cualquier estructura subyacente a una matriz de datos, y

b) tratan de representar esta estructura de forma que sea mucho más accesible al ojo humano, especialmente en forma de figura geométrica.

En el resto de este artículo vamos a desarrollar un ejemplo original siguiendo en parte la exposición de Rabinowitz (1975), que permitirá ilustrar el fundamento del análisis multidimensional no métrico. Explicaremos el modelo analítico básico y abordaremos los problemas que se plantean a la hora de interpretar los resultados del análisis.

## 2. *Ejemplo ilustrativo*

Supongamos que estamos interesados en conocer la imagen de cinco países: Egipto, España, Japón, México y Polonia. Obtendríamos una información de interés si pidiéramos a los entrevistados que clasificaran a los países por parejas, en función del grado de semejanza que vieran en ellos. Si el entrevistado organiza su imagen de los países según un criterio de ricos y pobres, podemos esperar que perciba como más semejantes a los países de igual renta; si ve los países según un criterio de pertenencia a los bloques, igualmente podemos esperar que sitúe más próximos los países que pertenecen a un mismo bloque; y así podría ocurrir con cada una de las dimensiones que el individuo en cuestión juzgara pertinentes a la hora de clasificar a los países, influyendo todas ellas en su clasificación (ordenación) de los países.

Con el fin de proceder a la clasificación pedimos a los entrevistados que puntúen en una escala de 0 a 100 cada pareja posible de países, entendiendo que dos países totalmente iguales recibirán la puntuación mínima, mientras que el 100 se reserva para aquella pareja de países totalmente diferentes. Presen-

---

<sup>3</sup> El mismo grupo de la nota anterior podría decir que España y México son dos países más semejantes que España y Japón, en parte porque los países son "objetivamente" diferentes, pero también en parte porque ellos tienen una visión "particular" de los mismos, con lo cual habría que medir las variabilidades de los estímulos y de los sujetos. Por lo tanto, no basta con hacer una evaluación media de las puntuaciones que dan los entrevistados, sino que hay que considerar las respuestas individuales —o por grupos homogéneos.

tamos los resultados obtenidos en forma de matriz triangular, puesto que esta misma matriz será la que sirva de *input* para el análisis de los datos mediante ordenador (tabla 1).

TABLA 1

*Matriz de datos. Cada número es la media de los valores asignados por los entrevistados a cada pareja*

		Egipto	España	Japón	México	Polonia
		1	2	3	4	5
Egipto	1					
España	2	70				
Japón	3	95	89			
México	4	40	50	80		
Polonia	5	75	80	90	70	

Según estos datos, Egipto y México son los países más semejantes, mientras que Japón y Egipto son los más diferentes; entre medias quedan las ocho restantes parejas. De la información contenida en la matriz sólo vamos a retener la *ordenación* de las parejas de países, olvidándonos de las diferencias numéricas entre ellas.

La ordenación de países por parejas nos dice poco sobre su clasificación y los criterios subyacentes a esta clasificación. Sin embargo, si utilizamos esta ordenación para representar los cinco países en un espacio, la estructura subyacente a la percepción de los países por parte de los individuos quedaría más clara. Esta representación tendría en cuenta la similaridad de los países en orden a su colocación: cuanto más similares sean dos naciones más próximas estarán en el espacio, y su localización vendrá dada por una serie de parámetros —tantos como dimensiones tuviera el espacio representacional.

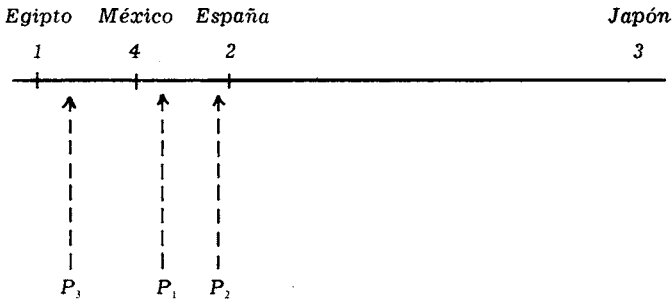
Según estos criterios, ¿es posible construir una escala con tales datos sobre una sola dimensión? Habría que colocar los nombres de los países en un continuo, respetando el principio de que «cuanto más similares sean dos países, más próximos estarán en el espacio».

Como los países más diferentes son Japón y Egipto, éstos estarán en los extremos de la escala. A continuación colocaríamos el par más similar, México-Egipto; por tanto, ambos deberán de aparecer uno junto a otro en la escala (véase gráfico 1). Consideremos ahora la colocación del punto correspondiente a España: nuestro país aparece como más semejante a México y Egipto, por

este orden; por tanto, colocándolo en el lugar en que aparece en el gráfico 1 satisface esta condición y al mismo tiempo se mantiene que México y Egipto sean los países más próximos de todos.

GRAFICO 1

*Representación de los puntos en una dimensión*

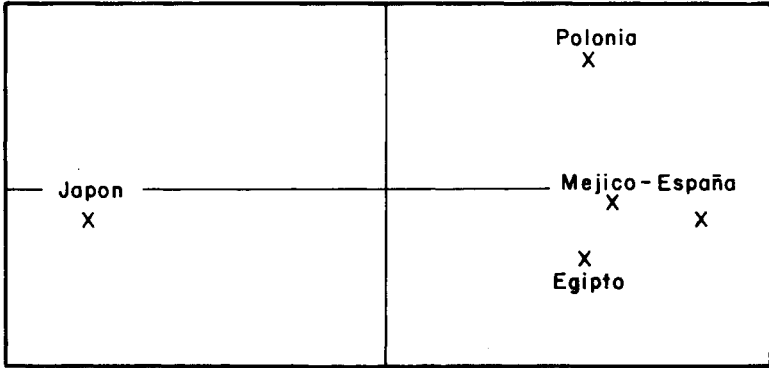


Hasta ahora vemos que los cuatro primeros puntos se pueden colocar en un eje (una sola dimensión). Veamos qué pasa cuando intentamos situar a Polonia. Este país aparece como más cercano a México, por ello habrá que colocarlo más cerca de México que de ningún otro país. Supongamos que lo colocamos en  $P_1$ . En esta situación cumple la condición mencionada pero rompe la condición de que México y Egipto sean los dos países más semejantes, ya que ahora la distancia que les separa es mayor que la existente entre México y  $P_1$ . Si lo colocamos en  $P_2$  se respeta la condición de que Egipto y México sean los países más semejantes pero se incumple la condición de que México sea el país más semejante a Polonia, por cuanto que ahora España aparece más próxima a  $P_2$  que México. Podríamos tratar de situar a Polonia en  $P_3$ . En este caso lo que pasa es que se rompe la condición de que México y Egipto sean los países más semejantes. Vemos, pues, que no hay solución a la representación gráfica de los puntos en una sola dimensión; se hace necesario recurrir a un espacio de más dimensiones. ¿Cuántas? Probo- mos una nueva solución tratando de situar los países en un espacio de dos dimensiones.

En el gráfico 2 colocamos los cinco puntos en un espacio bidimensional, distribuidos de forma arbitraria. En este caso tampoco la configuración responde al orden de las respuestas. Japón y España aparecen como los países más lejanos, mientras que en la ordenación de los entrevistados los más diferentes son Japón y Egipto. Tendríamos, pues, que ir moviendo los puntos de forma tal que fueran satisfaciendo el criterio de la escala. Para ello podemos ejecutar los desplazamientos de forma intuitiva o bien encontrar un procedimiento analítico que nos ayude en estos movimientos. Según el primero de los métodos tendríamos que ir moviendo cada punto, teniendo en cuenta que

GRAFICO 2

Representación (arbitraria) de los puntos en dos dimensiones



está interrelacionado con todos los demás y que por tanto la variación de un punto (país) afecta a las distancias que mantiene con el resto. Mientras el número de estímulos sea pequeño, el método es viable; pero a partir de tres puntos resulta prácticamente imposible encontrar la solución de forma intuitiva. El método correcto sería crear una serie de valores que nos guiasen en los movimientos. A estos valores les vamos a llamar las disparidades (*disparities*).

Kruskal (1964) define las disparidades,  $\hat{d}_{ij}$ , como «una secuencia monótona de números, elegidos tan próximos a las distancias entre los estímulos,  $d_{ij}$ , como sea posible, y que se utilizan como referencia para medir la no-monotonía de estos valores». Se trata de un conjunto de valores que han de ser monótonos respecto de los datos y que al mismo tiempo minimicen la suma de cuadrados  $\sum_{ij} (d_{ij} - \hat{d}_{ij})^2$ . Dicho de otra manera, las disparidades son una transformación monótona de los datos que al mismo tiempo han de asemejarse a las distancias entre los  $n$  estímulos en un sentido mínimo cuadrático.

