

Análisis de Multidimensional Scaling

Medidas de similitud

En realidad, es bastante subjetivo el hecho de elegir una medida de similitud ya que depende de las escalas de medida. Se pueden agrupar observaciones según la similitud expresada en términos de una distancia. Si se agrupan variables, es habitual utilizar como medida de similitud los coeficientes de correlación en valor absoluto. Para variables categóricas existen también criterios basados en la posesión o no de los atributos (tablas de presencia-ausencia).

Dados dos vectores $\mathbf{x}_i, \mathbf{x}_j$ pertenecientes a \mathbb{R}^k , diremos que hemos establecido una distancia entre ellos si definimos una función d con las propiedades siguientes:

1. $d : \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}^+$, es decir $d(\mathbf{x}_i, \mathbf{x}_j) \geq 0$;
2. $d(\mathbf{x}_i, \mathbf{x}_i) = 0 \quad \forall i$, la distancia entre un elemento y sí mismo es cero.
3. $d(\mathbf{x}_i, \mathbf{x}_j) = d(\mathbf{x}_j, \mathbf{x}_i)$, la distancia es simétrica
4. $d(\mathbf{x}_i, \mathbf{x}_j) \leq d(\mathbf{x}_i, \mathbf{x}_p) + d(\mathbf{x}_p, \mathbf{x}_j)$, la distancia verifica la propiedad triangular.

Estas propiedades generalizan la noción intuitiva de distancia euclídea entre dos puntos.

Ejemplos de distancias entre objetos

Distancia euclídea

Dados dos objetos I_1 y I_2 medidos según dos variables x_1 y x_2 , la distancia euclídea entre ambos es:

$$d_{I_1 I_2} = \sqrt{(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2}.$$

Con más dimensiones (o variables que se miden) es equivalente a:

$$d_{I_1 I_2} = \sqrt{\sum_{k=1}^p (x_{1k} - x_{2k})^2}$$

En notación vectorial se expresa como

$$d_{I_i I_j}^2 = (x_i - x_j)'(x_i - x_j).$$

Si se consideran n objetos para $i, j \in \{1, \dots, n\}$, la distancia total es

$$\mathbf{d} = \sum_{i=1}^n \sum_{j=1}^n \left(\sum_{k=1}^p (x_{ik} - x_{jk})^2 \right)^{1/2}.$$

Distancia de Minkowski

$$d_{I_i I_j} = \left[\sum_k |x_{ik} - x_{jk}|^m \right]^{1/m}$$

donde $m \in \mathbb{N}$.

Si $m = 1$, se tiene la distancia en valor absoluto y si $m = 2$, la euclídea.

Distancia de Mahalanobis

Se define como

$$d_{I_i I_j}^2 = (x_i - x_j)'W^{-1}(x_i - x_j)$$

donde W es la matriz de covarianzas entre las variables. De este modo, las variables se ponderan según el grado de relación que exista entre ellas, es decir, si están más o menos correlacionadas. Si la correlación es nula y las variables están estandarizadas, se obtiene la distancia euclídea.

Ejemplos de distancias entre variables

Coefficiente de correlación de Pearson

Se define como:

$$r = \frac{S_{xy}}{S_x S_y}$$

donde S_{xy} es la covarianza muestral entre x e y , S_x y S_y son las desviaciones estándar de x e y respectivamente.

Coefficiente de correlación de rangos de Kendall

Se comparan las ordenaciones que dan dos variables, es decir, los datos se ordenan según dos criterios o características y se establece el número de concordancias y discordancias.

Método:

1. Calculo todas las posibles parejas. Tomo una pareja (i, j) . Si están ordenados igual según las dos variables o criterios, se marca una *concordancia* (es decir, si el elemento i está delante del elemento j según ambas variables o criterios). Si no lo están, se establece una *discordancia*.
2. El número total de parejas distintas que se pueden hacer con n elementos es $\binom{n}{2} = \frac{n(n-1)}{2}$. Se cuenta, además
 $a =$ número total de concordancias,
 $b =$ número total de discordancias,
3. Se define el coeficiente de correlación de rangos como:

$$\tau = \frac{a - b}{\frac{n(n-1)}{2}}$$

Coefficiente de correlación de rangos de Spearman

Se consideran, igual que antes, n objetos clasificados según dos variables o criterios.

