Computers Math. Applic. Vol. 17, No. 1-3, pp. 137-145, 1989
Printed in Great Britain.

0097-4943/89 $3.00 + 0.00

ARRANGEMENTS OF MINIMAL VARIANCE— MULTIDIMENSIONAL SCALING IN THE SYMMETRICAL CASE

L. TELEGDI
Computer and Automation Institute, Hungarian Academy of Sciences, Budapest, P.O. Box 63, H-1502, Hungary

Abstract—We deal with the problem of how to arrange n points in the plane with a given mean and minimal variance. Ordinary and multiple multidimensional scaling are outlined, investigated in the symmetrical case and applied to congenital abnormalities.

1. ARRANGEMENT OF POINTS IN THE PLANE WITH DISTANCE OF MINIMAL VARIANCE

Certain extremum requirements imply regularity. This was illustrated by some examples in Ref. [1]: the area of an n-gon of given perimeter is maximized by the regular n-gon; among all convex polyhedra containing a ball, the circumscribed cube has the least total edge-length etc. In what follows we deal with the problem how to arrange n points in the plane with a given mean and minimal variance. (The solution conjectured and given below is not regular for $n > 6$, but it is symmetrical and consists of regular parts.) This problem is equivalent with the following: determine the points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ in the plane so as to minimize the sum of squares

$$V = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (K - \|\mathbf{x}_i - \mathbf{x}_j\|)^2$$

of the deviations of the distances between pairs of them from a given constant K. Moreover, we would like to determine the value $\tilde{V}$ of V for this optimal configuration. The latter formulation of the problem suggests its connection with mechanics: V can be regarded as the potential of C_n^2 springs with length K and spring constant 2 between each pair $\mathbf{x}_i$, $\mathbf{x}_j$ of points.

As V depends on the $\mathbf{x}_i$s only through their distances, the configuration obtained for the solution of the problem is obviously indeterminate with respect to translation, rotation and reflection. Almost the same can be true for uniform expansion or contraction: they result in similar configurations, and obviously $\tilde{V}(K) = K^2\tilde{V}(1)$. Therefore, we assume that $K = 1$.

The general solution of the problem is not known. As V is the function of the $\mathbf{x}_i$s, i.e.—denoting by x_{ik} the kth coordinate of the ith point—that of the unknown variables $x_{11}, x_{12}, x_{21}, x_{22}, \ldots, x_{n1}, x_{n2}$, a solution can be obtained by the minimization of a $2n$-variate function without constraint. Such minimizing procedures are usually gradient- or Fletcher-type algorithms, which are not able to distinguish between a local and a global minimum, therefore they need a good initial configuration.

Let $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n$ be the vertices of the $(n-1)$-dimensional regular unit n-hedron, then obviously V is 0, i.e. minimal for these $\mathbf{z}_i$s. Let

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \ldots & 1 \end{bmatrix}$$

be the identity,

$$\mathbf{J} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

the unit matrix, and

$$\mathbf{B} = \frac{1}{2}\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right).$$

Let $\lambda_1 \geqslant \lambda_2 > 0$ be the two largest eigenvalues of **B**.

Definition

Let $\mathbf{v}_{(i)}$ be the ith eigenvector of **B**,

$$\mathbf{B}\mathbf{v}_{(i)} = \lambda_i \mathbf{v}_{(i)},$$

normalized according to $\mathbf{v}_{(i)}^{\mathrm{T}}\mathbf{v}_{(i)} = \lambda_i$, $i = 1, 2$. The rows of the matrix $[\mathbf{v}_{(1)}, \mathbf{v}_{(2)}]$ are called the *principal coordinates* of the configuration $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n)$ in two dimensions.

Remark

$\lambda_1 = \lambda_2 = \frac{1}{2}$; $\mathbf{v}_{(1)}$ and $\mathbf{v}_{(2)}$ are two arbitrary orthogonal vectors of the $(n-1)$-dimensional subspace orthogonal to the vector $(1, 1, \dots, 1)^{\mathrm{T}}$.