Según la definición de las disparidades vemos que para su cálculo es necesario conocer la distancia entre los puntos y explicar qué es la no-monotonía de las distancias. Las distancias entre los puntos son fácilmente calculables. En el modelo Euclideo <sup>4</sup> la distancia entre dos puntos en un espacio de  $r$  dimensiones es igual a:

$$d_{ij} = \left( \sum_{m=1}^r (a_{im} - a_{jm})^2 \right)^{1/2}$$

<sup>4</sup> Existen otros modelos espaciales, como por ejemplo el de Attneave, pero nosotros sólo vamos a referirnos en este artículo al espacio Euclideo. Sobre los modelos espaciales véase TORGENSON (1958, pp. 251-259).

$d_{ij}$  es la distancia entre los estímulos  $i$  y  $j$  ( $i, j=1, \dots, n$ )  
 $m$  es un subíndice para los ejes ortogonales del espacio ( $m=1, \dots, r$ )  
 $a_{im}$  es la proyección del estímulo  $i$  en el eje  $m$

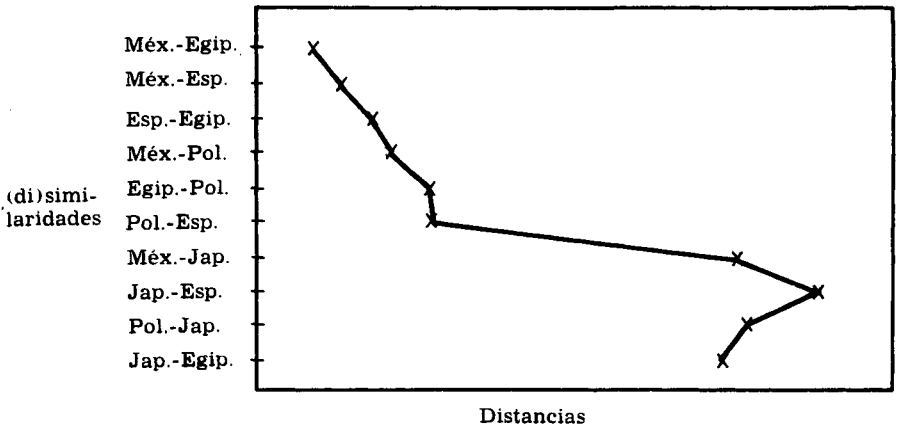
Sabiendo las coordenadas de los puntos en el espacio del gráfico 2 basta utilizar esta fórmula para calcular sus distancias. Por ejemplo, siendo las coordenadas de Polonia (2.7,1.6) y las de Japón (4.0, .5), la distancia entre ambos países es

$$d = [(2.7 + 4.0)^2 + (1.6 - .5)^2]^{1/2} = 6.8$$

Cuando tenemos un espacio de dos dimensiones se puede resolver el problema simplemente utilizando una regla. Si medimos la distancia que hay entre Polonia y Japón en la figura 2 vemos que coincide con la distancia que acabamos de calcular. Procediendo de esta manera construimos el gráfico 3, donde cada punto del gráfico representa una pareja de países.

GRAFICO 3

*Distancias entre los puntos en función de las (di)similaridades*



En la ordenada aparecen las parejas por orden de semejanza, mientras que en la abcisa se muestran las distancias entre las parejas.

El otro concepto al que se hacía referencia en la definición que hace Kruskal de las disparidades es el de «monotonía de las distancias». Ello quiere decir que la línea que une los puntos del gráfico 3 ha de ser monótona; es decir, que siempre se ha de mover de arriba abajo y de izquierda a dere-

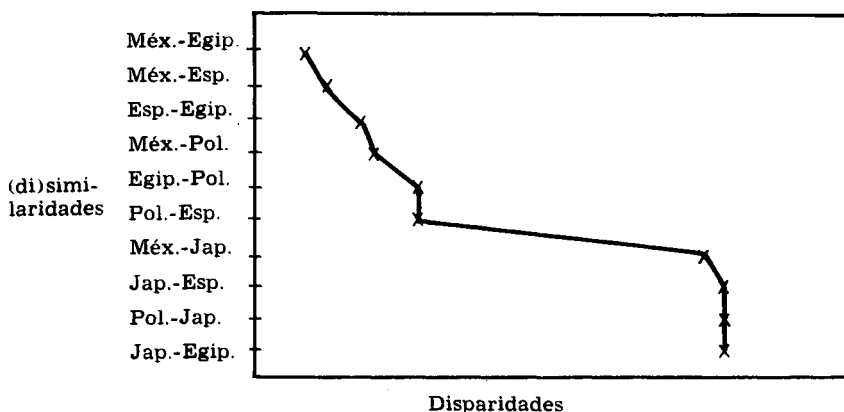
cha. Este es un requisito del análisis multidimensional no-métrico: las distancias en la solución han de ser una transformación monótona de las relaciones de (di)similaridad entre los estímulos (a mayor disimilaridad entre los estímulos mayor distancia entre los mismos).

La exigencia de monotonía en las distancias entre los estímulos es una de las diferencias básicas entre el análisis no-métrico y el métrico, donde se hace necesario que las distancias entre los estímulos sean una transformación *lineal* de las distancias que se obtienen en los datos <sup>5</sup>.

Volviendo a nuestro ejemplo vemos en el gráfico 3 que a partir de la pareja Japón-España se rompe la monotonía. Precisamente, con el fin de lograr que la línea de esta figura sea una función monótona (a mayor disimilaridad mayor distancia) nos vamos a servir de las disparidades. El gráfico 4 muestra las disparidades. Esta línea se construye de tal forma que sea tan semejante a la línea de las distancias como sea posible, con la restricción de que nunca se mueva hacia abajo y hacia la izquierda, preservando de esta manera la monotonía. En el apartado siguiente ofrecemos la explicación de su cálculo.

### GRAFICO 4

*Disparidades entre los puntos*



<sup>5</sup> En el modelo métrico, a partir de las (di)similaridades entre los estímulos hay que obtener las distancias entre ellos (distancias en los datos) (véase nota 12); estas distancias son las que sirven para predecir (mediante una *regresión lineal*) las distancias entre los estímulos, y por tanto, mediante la comparación entre las distancias previstas y las distancias reales se llega a sopesar la bondad de la solución obtenida. Para obtener una explicación de cómo se obtienen las distancias entre los estímulos a partir de los datos originales, véase el capítulo sobre escalas multidimensionales en la obra de TORGENSON (1958, cap. 11). Una exposición completa del método métrico se puede ver en el libro que acabamos de mencionar o en BUGEDA (1973, 1974).



3. Procedimiento analítico

3.1. Regresión monótona de las (di)similaridades

Hemos indicado que la estrategia utilizada para calcular la representación espacial de los puntos va a consistir en el cálculo de unos valores, las disparidades, a partir de las (di)similaridades en los datos y las distancias entre los puntos en el espacio, moviendo posteriormente los puntos con objeto de hacer que esas distancias se aproximen a las disparidades. Puesto que en el método no-métrico los valores de las (di)similaridades no son cantidades aritméticas —sólo se usa la información contenida en el orden existente entre los estímulos—, las disparidades, que sí son métricas, nos sirven de referencia para ver en qué medida la solución espacial obtenida se ajusta a la información que tenemos de partida. Es decir, dada una información sobre el orden de las parejas de estímulos tenemos que encontrar una configuración euclídea de puntos (representativa de los estímulos) cuyas distancias resulten tan próximas como sea posible a las disparidades.

Estas disparidades se pueden obtener de dos formas: según el método de los «mínimos cuadrados», desarrollado por Kruskal; y según el método del *rank image*, debido a Guttman

Siguiendo a Kruskal, nuestros cinco estímulos generan 10 parejas de (di)similaridades, cuyo orden es el que aparece en la tabla 2.

TABLA 2

Matriz de (di)similaridades

		Egipto	España	Japón	México	Polonia
		1	2	3	4	5
Egipto	1	X				
España	2	3	X			
Japón	3	10	8	X		
México	4	1	2	7	X	
Polonia	5	5	6	9	4	X

Siendo las distancias entre los puntos,  $d_{ij}$ , las que aparecen en el gráfico 2, la forma de calcular las disparidades,  $\hat{d}_{ij}$ , sería la siguiente (tabla 3):

En tanto no se rompa la monotonía de la sucesión reproducimos como disparidades los valores de las distancias. Al llegar al punto noveno, el constituido por la pareja (5.3) vemos que la sucesión decrece de 8.1 a 6.9, por

lo que la monotonía se rompe. Para calcular las disparidades correspondientes a estos puntos tenemos que calcular la media de los dos pares, con lo cual el valor obtenido, 7.5, mantiene el criterio de monotonía. En la décima distancia vemos que de nuevo se rompe la monotonía; si como en el caso anterior calculásemos la media de las dos distancias (7.5 y 6.6), el valor obtenido, 7.05, al ser inferior a los 7.5 precedentes rompería el criterio de monotonía; por ello, lo que hacemos es calcular la media de los tres valores (8.1, 6.9 y 6.6), siendo este valor, 7.2, el que utilizamos como disparidad. De esta forma es como se obtuvo la línea de las disparidades en el gráfico 4.