Por ejemplo, supongamos dos variables x e y que toman n valores emparejados (x_1, y_1) , $(x_2, y_2), \dots, (x_n, y_n)$. Se definen los rangos sobre cada una de las variables, de modo que se emparejan $(r_{x_1}, r_{y_1}), (r_{x_2}, r_{y_2}), \dots, (r_{x_n}, r_{y_n})$:

$$\begin{array}{cc|cc} x_1 & y_1 & r_{x_1} & r_{y_1} \\ x_2 & y_2 & r_{x_2} & r_{y_2} \\ \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & r_{x_n} & r_{y_n} \end{array}$$

Se definen las diferencias $d_i = (r_{x_i} - r_{y_i})$, es decir, las diferencias de la posición del individuo i -ésimo según la clasificación (rango) dada por x y la clasificación (rango) dada por y .

El coeficiente de correlación se define, entonces, como

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}.$$

Coeficientes de asociación (*matching types*)

Se consideran variables dicotómicas que toman como posibles valores 0 ó 1, del tipo *presencia - ausencia*. Existen diferentes formas de medir las coincidencias.

Ejemplo: Se tienen dos observaciones en las que se consideran 5 variables dicotómicas (*sí / no*).

Sea Sí = 1 y No = 0

individuos \ variables	x_1	x_2	x_3	x_4	x_5
A	1	1	0	0	1
B	0	1	0	1	0

Un posible coeficiente de similitud sería: m/N donde m = número de variables comunes a los dos elementos y M es el número total de variables. En este ejemplo, sería $2/5$.

Antes de mostrar una serie de medidas habituales, se tienen que definir los siguientes términos para 2 individuos dados.

X_{Aj} = valor del individuo A en la variable j -ésima $\in \{1, 0\}$.

X_{Bj} = valor del individuo B en la variable j -ésima $\in \{1, 0\}$.

$$V = \sum_j X_{Aj} (1 - X_{Bj}) \quad \text{N}^\circ \text{ de atributos donde A es 1 y B es 0}$$

$$R = \sum_j X_{Aj} X_{Bj} \quad \text{N}^\circ \text{ de atributos donde A y B son 1}$$

$$S = \sum_j (1 - X_{Aj}) (1 - X_{Bj}) \quad \text{N}^\circ \text{ de atributos donde A y B son 0}$$

$$T = \sum_j (1 - X_{Aj}) X_{Bj} \quad \text{N}^\circ \text{ de atributos donde A es 0 y B es 1}$$

$$U = R + S + T + V \quad \text{N}^\circ \text{ total de atributos o variables}$$

En el ejemplo anterior,

$$V = 1(1 - 0) + 1(1 - 1) + 0(1 - 0) + 0(1 - 1) + 1(1 - 0) = 2$$

$$R = 1$$

$$S = 1$$

$$T = 1$$

$$U = 5$$

Esto da lugar a distintos índices de similaridad, por ejemplo,

Índice de Russel-Rao

$$C = \frac{R}{U}$$

En el ejemplo es $1/5$.

Índice de Kendall

$$C = 1 - \frac{V + T}{U}$$

En el ejemplo es $2/5$.

Indice de Jaccard

$$C = \frac{R}{R + T + V}$$

En el ejemplo es 1/4.

Indice de Dice-Sorensen

$$C = \frac{2R}{2R + T + V}$$

En el ejemplo es 2/5.

Los índices más habituales son los de Jaccard y Dice-Sorensen.