Theorem

Amongst all projections of **Z** onto planes, the quantity

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[1 - \|\mathbf{x}_i - \mathbf{x}_j\|^2\right]$$

(which is positive, because projecting a configuration reduces the interpoint distances) is minimized when **Z** is projected onto its principal coordinates in two dimensions [2].

This theorem suggests a possible choice of initial configuration: choose the configuration in the plane whose coordinates are determined by the first two eigenvectors of **B**. V can be minimized iteratively, starting with this configuration. The iteration can be the repeated application of the following procedure: the negative gradient vector of the $2n$-variate function $V = V(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ is determined (explicitly calculated), and by performing a line search in the direction of this vector the new coordinates of the points $\mathbf{x}_i$ are calculated ($i = 1, 2, \dots, n$). At computer realization it is worthwhile taking the gradients of the functions

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} \left(1 - \|\mathbf{x}_i - \mathbf{x}_j\|\right)^2, \quad i = 1, 2, \dots, n,$$

instead of the gradient of V (moving only one point at the same time instead of n points. By this both computer time and memory demand decrease).

Even in case of the above initialization it can occur that only a local minimum of V is produced. The danger of its occurrence can be decreased in such a way that sum of qth powers is written instead of sum of squares in V, and the value of q is changed in the course of the algorithm. According to computational experiences, the algorithm is the most effective if after the initialization q is chosen for 3, then as soon as V already decreases only to a small extent (the points $\mathbf{x}_i$ change scarcely), its value is taken first for 2, then for 1, and finally again for 2.

Proceeding in the above fashion, the algorithm gives arrangements in which the points are situated—for $n = 3, 4, \dots, 65$—on vertices of concentric regular polygons (regular polygons inscribed into concentric circles). For example, for $n = 9$, 16 and 23 we have, respectively, (i) a regular octagon and its centre; (ii) a regular 12-gon and within this a concentric square; (iii) a

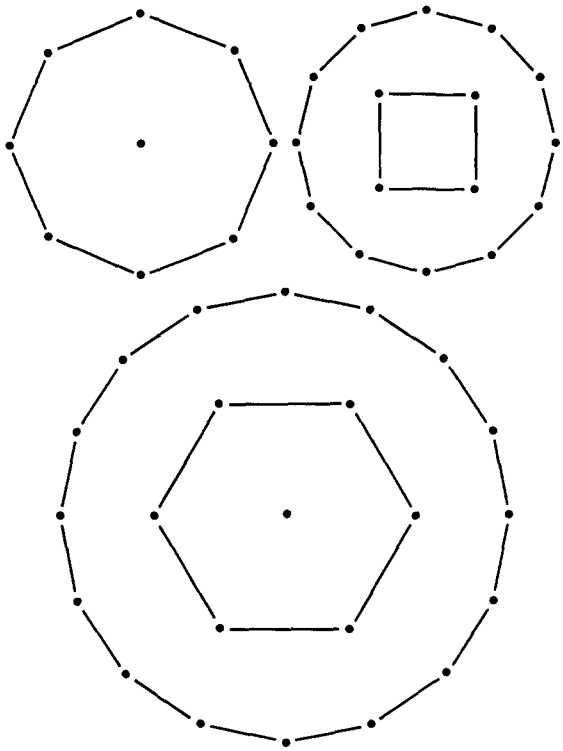

Fig. 1. Arrangements of minimal variance for $n = 9$, 16 and 23.

regular 16-gon, within this a concentric regular hexagon, and their common centre (Fig. 1). The numbers m_ν of vertices of the various polygons for some other ns are exhibited in Table 1 (from outside inwards). For the final value $\tilde{V}$ of the term V the equation

$$\tilde{V} = B_n C_{n-2}^2, \tag{1}$$

and here for B_n under $4 \leqslant n \leqslant 65$ the inequality

$$0.1716 \leqslant B_n \leqslant 0.1807,$$

was obtained

$$(0.1716 = 3 - 2\sqrt{2} = B_4;\ 0.1807 = \frac{75 - 36\sqrt{3}}{70} = B_7;$$

as the value of n approaches 65, the value of B_n approaches 0.176).