TABLA 3

*Cálculo de las disparidades según el método de Kruskal*

(pareja)	$i, j$	4,1	4,2	2,1	5,4	5,1	5,2	4,3	3,2	5,3	3,1
[(di)similaridad]	$S_{ij}$	1	2	3	4	5	6	7	8	9	10
(distancia)	$d_{ij}$	0.8	1.2	1.6	1.8	2.5	2.6	7.0	8.1	6.9	6.6
(disparidad)	$\hat{d}_{ij}$	0.8	1.2	1.6	1.8	2.5	2.6	7.0	7.5	7.5	7.2
									7.05	7.05	7.2
									7.2	7.2	7.2

Según el método de Guttman (1968), y para los mismos datos, lo que se hace para obtener las disparidades es intercambiar los valores de las distancias,  $d_{ij}$ , con objeto de mantener el orden de disimilaridades,  $S_{ij}$  (tabla 4).

TABLA 4

*Cálculo de las disparidades según el método de Guttman*

(pareja)	$i, j$	4,1	4,2	2,1	5,4	5,1	5,2	4,3	3,2	5,3	3,1
(disimilaridad)	$S_{ij}$	1	2	3	4	5	6	7	8	9	10
(distancia)	$d_{ij}$	0.8	1.2	1.6	1.8	2.5	2.6	7.0	8.1	6.9	6.6
(disparidad)	$d_{ij}^*$	0.8	1.2	1.6	1.8	2.5	2.6	6.6	6.9	7.0	8.1

Una vez calculadas las disparidades, utilizamos estos valores para mover los puntos en el espacio con objeto de que las distancias entre los mismos sean lo más semejantes posible a las disparidades. Y el criterio empleado será el de

detenerse en el proceso de ver la diferencia entre las distancias y las disparidades, y mover los puntos con objeto de reducir esta diferencia, cuando la suma de los cuadrados de las diferencias sea mínima. A la cantidad obtenida en esta operación es lo que se llama el Stress.

Según la tabla 3, si queremos que las distancias sean una función monótona de las disimilaridades hemos de mover los puntos (3,2), (5,3) y (3,1) según nos indican las disparidades; ello se consigue haciendo que las distancias entre 3 y 2 (Japón y España) disminuya y aumenten las relativas a 5,3 (Polonia y Japón) y 3,1 (Japón y Egipto). Siguiendo estas indicaciones construiríamos una nueva representación, y para saber si era correcta volveríamos a repetir el procedimiento precedente: cálculo de las distancias, cálculo de las disparidades y cálculo del stress. El procedimiento concluiría cuando el Stress fuera pequeño.

### 3.2. El stress (medida de la bondad del ajuste)

A partir de los datos de la tabla 3, según Kruskal el stress sería

$$\sum (d_{ij} - \hat{d}_{ij})^2 = 1.26.$$

A partir de los datos de la tabla 4, según Guttman el stress sería

$$\sum (d_{ij} - \hat{d}_{ij})^2 = 3.86.$$

Estas fórmulas del Stress tienen algunas propiedades que no son convenientes. En el caso de que una solución particular se redujese en una cantidad constante,  $K$ , el Stress se reduciría también por un factor de  $K^2$ . Por tanto, la forma de minimizar el Stress consistiría en hacer la representación muy pequeña. Con objeto de evitar esta situación es necesario utilizar un elemento que mantenga el Stress invariante ante los cambios de escala en la representación espacial. De los diferentes factores posibles nos vamos a referir aquí a dos, definiendo según se use uno u otro el Stress 1 o el Stress 2.

$$\text{Stress 1} = \frac{\sum_i \sum_j (d_{ij} - \hat{d}_{ij})^2}{\sum_i \sum_j d_{ij}^2}$$

$$\text{Stress 2} = \frac{\sum_i \sum_j (d_{ij} - \hat{d}_{ij})^2}{\sum_i \sum_j (\overline{d} - d_{ij})^2}$$

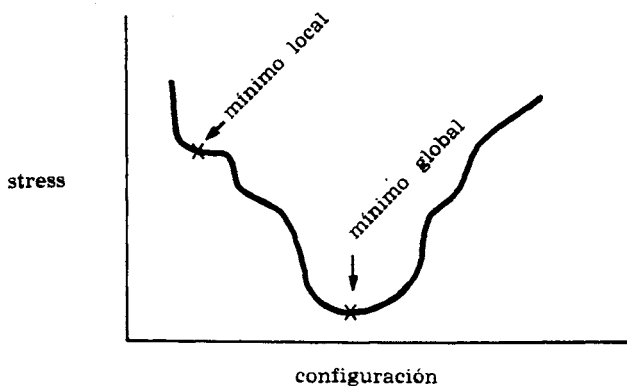
$\overline{d}$  = media de la distancia de todos los pares.

Estas fórmulas del Stress permiten medir la bondad del ajuste de la representación espacial con relación a las disimilaridades que teníamos como información de partida. Por tanto, si fijamos unos valores indicativos de lo que sería un buen ajuste, podemos saber cuándo terminar el proceso de movimiento de los puntos con objeto de llegar a la solución final<sup>6</sup>.

Una vez calculado el Stress estamos en condiciones de dar solución analítica: *a)* al problema de encontrar una representación espacial que respete las disimilaridades, y *b)* al procedimiento de mover los puntos en las sucesivas representaciones obtenidas, con objeto de ir disminuyendo el Stress. Analíticamente, el primer problema se resuelve encontrando una solución espacial (una configuración) cuyo Stress sea mínimo<sup>7</sup>. Con relación al segundo problema se puede pensar que cada configuración está representada por un único punto<sup>8</sup> y que el Stress mide el ajuste de la configuración con los datos originales; de este modo el conjunto de puntos se puede ver como una función del Stress (véase gráfico 5).

### GRAFICO 5

*Representación de los puntos-configuración como función del Stress*



El método utilizado por Kruskal<sup>9</sup> para pasar de una representación espacial a la siguiente, consiste en determinar la pendiente de la función,  $G_{ia}$ , con

<sup>6</sup> En el apartado 4.2 se proporcionan estos valores indicativos.

<sup>7</sup> En esta parte seguimos la exposición que sobre el tema hace COXON (1978; págs. 20-26).

<sup>8</sup> La configuración consiste en  $n \times r$  parámetros — las coordenadas de los  $n$  puntos en las  $r$  dimensiones. Un "espacio de configuraciones" se puede conceptualizar como un espacio  $n \times r$  dimensional donde cada parámetro actúa como una coordenada que fija una configuración particular como un punto en ese espacio.

<sup>9</sup> Se puede ver otro método diferente para modificar la configuración, desarrollado por GUTTMAN-LINGOES y ROSKAM, en COXON (1978; págs. 26-29).

respecto a la configuración actual; es decir, hallar la derivada del Stress con respecto a la configuración:

$$G_{ia} = d S / d X_{ia}$$

$S$  = valor del *stress*

$i$  = subíndice dado a los estímulos ( $i = 1, 2, \dots, n$ )

$a$  = subíndice dado a las dimensiones ( $a = 1, 2, \dots, r$ )

$X_{ia}$  = coordenadas del estímulo  $i$  en la dimensión  $a$

Si la configuración se mueve en la dirección de la pendiente negativa, tendrá que hacer más pequeño el Stress de la siguiente configuración. Cuando la pendiente sea cero, es que se ha alcanzado un mínimo<sup>10</sup> y el proceso se puede dar por concluido.

Una vez calculada la pendiente<sup>11</sup> el problema está en determinar cuánto hay que mover la configuración en la dirección de un Stress más pequeño. Si se elige una cantidad (*step size*) muy pequeña, se gastará más tiempo de cálculo; mientras que si movemos mucho la configuración, podemos pasarnos del mínimo global buscado. En la medida en que no hay una solución analítica a este problema, cada programa utiliza sus propias reglas.

Finalmente, una vez determinada la cantidad en que hay que mover la configuración y la pendiente, la fórmula para calcular las nuevas coordenadas es como sigue:

$$\text{Nueva } X_{ia} = \text{Vieja } X_{ia} - \frac{\text{Step size}}{\text{Norming factor}} \text{ Pendiente}$$

*Norming factor* es el tamaño relativo de la pendiente (comparada con otras pendientes).