Cuando se consideran variables categóricas otra posible medida de distancia se construye considerando la tabla de asociación entre variables como una tabla de contingencia y calculando el valor de la chi-cuadrado, χ^2 , de modo que se puede definir la distancia como el *coeficiente de contingencia*:

$$d_{ij} = 1 - \sqrt{\frac{\chi^2}{n}}$$

Introducción al Multidimensional Scaling (*MDS*)

Supongamos que se dispone de una matriz, \mathbf{D} , $n \times n$ de distancias o disimilaridades entre los n elementos de un conjunto. Por ejemplo, los n elementos pueden ser distintos productos fabricados por una empresa, candidatos políticos o apartados de un cuestionario. Las distancias se han obtenido preguntando a un grupo de evaluadores que ha estimado su proximidad relativa. El objetivo del MDS es representar las distancias observadas mediante unas variables y_1, \dots, y_k , donde $k < n$ tales que las distancias euclídeas entre las coordenadas de los elementos respecto a estas variables sean iguales (o lo más próximas posibles) a las distancias o disimilaridades de la matriz partida. De esta mane-

ra, la representación gráfica en k dimensiones será una reproducción fiel de la estructura observada.

El procedimiento consiste en partir de la matriz cuadrada $n \times n$ de distancias o similitudes, \mathbf{D} , entre individuos de una población con elementos tales que $d_{ii} = 0$, $d_{ij} \geq 0$ y $d_{ij} = d_{ji}$. Se pretende obtener una matriz $n \times k$ que represente los valores de k variables en los n individuos y tal que la distancia euclídea entre los elementos, cuando se miden mediante estas variables, reproduzca aproximadamente la matriz \mathbf{D} inicial.

Cuando $k > 2$, las variables pueden ordenarse en importancia y suelen hacerse representaciones gráficas en dos y tres dimensiones para entender la estructura existente.

Este planteamiento presenta dos interrogantes: ¿Es siempre posible encontrar estas variables? ¿Cómo construirlas?

En general no es posible encontrar k variables que reproduzcan *exactamente* las distancias iniciales; sin embargo, es frecuente encontrar variables que reproduzcan aproximadamente las distancias iniciales.

Reconstrucción de las variables a partir de las distancias entre puntos

Dado un conjunto k de variables en n individuos representados en la matrix \mathbf{X} , $n \times k$, podemos construir dos tipos de matrices cuadradas y semidefinidas positivas: la matriz de covarianzas \mathbf{S} (definida por $\frac{1}{n}\mathbf{X}'\mathbf{X}$), si las variables tienen media cero, y la matrix de productos cruzados $\mathbf{Q} = \mathbf{X}\mathbf{X}'$.

El método de Componentes Principales consiste en un análisis de la matrix \mathbf{S} . El MDS puede verse como un análisis de la matrix \mathbf{Q} .

La matrix \mathbf{Q} puede interpretarse como una matrix de similitudes entre las observaciones ya que sus términos, q_{ij} , contienen el producto escalar de las observaciones de dos elementos dados

$$q_{ij} = \sum_{p=1}^k x_{ip}x_{jp} = \mathbf{x}'_i\mathbf{x}_j. \quad (1)$$

En efecto, como $q_{ij} = |\mathbf{x}_i| |\mathbf{x}_j| \cos \theta_{ij}$, si los dos elementos tienen coordenadas similares, $\cos \theta_{ij} \simeq 1$ y q_{ij} será grande. Por el contrario, si los dos elementos tienen valores muy distintos $\cos \theta_{ij} \simeq 0$ y q_{ij} será pequeño.

Cuando conocemos los valores de las variables en dos elementos podemos calcular fácilmente su distancia euclídea. Sea \mathbf{X} la matriz $n \times k$ de variables y sea $\mathbf{Q} = \mathbf{X}\mathbf{X}'$ la matriz de $n \times n$ de productos cruzados entre estas variables que será singular de rango k . La distancia euclídea al cuadrado entre dos elementos se define por:

$$d_{ij}^2 = \sum_{p=1}^k (x_{ip} - x_{jp})^2 = \sum_p x_{ip}^2 + \sum_p x_{jp}^2 - 2 \sum_p x_{ip} x_{jp}$$

y se obtiene inmediatamente en función de los términos de la matriz Q , mediante la expresión

$$d_{ij}^2 = q_{ii} + q_{jj} - 2q_{ij}. \quad (2)$$

Por tanto, dada la matriz \mathbf{X} podemos construir la matriz \mathbf{Q} y a partir de esta matriz es fácil obtener la matriz de distancias al cuadrado con ayuda de las expresiones (1) y (2).