Table 1. Numbers of vertices of polygons

n	m_1	m_2	m_3	m_4
3	3	—	—	—
4	4	—	—	—
5	5	—	—	—
6	6	—	—	—
7	6	1	—	—
10	9	1	—	—
15	12	3	—	—
20	15	5	—	—
25	17	7	1	—
30	19	9	2	—
35	21	10	4	—
40	24	11	5	—
45	25	12	7	1
50	28	13	8	1
55	29	15	8	3
60	30	16	10	4
65	32	17	12	4

2. MULTIDIMENSIONAL SCALING

Let us assume that we are investigating M objects and n arbitrary variables characterizing them. Most of multivariate statistical methods work with data of the variables observed on the objects. In contrast with these methods, in case of *multidimensional scaling* (MDS) we can't—or don't want to—observe directly the data points as M points in n-dimensional Euclidean space, but we have only indirect information about them. This information may concern the *distance* (*dissimilarity*) or—on the contrary—*proximity* (*similarity*) of the objects or/and variables. MDS deals with the following problem: how can the objects or/and variables be drawn in the space on the basis of an $(M \times M)$, $(n \times n)$ or $(M \times n)$ distance of similarity matrix (data matrix of MDS), in other words how can an M-tuple or/and an n-tuple of points be constructed in low-dimensional Euclidean space with the property that Euclidean distance of the points should reflect distance (dissimilarity) of the objects or/and variables as well as possible?

Tasks involving MDS can be classified according to a few organizing concepts. A major one is whether the data of MDS (i.e. the distances or similarities) represent one or two sets of things. If they represent one set of things (either the objects, or the variables), they are called *one-mode*, if two sets of things (the objects *and* the variables), they are called *two-mode*.

In case of MDS with one-mode data the set of things represented by the rows of the data matrix of MDS is the same as the set of things represented by the columns. In such a case the data matrix is square and symmetric. Depending on whether the rows and columns represent the objects or the variables, we speak of MDS of the objects or that of the variables.

In case of MDS with two-mode data the set of things represented by the rows differs from the set of things represented by the columns. In such a case the data matrix is rectangular. MDS with two-mode data is called also *multidimensional unfolding* (MDU). In what follows we deal with the MDS of the variables (i.e. MDS with one-mode data, representing the variables). For more details on MDS [3, 4], Chap. 14 of Ref. [2], Chap. 5 of Ref. [5], furthermore Refs [6–8] are recommended to the reader.

Let $\mathbf{D} = [d_{ij}]$ be an arbitrary *distance matrix* (it means the following: $\mathbf{D}$ is symmetric, and

$$d_{ii} = 0, \quad d_{ij} \geqslant 0, \quad i \neq j).$$

k-dimensional points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are to be determined in such a way that denoting by $\hat{d}_{ij}$ the Euclidean distance of $\mathbf{x}_i$ and $\mathbf{x}_j$, the matrix $[\hat{d}_{ij}]$ should be "similar" to $\mathbf{D}$ in some sense. Usually not only the points $\mathbf{x}_i$, but also the dimension k is unknown. In practice, this latter one is mostly chosen for 1, 2 or 3, because then the variables are in fact drawn by the points. In many cases there is a point configuration in some p-dimensional Euclidean space R^p the interpoint distance matrix of which is just $\mathbf{D}$ (i.e. $\mathbf{D}$ is *Euclidean*). This configuration can be accepted as solution to the MDS problem only if p can be chosen for k. However, in practice p is usually too large for this.