*Step size* es el tamaño del movimiento hecho en la configuración.

<sup>10</sup> Puede tratarse de un mínimo "local" y no de un mínimo "global". La forma de determinar cuál es el caso consiste en ver, cuando la derivada parcial del Stress es cero, el tamaño del Stress. Si el Stress es "suficientemente pequeño" podemos decir que estamos ante un mínimo global, en caso contrario tendríamos un mínimo local.

<sup>11</sup> LINGOES y ROSKAM (1971; pág. 29) expresan la ecuación para su cálculo de la siguiente forma:

$$G_{ia} = \frac{d S}{d X_{ia}} = \frac{1}{\sum d_{ij}^2 S} \sum_j \left(1 - \frac{\hat{d}_{ij}}{d_{ij}} - S^2\right) (X_{ia} - X_{ja})$$

Tomado de COXON (1978; pág. 25).

3.3. *La elección de la configuración inicial*

En el ejemplo sobre los países decíamos que la configuración inicial se obtenía por el procedimiento de colocar los cinco puntos en un espacio euclídeo bidimensional de forma totalmente arbitraria. A partir de aquí íbamos moviendo los puntos con objeto de lograr una solución más adecuada (mínimo Stress). Esta no sería sino una de las posibles formas de dar comienzo el proceso, atribuible a Kruskal (1964, 120).

Alternativamente, Shepard (1962, 133) sugiere el uso de una «malla» regular, en la que todos los puntos estén separados entre sí por la misma distancia (véase gráfico 6), y a partir de aquí comenzar los sucesivos movimientos de los puntos hasta alcanzar la solución final.

GRAFICO 6

*Ejemplo de configuración inicial atribuible a Shepard*

X	X	X	X	X	X
X	X	X	X	X	X
X	X	X	X	X	X
X	X	X	X	X	X

Un tercer método consiste en tratar la información original según el método de análisis multidimensional clásico de Torgerson<sup>12</sup>. De esta forma, la solución del análisis clásico puede servir como configuración inicial del análisis no métrico.

<sup>12</sup> En el análisis multidimensional clásico, siguiendo el ejemplo de Torgerson sobre la percepción de los colores, se trata de obtener, a partir de los juicios de similitud entre los colores, y mediante el método completo de las triadas, las distancias entre los mismos. Estas distancias son *comparativas*, referidas a un origen arbitrario, por lo que no pueden ser utilizadas en un modelo especial ya que la configuración estaría en función del cero arbitrario que seleccionásemos. Con objeto de superar este problema hay que transformar estas distancias *comparativas* en *absolutas*. Para ello hay que encontrar una constante, *c*, en la ecuación

$$d_{ij} = h_{ij} + c$$

$d_{ij}$  = distancia absoluta entre los estímulos ( $i, j = 1, 2 \dots, n$ )

$h_{ij}$  = distancia comparativa entre los estímulos ( $i, j = 1, 2 \dots, n$ )

que permita la representación de los estímulos en un espacio Euclideo de la menor dimensionalidad posible.

La matriz de distancias absolutas,  $D^0$ , se transforma en una matriz de productos escalares,  $B^*$ , referidos al "centroide" de todos los puntos. Después esta matriz se somete a un análisis factorial, con objeto de obtener la matriz de coor-

Cuando en el análisis clásico las *distancias comparativas* entre los estímulos no tengan error y la *constante aditiva* que se utiliza para calcular las *distancias absolutas* haya sido bien elegida, la solución final, tras el *scaling* no-métrico, no ganará mucho con relación a la configuración inicial —es decir, nos podíamos haber ahorrado el análisis no-métrico. Pero en el caso contrario, el análisis no-métrico mejorará la configuración inicial (la solución del análisis clásico).

Es importante la elección de la configuración inicial, pues una mala configuración causaría:

1. Un uso prolongado del ordenador debido a la necesidad de efectuar un mayor número de iteraciones hasta lograr una solución satisfactoria.
2. Una mayor probabilidad de encontrar una solución que sea un «mínimo local» (véase gráfico 5).

Estratégicamente es preferible, en general, la utilización de una configuración inicial racional y no elegida al azar.

Como resumen de esta parte, y antes de entrar en los problemas que se plantean a la hora del «diagnóstico» y de la interpretación de la solución obtenida, conviene considerar el siguiente esquema-resumen preparado por Charles Jones (1979):

### 3.4. Resumen del procedimiento

Dados unos datos en forma de una matriz cuadrada y simétrica con  $\frac{1}{2} N(N-1)$  (di)similaridades (una para cada par de estímulos), hay una «solución espacial» en la cual los estímulos son puntos con coordenadas en un pequeño número de dimensiones.

Ahora bien, la distancia entre los puntos (estímulos) en esta «solución espacial» está relacionada con sus correspondientes (di)similaridades por:

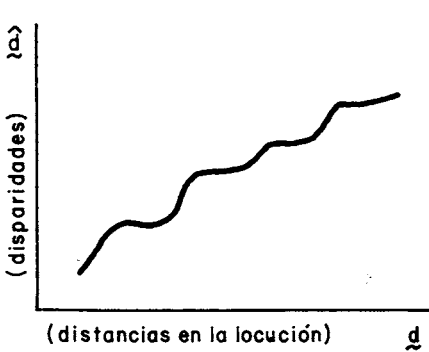
- a) Una función lineal (en el análisis métrico).
- b) Una función monótona (en el análisis no-métrico).

---

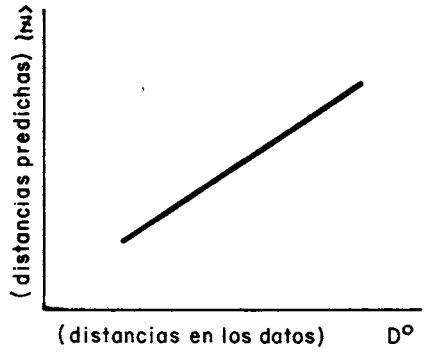
denadas  $X$  de los estímulos en las diferentes dimensiones o ejes. Resumiendo:  
Distancia en los Datos  $D^0 \rightarrow$  productos escalares  $B^*$   $\rightarrow$  cálculo  $\rightarrow$   $X$  (config.)  
valores  
y vectores  
propios

Para una exposición detallada véase TORGESON (1958) y BUGEDA (1973 y 1974).

---

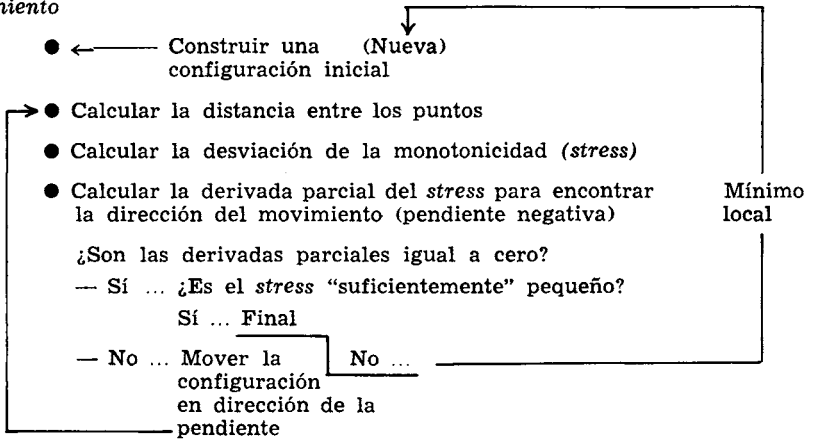


Análisis No-Métrico



Análisis Métrico

Procedimiento



4. El problema de la interpretación

Hasta aquí nos hemos limitado a ofrecer el modelo básico del análisis multidimensional. Siguiendo los pasos indicados se llegaría a una solución que sería preciso interpretar. Tres problemas surgirían entonces.

En primer lugar, con objeto de facilitar la comprensión hasta ahora hemos recurrido a un ejemplo en el que se utilizaba un espacio bidimensional. No siempre será éste el caso y en cada análisis habrá que determinar el número de dimensiones necesarias para representar la configuración de estímulos.

Por otra parte, el número de dimensiones está relacionado con el «error» que estemos dispuestos a cometer. Habida cuenta que todo conjunto  $n$  de objetos se puede representar en un espacio  $n$  dimensional, nosotros lo que



pretendemos es ganar en la interpretación reduciendo ese número de dimensiones, aun a costa de forzar algo la representación. Ese «error» vendrá determinado por el Stress. El límite del valor que puede alcanzar el Stress será otro tema que habrá que examinar.