El problema que se aborda en MDS es el inverso: dada una matriz de distancias al cuadrado, \mathbf{D} , con elementos d_{ij}^2 se trata de reconstruir la matriz \mathbf{X} .

Lo primero que se plantea es obtener la matriz \mathbf{Q} dada la matriz \mathbf{D} . Para ello, observemos, que sin pérdida de generalidad, siempre podemos suponer que las variables \mathbf{X} tienen media cero. En efecto, las distancias entre los puntos, d_{ij}^2 no varían si expresamos las variables en desviaciones a la media, ya que

$$d_{ij}^2 = \sum_{p=1}^k (x_{ip} - x_{jp})^2 = \sum_{p=1}^k [(x_{ip} - \bar{x}_p) - (x_{jp} - \bar{x}_p)]^2 \quad (3)$$

y, por tanto, podemos suponer siempre que las variables que buscamos tienen media cero. Por ello, como resulta que $\mathbf{X}'\mathbf{1} = 0$ se debe verificar que $\mathbf{Q}\mathbf{1} = 0$, es decir, la suma de todos los elementos de una fila de la matriz \mathbf{Q} (y de una columna ya que la matriz es simétrica) debe de ser cero. Para imponer estas restricciones, sumamos en (2) por filas:

$$\sum_{i=1}^n d_{ij}^2 = \sum_{i=1}^n q_{ii} + nq_{jj} = t + nq_{jj} \quad (4)$$

donde $t = \sum_{i=1}^n q_{ii} = \text{traza}(\mathbf{Q})$, y sabiendo que $\sum_{i=1}^n q_{ij} = 0$. Sumando (2) por columnas se tiene

$$\sum_{j=1}^n d_{ij}^2 = nq_{ii} + t \quad (5)$$

y sumando ahora (4) por filas de nuevo se tiene

$$\sum_{i=1}^n \sum_{j=1}^n d_{ij}^2 = 2nt.$$

Sustituyendo (3) y (4) en (2), tenemos que

$$d_{ij}^2 = \frac{1}{n} \sum_{i=1}^n d_{ij}^2 - \frac{t}{n} + \frac{1}{n} \sum_{j=1}^n d_{ij}^2 - \frac{t}{n} - 2q_{ij},$$

y llamando $d_{i\cdot}^2$ y $d_{\cdot j}^2$ a las medias por filas y por columnas y utilizando (5), tenemos que

$$d_{ij}^2 = d_{i\cdot}^2 + d_{\cdot j}^2 - d_{\cdot\cdot}^2 - 2q_{ij}. \quad (6)$$

donde $d_{\cdot\cdot}^2$ es la media de todos los elementos de \mathbf{D} , dada por

$$d_{\cdot\cdot}^2 = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d_{ij}^2$$

Finalmente, despejando de (6) resulta que

$$q_{ij} = -\frac{1}{2}(d_{ij}^2 - d_{i\cdot}^2 - d_{\cdot j}^2 + d_{\cdot\cdot}^2)$$

expresión que indica cómo construir la matriz \mathbf{Q} a partir de la matriz \mathbf{D} de distancias con elementos d_{ij}^2 .

De una manera equivalente, se puede obtener la matriz \mathbf{X} a partir de la matrix \mathbf{Q} . Si ésta es definida positiva de rango k entonces puede representarse por

$$\mathbf{Q} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}'$$

donde \mathbf{V} es una matriz $n \times k$ y $\mathbf{\Lambda}$ es una matriz diagonal $k \times k$ y contiene los vectores propios no nulos de \mathbf{Q} . Entonces, podemos escribir

$$\mathbf{Q} = (\mathbf{V}\mathbf{\Lambda}^{1/2})(\mathbf{\Lambda}^{1/2}\mathbf{V}')$$

y tomando

$$\mathbf{Y} = \mathbf{V}\mathbf{\Lambda}^{1/2}$$

se obtienen k variables incorreladas que reproducen la métrica. Observemos que las variables \mathbf{Y} no son las variables originales \mathbf{X} , sino sus componentes principales. En general no es posible encontrar las variables \mathbf{X} , pero sí podemos encontrar un conjunto de variables de la misma dimensión que sean combinación lineal de ellas. La razón es que dada la matriz \mathbf{Q} como

$$\mathbf{Q} = \mathbf{X}\mathbf{X}' = \mathbf{X}\mathbf{A}\mathbf{A}'\mathbf{X}$$

para cualquier matriz \mathbf{A} ortogonal, la matriz \mathbf{Q} sólo contiene información sobre el espacio generado por las variables \mathbf{X} y cualquier rotación de las variables \mathbf{X} deja la matriz \mathbf{Q} invariable.