Por último, una vez determinado el número de dimensiones, hay que interpretar su significado con objeto de dar un sentido al análisis; éste será el último punto que consideraremos.

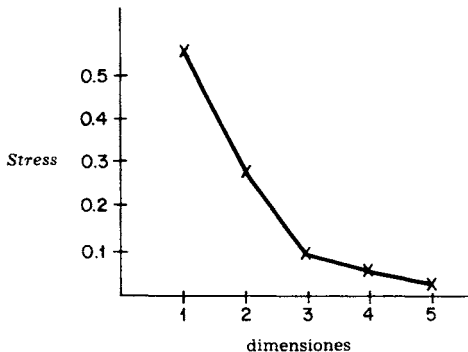
#### 4.1. *El número de dimensiones*

Como señala Shepard (1972, 9), el número de dimensiones que se pueden elegir en un análisis multidimensional constituye un punto de equilibrio entre la bondad del ajuste y la visualización y «economía» (parsimonia) obtenida con la solución. Cuanto mayor es el número de dimensiones mejor es el ajuste y por tanto menor es el *stress*; pero al mismo tiempo, cuando aumenta el número de dimensiones se pierde la ventaja que tiene el análisis multidimensional para representar gráficamente la estructura subyacente a un conjunto de datos, ya que resulta imposible de *ver* una solución en más de dos o tres dimensiones. Y no sólo eso, sino que también se pierde el interés del análisis basado en su capacidad para reducir la información a un número de dimensiones menor que el de los datos disponibles.

El valor obtenido por el Stress en una solución con diferentes dimensiones se puede utilizar para fijar el número de dimensiones más apropiado al caso. Si representamos gráficamente el valor del Stress para un número determinado de dimensiones (véase gráfico 7), se puede tomar como solución el número de dimensiones a partir del cual la curva deja de decrecer marcadamente.

GRAFICO 7

*Valor del Stress según el número de dimensiones*



En el gráfico 7 se puede ver que el número de tres dimensiones sería el apropiado. Sin embargo, este criterio habría que relacionarlo con los niveles del *valor* del Stress para los tres dimensiones, lo que consideraremos al ocuparnos del tamaño del Stress.

Existe igualmente una tabla, diseñada por Spence y Graef, con los valores *esperados* del Stress para matrices de datos generadas en dimensiones verdaderas (de una a cinco dimensiones), con cinco diferentes niveles de error y para un número de puntos que va de 12 a 36. En el análisis, el valor del *stress* correspondiente a las soluciones de una a cinco dimensiones se compara con estas tablas, pudiendo determinarse el número más correcto de dimensiones de la solución y el porcentaje de error. Cada uno de los valores *esperados* del Stress representa la media de un número de valores estimados en diferentes estudios de Monte Carlo. Toda la tabla está incorporada en una subrutina llamada MSPACE, que se incluye en algunas versiones del programa numérico MINISSA. Para una descripción del método véase I. Spence y J. Graef (1974) y I. Spence y F. W. Young (1978).

En el análisis de los datos de los cinco países hemos utilizado el programa MINISSA, incorporado en el paquete MDS (X), obteniendo las posibles soluciones en espacios de una a cinco dimensiones. Debido al pequeño número de estímulos que utilizamos, a partir de dos dimensiones el *stress* se hace igual a cero, por lo que elegimos este espacio bidimensional como solución final.

El programa necesario para el análisis consta de una serie de sencillas instrucciones, que incluimos a continuación:

TABLA 5

*Sentencias del programa MINISSA*

```

RUN NAME      J A V I E R
N OF STIMULI  5
PRINT DATA   YES
INPUT FORMAT   (4F2.0)
PARAMETERS    MIND (1), MAXD (5), DATA (1), OUTPUT (1)
READ MATRIX
    
```

INPUT MATRIX

```

ROW 2  70.00000
ROW 3  95.00000  39.00000
ROW 4  40.00000  50.00000  80.00000
ROW 5  75.00000  80.00000  90.00000  70.00000
    
```

COMPUTE

Entre otra información el *output* del análisis proporciona las coordenadas de la solución, las distancias entre los puntos y su representación espacial. La tabla 6 y el gráfico 8 muestran esta información (aplazamos la interpretación de los resultados hasta el apartado 4.3).

TABLA 6

a) *Coordenadas de los cinco puntos en dos dimensiones.* b) *Distancias entre los cinco puntos*

<i>Final configuration</i>		
	<u>1</u>	<u>2</u>
1	-0.6479	-0.1689
2	-0.3494	-0.6836
3	1.5342	-0.1296
4	-0.3512	-0.1988
5	-0.1856	1.1809
MEAN	-0.0000	0.0
SIGMA	0.7815	0.6240

a)

<i>Distances</i>				
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>
2	0.5950			
3	2.1824	1.9634		
4	0.2981	0.4849	1.8357	
5	1.4267	1.8717	2.1622	1.3896

b)

#### 4.2. *Tamaño del Stress*

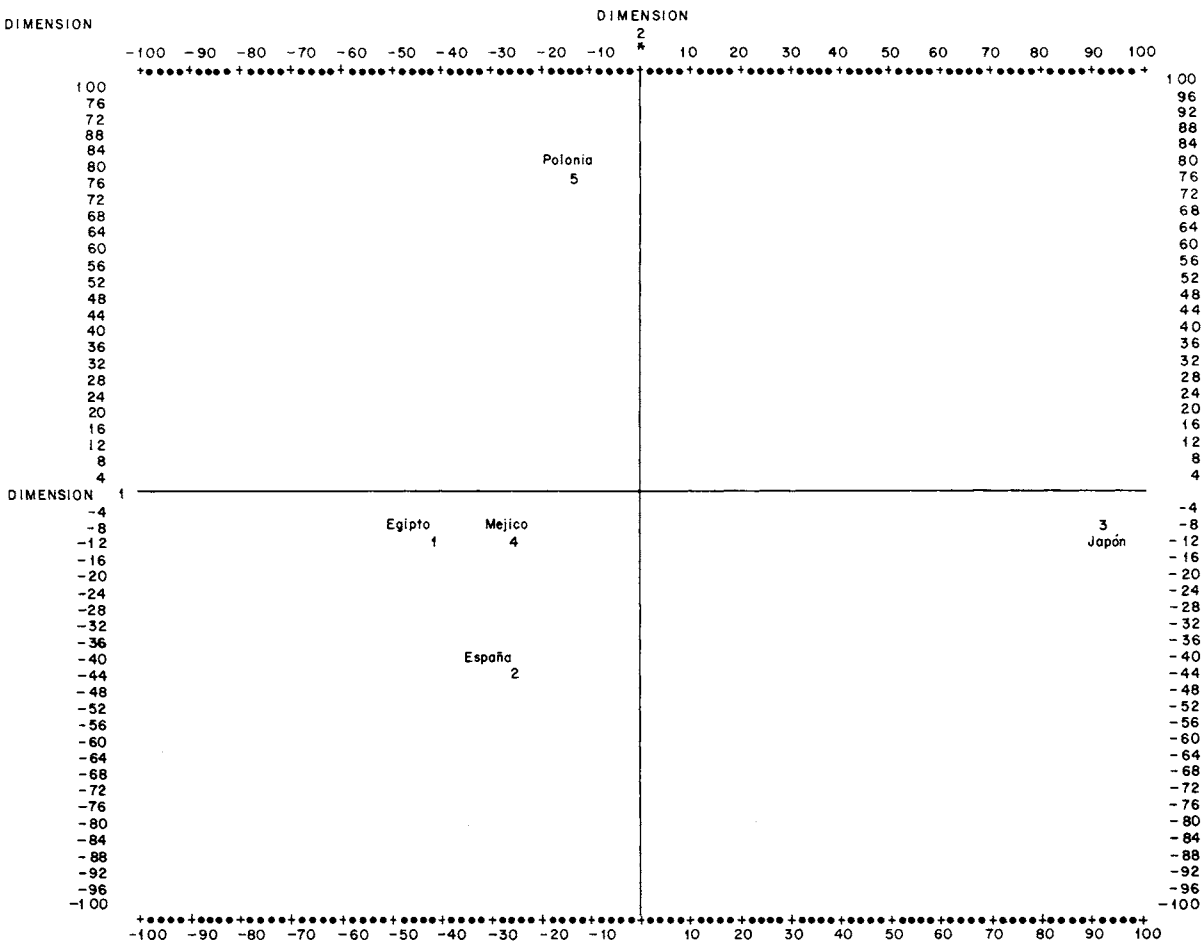
Tal como señalábamos, el Stress mide la bondad del ajuste obtenido (la linealidad o monotonicidad, según se trate de un análisis métrico o no-métrico), entre las (di)similaridades de los datos y las distancias entre los puntos en la solución. Igualmente afirmábamos que el Stress debería ser «suficientemente» pequeño. Con objeto de poder utilizar un criterio para determinar los límites aceptables que puede alcanzar el Stress, Kruskal ofreció la siguiente tabla:

#### *Stress*

	0.01 = Excelente
0.01 a 0.05	= Bueno
0.05 a 0.10	= Correcto
0.10 a 0.15	= Moderado
0.15	= Pobre

GRAFICO 8

*Representación de los puntos en dos dimensiones*



La tabla tiene validez tratándose del análisis no-métrico y utilizando el Stress 1. Es decir, la tabla no tiene una utilidad general, pues existe toda una serie de factores que condicionan el tamaño del Stress, haciendo que su valor sea relativo. Como factores que determinan que el Stress sea más pequeño, podemos considerar:

- El uso del análisis no-métrico (tiene un Stress inferior) frente al análisis métrico.
- El uso de la  $d$  (monotonicidad débil) ofrece un Stress inferior al uso de  $d'$  (monotonicidad fuerte).
- El enfoque «primario» de los «empates» (*ties*) frente al enfoque «secundario»<sup>13</sup>.
- El Stress 1 frente al Stress 2.
- Utilizando un número elevado de dimensiones el Stress es menor.
- En un análisis con un número pequeño de puntos también el Stress es menor que en el caso de que existan muchos puntos. Sin embargo, esta consideración hay que relacionarla con la *recomendación de utilizar, al menos, doce puntos en el análisis*; y ello porque es necesario que exista un número mínimo de puntos a fin de que las constricciones impuestas sean suficientes para que la información ordinal pueda determinar un resultado métrico.

<sup>13</sup> Digamos brevemente que uno de los problemas que se plantean en el análisis no métrico es la existencia de parejas de puntos con (di)similaridades y/o disparidades iguales (con empates). Esto plantea problemas, pues la única información de que disponemos para calcular la solución es el orden entre las (di)similaridades y entre las disparidades, y este orden lo perdemos por el hecho de que hay varias (di)similaridades y disparidades que son iguales.

$i, j$	4,1	4,2	5,4	2,1	5,1	5,2	4,3	3,2	5,3	3,1
$S_{ij}$	1	2	3	3	5	6	6	8	9	10
$d_{ij}$	.8	1.2	1.8	1.6	2.5	2.6	7.0	8.1	6.9	6.8
$\hat{d}_{ij}$	.8	1.2	1.7	1.7	2.5	2.6	7.0	7.2	7.2	7.2

Tomando los datos originales (tabla 1) y representándolos en la tabla anterior, según el enfoque "primario" para el caso de las disparidades iguales a 1.7, la solución consistiría en cambiar el 1.7, 1.7 por 1.6 y 1.8; o lo que es lo mismo, alternar el orden de los puntos (5,4 y 2,1).

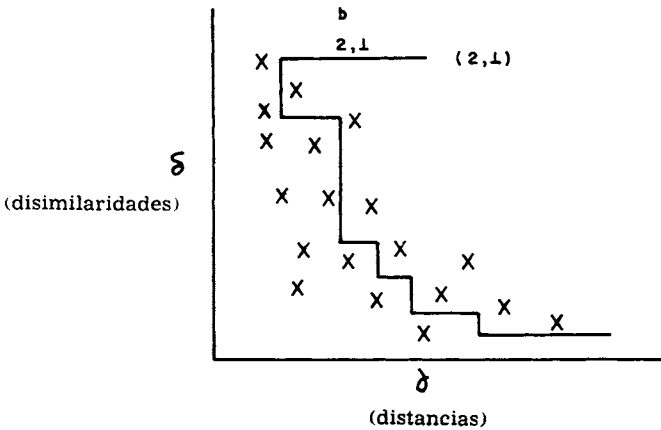
En el enfoque "secundario" dejaríamos los valores sin cambiar. En el enfoque primario, pues, se rompen los empates (*ties*) con objeto de reducir el *stress*; mientras que en el "secundario" se dejan los empates tal como están, con lo cual se tiende a aumentar el *stress*. En cualquier caso, cuando existen muchos empates (*ties*) en los datos es preciso el análisis métrico y el uso del enfoque "secundario"; y en última instancia rechazar los datos.

El programa que hemos utilizado, MINISSA, hace uso del enfoque primario; por esa razón ya de partida, en el gráfico 3, hemos colocado el punto 2,1 antes que el 5,4. Igual argumento sirve para justificar la colocación del punto 5,2 antes que el 4,3.

Estas consideraciones sobre el Stress hacen ver que antes de interpretar la solución obtenida es conveniente examinar la «calidad» de la solución; a esta operación es a lo que Coxon llama el «diagnóstico». Se trataría, en definitiva, de hacer lo mismo que en cualquier análisis de la regresión: no basta con ver los coeficientes de regresión, sino que hay que mirar la posible existencia de heteroscedasticidad, multiconinealidad y/o autocorrelación de los datos, factores que son los que en definitiva justificarán la validez del análisis. En este sentido, Coxon sugiere también que antes de interpretar la solución obtenida en el análisis multidimensional se haga un diagrama de dispersión, para ver el ajuste de los puntos a la línea de regresión (monótona en nuestro caso).

GRAFICO 9

Diagrama de Shepard



A simple vista, la existencia de algunos puntos, como el 2.1 (véase gráfico 9), nos puede sugerir que parte del Stress está causado por la contribución de estos puntos, que una vez suprimidos pueden hacer que el tamaño del Stress sea aceptable<sup>14</sup>.

Tal como indicábamos en el apartado 4.1, en nuestro ejemplo el Stress tiene un valor de cero a partir de dos dimensiones (véase tabla 7).

Sin embargo, siempre que se utilice un número tan pequeño de puntos hay que ser precavido a la hora de observar el valor del Stress, pues la interpretación de este valor depende del número de estímulos,  $I$ , y de la dimen-

<sup>14</sup> Analíticamente, se puede detectar la existencia de estos puntos «malos» calculando el valor de los residuos para cada punto (véase gráf. 9)  $(d_{ij} - d^0_{ik})$ , donde  $d_{jk}$  puede ser  $\hat{d}$  o  $d^*$ .

Aquellos puntos que tienen un residuo grande son los que contribuyen principalmente a que el tamaño del stress también sea grande.

TABLA 7.

Valores de Stress utilizando Stress 1 y d (DSTAR)

METRIC IS EUCLIDEAN	
DSTAR IS USED	
ITER GRADIENT ANGLE	
1	
2 DIMENSIONS	
ANGLE FACTOR IS USED	
RAW STRESS DHAT STRESS DHAT	
0.0	0.0
SFORM 1 SOFT SQUEEZE	
RAW STRESS DSTAR COEF. ALIEN DSTAR	
0.0	0.0

sionalidad,  $R$ , de la solución. Kruskal y Wish (1978) indican que sólo a partir de que  $I > 4R$  cabe decir que la interpretación del Stress es independiente de  $I$  y de  $R$ . Como ejemplo, los autores mencionados señalan que en una solución de siete objetos en tres dimensiones cabe esperar que datos generados al azar lleven a un Stress tan pequeño como 0.02 en un 50 por 100 de las ocasiones (A. C., p. 52). Por tanto, reiteramos que las consideraciones previas sobre el Stress son válidas siempre y cuando  $I$  sea relativamente grande respecto de  $R$ .

Una vez determinado el número de dimensiones y realizado el «diagnóstico» de la solución obtenida, se estaría en condiciones de pasar a su interpretación.

#### 4.3. Interpretación de la configuración

Para determinar el significado de las diferentes dimensiones pueden utilizarse dos métodos diferentes. Según el primero, y a semejanza con la interpretación tradicional en el análisis factorial, se trataría de ver los *cluster* de puntos y, tras rotar los ejes según algún criterio<sup>15</sup>, determinar lo que «parece» más probable que signifiquen las dimensiones. En la solución del análisis multidimensional los ejes obtenidos para localizar los puntos son arbitrarios, ya que cualquier rotación deja invariantes las distancias entre los puntos. Por ello, es perfectamente factible rotar los ejes o incluso reflejarlos, sin que se altere la solución.

<sup>15</sup> El criterio utilizado será hacer que los ejes pasen por los *cluster* de puntos de modo que las coordenadas sean grandes o, alternativamente, próximas a cero — el equivalente al criterio de “estructura simple” (*simple structure*) en análisis factorial.

La necesidad de rotar los ejes es mayor cuando la solución viene dada en un número elevado de dimensiones, ya que en este caso se hace difícil la interpretación de los resultados a simple vista.

Un segundo método, que permite interpretar las dimensiones con mayor garantía, está basado en la utilización de algún criterio externo a la propia configuración. Junto a la recogida de la información pertinente para el análisis multidimensional se trata de obtener información sobre la puntuación de los estímulos respecto a una o varias propiedades. Estas propiedades posteriormente se representarán en la configuración mediante nuevos ejes, comprobando si las proyecciones de los puntos sobre éstos son similares o diferentes a las proyecciones de los puntos sobre el eje que queremos interpretar. En caso afirmativo, el eje interno (dimensión obtenida en el análisis multidimensional) se puede interpretar como la propiedad representada por el eje externo. En caso de que no coincidan, la configuración se puede interpretar en función del eje de las propiedades (olvidándonos de las dimensiones o tratando de encontrar su interpretación por otro camino).

Supongamos que el resultado del análisis multidimensional para  $M$  puntos es una matriz,  $X$ , de coordenadas en un espacio de  $r$  dimensiones.

Si tenemos las puntuaciones para cada uno de los  $M$  puntos respecto a alguna propiedad,  $P$ , se trataría de ver cómo se relaciona ésta con  $X$ . En términos analíticos se trata de un problema de regresión múltiple, donde la propiedad es la variable dependiente y las columnas  $X$  son las «variables independientes».

La ecuación de regresión vendría dada por

$$\text{propiedad} = b_1 e_{j1} + b_2 e_{j2} + \dots + b_r e_{jr} + \text{error}$$

donde las  $b$  son los coeficientes de regresión no estandarizados.

En la representación, la pendiente del nuevo eje de la propiedad vendría determinado por el cociente  $b_2/b_1$ , en el caso de que sólo hubiera dos dimensiones (véase gráfico 10)<sup>16</sup>; y la bondad del ajuste estaría en función del coeficiente de correlación múltiple<sup>17</sup>.

El número de propiedades a utilizar es opcional; en líneas generales hay dos alternativas: se puede utilizar un número de propiedades semejante al número de dimensiones y tratar de verificar alguna hipótesis sobre el signi-

<sup>16</sup> En general, la dirección de los ejes externos vendrá determinada por el coseno direccional, cuando los coeficientes están normalizados; es decir, cuando  $\sum_i b_i^2 = 1$ . En esta circunstancia el  $\cos \theta_i = b_i$ .

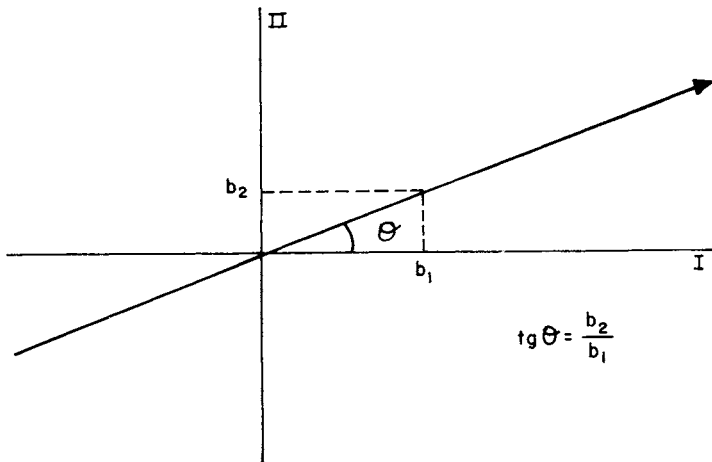
<sup>17</sup> Sólo en caso de existir una correlación elevada está justificado utilizar la propiedad como ayuda para la interpretación de la solución.



ficado de estas dimensiones<sup>18</sup>, o bien se puede utilizar un número superior de propiedades con una intención exploratoria<sup>19</sup>.

GRAFICO 10

*Dirección del eje de la propiedad*



Este análisis de las propiedades se puede hacer mediante cualquier programa de regresión múltiple o bien utilizando el programa PROFIT (*Property Fitting*), diseñado a tal efecto.

En nuestro ejemplo, junto con la información sobre las semejanzas entre los países se han puntuado los cinco países en dos escalas diferentes, ambas con un recorrido de 0 a 10 puntos. Una escala medía la riqueza de los países (RPC), de acuerdo con la opinión de los entrevistados, y otra su alineamiento con los bloques (USA y la URSS). En el primer caso un valor de cero significaba que el país en cuestión era considerado como «muy pobre» y diez, «muy rico»; y en el segundo, cero significaba total alineamiento con los Estados Unidos de América y diez total alineamiento con la Unión Soviética. Las puntuaciones de los cinco países en las dos escalas fueron las siguientes:

<sup>18</sup> Un ejemplo de este tipo se ofrece en el mencionado artículo de RABINOWITZ (1975), tratando de interpretar las dos dimensiones del análisis de las similitudes entre los políticos en términos de "afiliación a un partido" y "corriente ideológica". En el caso extremo, una superposición de los ejes internos (las dimensiones) y externos (las propiedades) permitiría interpretar las dimensiones en términos de esas propiedades.

<sup>19</sup> El estudio de SHERMAN y DOWDLE (1974), en el que se utilizan hasta un total de 32 propiedades, de las que 10 quedan como significativas, es un buen ejemplo de este tipo de análisis.

	Riqueza	Alineamiento
Egipto ... ..	3	3
España ... ..	5	1
Japón ... ..	9	1
México ... ..	5	3
Polonia ... ..	6	8

Utilizando la riqueza como variable dependiente se calcularon los coeficientes de regresión múltiple para la ecuación

$$\text{riqueza} = a + b_1 (\text{dimensión I}) + b_2 (\text{dimensión II}) + \text{error}$$

Igual se hizo cuando la escala de alineamiento con los bloques era la variable dependiente:

$$\text{alineamiento} = a + b_1 (\text{dimensión I}) + b_2 (\text{dimensión II}) + \text{error}$$

Los resultados obtenidos se reflejan en la tabla 8.

TABLA 8

*Coefficientes de las regresiones múltiples de las escalas en las dimensiones*

Escalas	Coef. de regresión				Coef. de correlación múltiple
	No estandarizados		Normalizados (cosenos direccionales)		
	DIM. I	DIM. II	DIM. I	DIM. II	
Riqueza ... ..	2.3733	.5137	.9774	.2114	.960
Alineamiento ... ..	-1.08	3.8686	-.269	.963	.998

La dirección de los ejes correspondientes a las escalas viene determinada por los cocientes  $b_2/b_1$  de cada ecuación, o por los  $\cos \theta_i = b_i$  (donde  $i$  es el coeficiente de regresión normalizado de mayor valor en cada ecuación). El eje riqueza forma un ángulo de 12.2 grados con el eje horizontal:

$$\text{tag} \theta_1 = b_2/b_1 = .5137/2.3733 = .216$$

$$\text{arc tag } .216 = 12.2 \text{ grados}$$

$$\text{o } \cos \theta_1 = .9774$$

$$\text{arc cos } .9774 = 12.2 \text{ grados}$$

El eje alineamiento forma un ángulo de 15.6 grados con el eje vertical (dimensión II):

$$\text{tag}\theta_2 = b_2/b_1 = 3.8686/-1.08 = 3.582$$

$$\text{arc tag } 3.582 = 74.4 \text{ grados; y su complemento } 15.6$$

$$\text{o } \cos\theta_2 = .963$$

$$\text{arc cos } .963 = 15.6 \text{ grados}$$

El gráfico 11 muestra la localización de las dos escalas en el plano definido por las dos dimensiones.

Puesto que los coeficientes de correlación múltiple tienen un valor muy elevado (.960 y .998 para la primera y la segunda escala, respectivamente), lo que significa que las coordenadas predicen bien los valores de las escalas, podemos interpretar las dimensiones obtenidas en el análisis como la «riqueza» y el «alineamiento» de los países. Nuestra percepción de los países en cuestión, sin coincidir exactamente con la de los entrevistados, parece confirmar los resultados del análisis: en el eje horizontal tenemos a Egipto como país más pobre y a Japón como país más rico; en el eje vertical, por un lado aparece Polonia, país perteneciente al bloque de los países socialistas, y por el otro España y Japón, países ambos ligados a los Estados Unidos.

Complementariamente a la interpretación de nuestros resultados en términos de dimensiones, se pueden observar los grupos de países que se forman, interpretándolos en línea con el análisis de *cluster*. Así, en nuestro análisis se aprecia la existencia de cuatro grupos o *clusters* de puntos: el primero constituido por Egipto y México, y los otros tres formados por Polonia, Japón y España, respectivamente. La interpretación de estos grupos se haría apoyándonos en el conocimiento que tenemos del significado de las dimensiones.