Construcción de coordenadas principales

En general, la matriz de disimilaridades o distancias no será compatible con una métrica euclídea, pero es frecuente que la matriz \mathbf{Q} tenga k valores propios positivos y más grandes que el resto. Si los restantes $n - k$ valores propios no nulos son mucho menores que los demás, podemos obtener una representación aproximada de los puntos utilizando los k vectores propios asociados a valores propios positivos. En este caso, las representaciones gráficas conservarán aproximadamente la distancia entre los puntos.

El procedimiento para obtener las coordenadas principales es el siguiente:

1. Formar la matriz de distancias al cuadrado, \mathbf{D} , cuyos elementos son los cuadrados de las distancias.
2. Construir la matriz \mathbf{Q} de productos cruzados.
3. Obtener los valores y vectores propios de \mathbf{Q} . Tomar las k mayores si podemos suponer que los restantes $n \times k$ son próximos a cero.

4. Obtener las coordenadas de los puntos en las variables mediante $\sqrt{\lambda_i \mathbf{v}_i}$, donde λ_i es el valor propio y \mathbf{v}_i el vector propio.

El método puede también aplicarse cuando la matriz de partida \mathbf{Q} es una matriz de similitudes cualesquiera. Entonces $q_{ii} = 1, q_{ij} = q_{ji}$ y $0 \leq q_{ij} \leq 1$. La matriz de distancias asociadas será, por (2),

$$d_{ij}^2 = q_{ii} + q_{jj} - 2q_{ij} = 2(1 - q_{ij})$$

y puede comprobarse que $\sqrt{2(1 - q_{ij})}$ es una distancia y verifica la desigualdad triangular al corresponder a la distancia euclídea para cierta configuración de puntos.

Se ha propuesto como medidas de la precisión en la aproximación mediante los k valores propios positivos los coeficientes:

$$m_{1,K} = 100 \times \frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^p |\lambda_i|}$$

ó

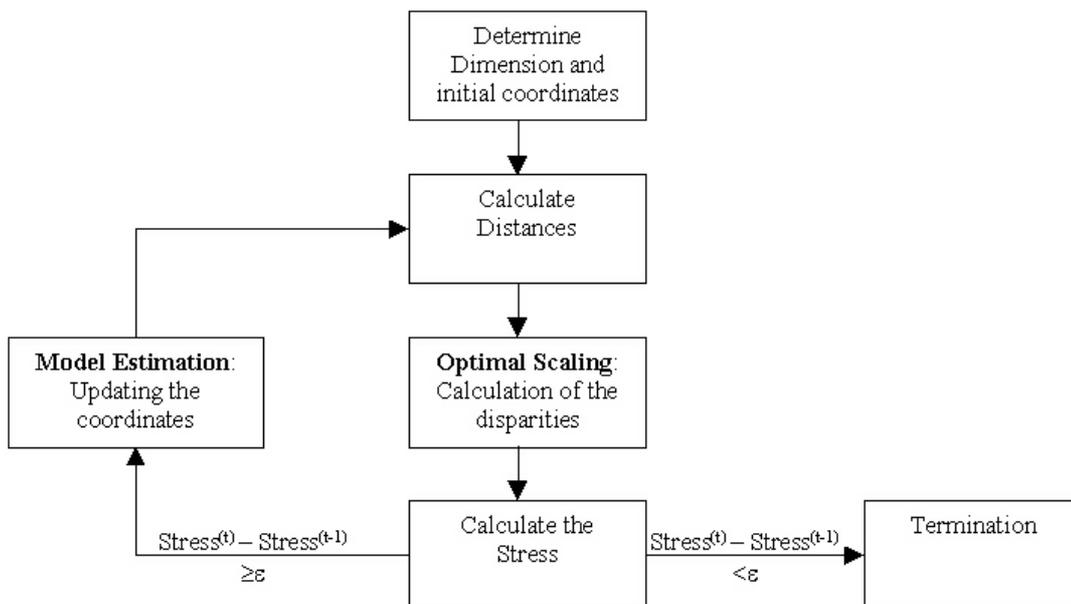
$$m_{2,K} = 100 \times \frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^p \lambda_i^2}$$