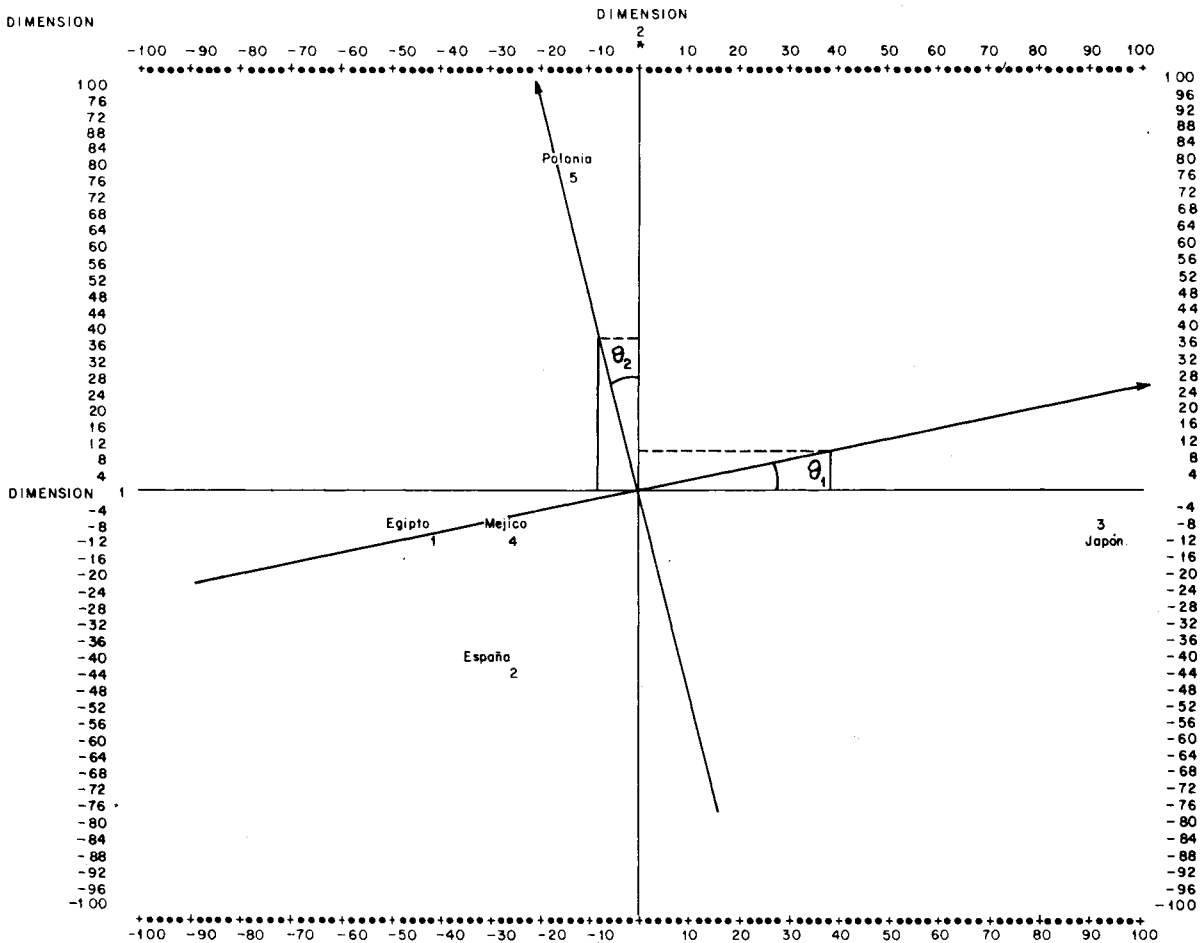
### *Programas de Ordenador*

La mayoría de los programas informáticos de escalas multidimensionales se encuentran en la serie MDS(X), desarrollada por un equipo de la Universidad de Edimburgo y del University College de Cardiff. En esta serie se halla la siguiente relación de programas:

- CANDECOMP (CANonical DECOMPosition).
- HICLUS (HIerarchical CLUString).
- INDSCAL-S (INDIvidual Differences SCALing).
- MDPREF (MultiDimensional PREFerence Scaling).
- MINICPA (Michigan-Israel-Nijmegen Integrated Series: Conditional Proximity Analysis).

GRAFICO 11

*Representación de los ejes de las propiedades*



- MINIRSA (MINI Rectangular Smallest Space Analysis).
- MINISSA (Michigan-Israel-Nijmegen Interated Smallest Space Analysis).
- MRSCAL (MetRic SCAling).
- MVNDS (Maximun Variance Non-Dimensional Scaling).
- PARAMAP (PARAMetric MAPping).
- PINDIS (Procustrean INdividual DIfferences SCaling).
- PREFMAP (PREFErrence MAPping).
- PROFIT (PROperty FITting).
- TRISOSCAL (TRIadic Similarities Ordinal SCAling).
- UNICON (UNIdimensional CONjoint measurement).

La serie está implementada en el centro de procesos de datos de la Universidad Complutense de Madrid.

#### BIBLIOGRAFIA

- BUGEDA, José: "La medida multidimensional de la opinión", *Revista Española de la Opinión Pública*, 34, pp. 37-66, 1973.
- *La medida de las ciencias sociales*, Madrid, Confederación Española de las Cajas de Ahorro, 1974.
- COXON, A. P. M.: *Multidimensional Scaling*, ECPR Summer School Monograph, Univ. de Essex (Inglaterra), 1975.
- *Elementary Introduction to Multidimensional Scaling MDS (X) Project*, Draft Paper, Univ. de Cardiff (Gran Bretaña), 1978.
- *The user's guide to multidimensional scaling*, Londres, Heinemann, 1982.
- y C. L. JONES: *The images of occupational prestige*, Londres, Macmillan, 1978.
- y C. L. JONES: "Multidimensional scaling", en D. MCKAY, N. SCHOFIELD y P. WHITLEY (eds.): *Data analysis and the social sciences*, Londres, France Pinter, pp. 171-225, 1983.
- y C. L. JONES: "EscalaS multidimensionales", en SÁNCHEZ CARRIÓN (ed.): *Técnicas de análisis multivariable en las ciencias sociales*, Madrid, CIS, 1984.
- GUTTMAN, Louis: "A general non metric technique for finding the smallest coordinate space for a configuration of points", *Psychometrika*, 33, 469-506, 1968.
- JONES, C. L.: *Multidimensional scaling*, notas del curso sobre el mismo tema de la Universidad de Essex (Inglaterra), 1979.
- KRUSKAL, J. B., y M. WISH: *Multidimensional scaling*, Londres y Beverly Hills, Sage, 1978.
- "Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis", *Psychometrika*, 29, 1-27, 1964.
- LINGOES, J. C., y E. E. ROSKAM: "A mathematical and empirical analysis of two multidimensional scaling algorithms", *Psychometrika*, suplemento monográfico, pp. 1-93. Reimpreso con correcciones en LINGOES y otros (eds.): *Geometric representations of relational data*, Ann Arbor, Matheis Press, pp. 371-473, 1973.
- MARTÍNEZ, M.: "EscalaS multidimensionales para el análisis de atributos", *CUPEMA*, vol. 1, núm. 11, pp 29-39, 1975.
- MAC DONALD, K. I.: "MDSCAL and distances between socio-economic groups", en HOPE, K. (ed.): *The analysis of social mobility*, Oxford, Clarendon Press, 1972.

- RABINOWITZ, G. B.: "An introduction to non-metric multidimensional scaling", *American Journal of Political Science*, vol. 19, pp. 343-390, 1975.
- SHEPARD, R. N.; A. K. ROMNEY y S. B. NERLOVE: *Multidimensional scaling: theory and applications in the behavioral sciences*, vol. 1: teoría; vol. 2: aplicaciones. Londres, Seminar Press, 1972.
- SHERMAN, R. C., y M. D. DOWDLES: "The perception of crime and punishment: a multidimensional scaling analysis", *Social Science Research*, vol. 3, pp. 109-126, 1974.
- SPENCE, I., y J. GRAEF: "The determination of the underlying dimensionality of an empirically obtained matrix of proximities", *Mult. Behav. Res.*, vol. 9, número 3, pp. 331-341, 1974.
- SPENCE, I., y F. W. YOUNG: "Monte-Carlo studies in non-metric scaling", *Psychometrika*, vol. 43, núm. 1, pp. 115-117, 1978.
- TORGENSON, W. A.: *Theory and methods of scaling*, Nueva York, Willey and Sons, 1958.
- VICENS, José: "Análisis multidimensional no-métrico en Marketing", *CUPEMA*, vol. 3, núm. 1, pp. 83-120.

---

# NOTAS DE INVESTIGACION