MDS no métrico

En el caso de datos ordinales se necesita disponer de otro procedimiento para considerar transformaciones lineales de las proximidades que sean invariantes bajo transformaciones monótonas. Una primera solución fue establecida por Shepard (1962) y fue posteriormente mejorada por Kruskal (1964). Ellos propusieron minimizar una función de adecuación denominada *stress* mediante un algoritmo iterativo. La función de stress es

$$S = \left[\frac{\sum_{i=1}^n \sum_{j>i}^n (\delta_{ij} - d_{ij})^2}{\sum_{i=1}^n \sum_{j>i}^n d_{ij}^2} \right]^{\frac{1}{2}},$$

y tiene en cuenta que no es realista esperar un ajuste perfecto del modelo a los datos. Por tanto, los valores δ_{ij} (denominados disparidades) se introducen en la función S como las aproximaciones óptimas de las proximidades observadas p_{ij} a las distancias d_{ij} en la representación geométrica posterior. Se utiliza, posteriormente, un algoritmo iterativo para minimizar la función de stress S :



El esquema anterior indica un procedimiento general que siguen diferentes variantes de MDS no métrico. Entre las más populares se encuentra el procedimiento ALSCAL, implementado en *SPSS* y en *SAS*, y el procedimiento PROXSCAL.

Ejemplo

Se tiene en la siguiente tabla las distancias en km por carretera entre las ciudades españolas siguientes: Madrid, Barcelona, Valencia, Sevilla, San Sebastián y La Coruña, que llamaremos matriz m .

	M	B	V	S	SS	LC
M	0	627	351	550	488	603
B	627	0	361	1043	565	1113
V	351	361	0	567	564	954
S	550	1043	567	0	971	950
SS	488	565	564	971	0	713
LC	603	1113	954	950	713	0

Puede comprobarse que la matriz Q viene dada por

$$Q = 1,0e + 05 \times$$

0.1176	-0.3908	-0.1795	0.3856	-0.3180	0.3852
-0.3908	3.0321	1.2421	-2.0839	0.7338	-2.5333
-0.1795	1.2421	0.7553	0.6095	-0.3989	-2.0285
0.3856	-2.0839	0.6095	3.6786	-2.0610	-0.5288
-0.3180	0.7338	-0.3989	-2.0610	1.6277	0.4165
0.3852	-2.5333	-2.0285	-0.5288	0.4165	4.2889

Los vectores propios de esta matriz son:

0.0960	-0.0443	-0.2569	0.1496	0.8566	0.4082
-0.6270	0.1400	-0.4155	-0.4717	-0.1593	0.4082
-0.2832	-0.2584	-0.0094	0.7670	-0.3130	0.4082
0.2934	-0.7216	0.2205	-0.4017	-0.1285	0.4082
-0.1241	0.4417	0.7812	-0.0687	0.0885	0.4082
0.6449	0.4426	-0.3198	0.0255	-0.3443	0.4082

y los valores propios

$$b = 1,0e + 05 \times$$

7.3792	5.9106	0.5947	-0.3945	0.0104	0.0000
--------	--------	--------	---------	--------	--------

La matriz Q tiene pues dos valores propios grandes y los otros cuatro son muy pequeños. Esto sugiere que las distancias pueden explicarse aproximadamente mediante dos variables. Tomando los dos vectores propios asociados a los mayores valores propios estandarizados por la raíz de su valor propio como coordenadas principales, resultan las siguientes coordenadas para cada ciudad

Madrid	82.44	-34.05
Barcelona	-538.61	107.67
Valencia	-243.29	-198.62
Sevilla	252.04	-554.79
San Sebastian	-106.60	339.55
La Coruña	554.02	340.25

Si se representan estas coordenadas en un gráfico las ciudades reproducen, aproximadamente, el mapa de España.

Ejemplo

Se consideran las distancias en relación a vuelos entre 10 ciudades norteamericanas:

	Atlanta	Chicago	Denver	Houston	L. Angeles	Miami	N York	S Francisco	Seattle	Washington
Atlanta	0.00	587.00	1212.00	701.00	1936.00	604.00	748.00	2139.00	218.00	543.00
Chicago	587.00	0.00	920.00	940.00	1745.00	1188.00	713.00	1858.00	1737.00	597.00
Denver	1212.00	920.00	0.00	879.00	831.00	1726.00	1631.00	949.00	1021.00	1494.00
Houston	701.00	940.00	879.00	0.00	1374.00	968.00	1420.00	1645.00	1891.00	1220.00
L Angeles	1936.00	1745.00	831.00	1374.00	0.00	2339.00	2451.00	347.00	959.00	2300.00
Miami	604.00	1188.00	1726.00	968.00	2339.00	0.00	1092.00	2594.00	2734.00	923.00
N York	748.00	713.00	1631.00	1420.00	2451.00	1092.00	0.00	2571.00	2408.00	205.00
S Francisco	2139.00	1858.00	949.00	1645.00	347.00	2594.00	2571.00	0.00	678.00	2442.00
Seattle	218.00	1737.00	1021.00	1891.00	959.00	2734.00	2408.00	678.00	0.00	2329.00
Washington	543.00	597.00	1494.00	1220.00	2300.00	923.00	205.00	2442.00	2329.00	0.00

```

aire.dist <- read.table("c:\\...\\aireMDS.txt")
dimnames(aire.dist)[[1]] <- c("Atlanta","Chicago","Denver","Houston",
"Los Angeles","Miami","New York","San Francisco","Seattle","Washington")

dimnames(aire.dist)[[2]] <- dimnames(aire.dist)[[1]]

# Se efectua un analisis clasico MDS metrico
aire.mds <- cmdscale(as.matrix(aire.dist),k=9,eig=T)

# Se Calculan los autovalores
aire.mds$eig

# Se normalizan los dos primeros autovalores
sum(abs(aire.mds$eig[1:2]))/sum(abs(aire.mds$eig))
sum(aire.mds$eig[1:2]^2)/sum(aire.mds$eig^2)

```

```
# La solución con dos dimensiones es adecuada
# Se muestran las coordenadas de las ciudades en las dos dimensiones
aire.mds$points[,1:2]

# Se dibujan las coordenadas de las ciudades en las dos dimensiones
par(pty="s")
plot(-aire.mds$points[,1],aire.mds$points[,2],type="n",xlab="Coordenada 1",
ylab="Coordenada 2",xlim = c(-2000,1500),ylim=c(-2000,1500))
text(-aire.mds$points[,1],aire.mds$points[,2],labels=row.names(aire.dist))
```

En R el MDS no métrico se puede implementar también, pero hay que cargar previamente la librería MASS.

```
library(MASS)
data(swiss)

# El fichero original de datos viene recogido ya en MASS:
swiss

swiss.x <- as.matrix(swiss[, -1])
swiss.dist <- dist(swiss.x)
swiss.mds <- isoMDS(swiss.dist)
# Las coordenadas son
swiss.mds$points

# Se dibujan los puntos
plot(swiss.mds$points, type="n", xlab="Coordenada 1", ylab="Coordenada 2")
text(swiss.mds$points, labels=as.character(1:nrow(swiss.x)))
```

Notas

Páginas interesantes sobre MDS y programas freeware asociados:

<http://www.ncl.ac.uk/mds/>

<http://www.granular.com/MDS/>

Programas en R del libro *Multivariate Statistics: Solutions and Exercises* de Hardle y Hlavka:

http://www.karlin.mff.cuni.cz/~hlavka/sms/MVAexercise_1.5.zip