The deviation between an arbitrary distance matrix $[d_{ij}]$ and the Euclidean distance matrix $[\hat{d}_{ij}]$ of some point configuration $\mathbf{X}$ in the course of MDS is measured mostly by one of the following terms:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (d_{ij} - \hat{d}_{ij})^2, \tag{2}$$

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} c_{ij}(d_{ij} - \hat{d}_{ij})^2, \tag{3}$$

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (d_{ij} - \hat{d}_{ij})^2 \Big/ \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \hat{d}_{ij}^2$$

[the c_{ij}s are given weighting factors; expression (2) is the special case of expression (3) belonging to $c_{ij} \equiv 1$]. Each of the above terms is the function of $\mathbf{X}$, i.e. an $(n \times k)$-variate function. Therefore the MDS problem is solved similarly to the algorithm described in the previous section. The problem of the arrangement of minimal variance is equivalent to MDS under $k = 2$, $d_{ij} \equiv 1$ and deviation measure (2).

3. MULTIPLE MDS

Let $y_1, y_2, \ldots, y_M$ be arbitrary objects characterized by the *dichotomous* variables $W_1, W_2, \ldots, W_n$ indicating the presence or absence of the characters $A_1, A_2, \ldots, A_n$. Let us assume that our task is MDS of the variables. In accordance with the previous section it means the following: distances are constructed between the variables, and the variables are to be put in low-dimensional Euclidean space in such a way that Euclidean distances of the points corresponding to the variables should differ from distances of the variables to as small extent as possible. In order to be well scalable, the variables must be consistent in the following sense: if two variables are near to a third one, they must be near to one another too. Let us assume for instance (Example 1) that $n = 3$, $M = 44$, the first 19 objects have the characters A_1 and A_2, the following 14 ones A_1 and A_3, and the last 11 ones A_2 and A_3. To these objects the distance matrix

$$\begin{bmatrix} 0 & 3 & 4 \\ 3 & 0 & 5 \\ 4 & 5 & 0 \end{bmatrix}$$

of the variables can be assigned. On its basis the variables can be well scaled (see Fig. 2). Let us assume now (Example 2) that $M = 33$, the first 19 objects have the characters A_1 and A_2, and the other 14 ones A_1 and A_3. To these objects the distance matrix

$$\begin{bmatrix} 0 & 3 & 4 \\ 3 & 0 & 60 \\ 4 & 60 & 0 \end{bmatrix}$$

of the variables can be assigned. On its basis the variables can be scaled only badly. Let us try therefore to scale them on the basis of the first 19 and the other 14 objects separately! To these sets of objects (which can be regarded as clusters) the distance matrices

$$\begin{bmatrix} 0 & 3 & 60 \\ 3 & 0 & 60 \\ 60 & 60 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 60 & 4 \\ 60 & 0 & 60 \\ 4 & 60 & 0 \end{bmatrix}$$

of the variables can be assigned. On their basis the variables—on two planes!—can already be well scaled (see Fig. 3).

Multiple multidimensional scaling (MMDS), see Ref. [9], deals with cases similar to Example 2, with the problem arising if the consistency is not fulfilled: the objects are to be divided into disjoint clusters as homogeneously as possible, where the homogeneity of a cluster is measured by the goodness of (ordinary) MDS of the variables under it. We would like to determine the positive integer p, the disjoint clusters

$$Y_1, Y_2, \ldots, Y_p \subset \{y_1, y_2, \ldots, y_M\} = Y$$

of the objects with the property that

$$\bigcup_{m=1}^{p} Y_m = Y,$$

furthermore, the points $\mathbf{x}_i^{(m)}$ of R^k ($i = 1, 2, \ldots, n$; $m = 1, 2, \ldots, p$) which represent the variables in the sense that the closeness of the points $\mathbf{x}_i^{(m)}$ and $\mathbf{x}_j^{(m)}$ corresponds to the proximity of the variables W_i and W_j under Y_m.

Fig. 2. MDS in the consistent case (Example 1).

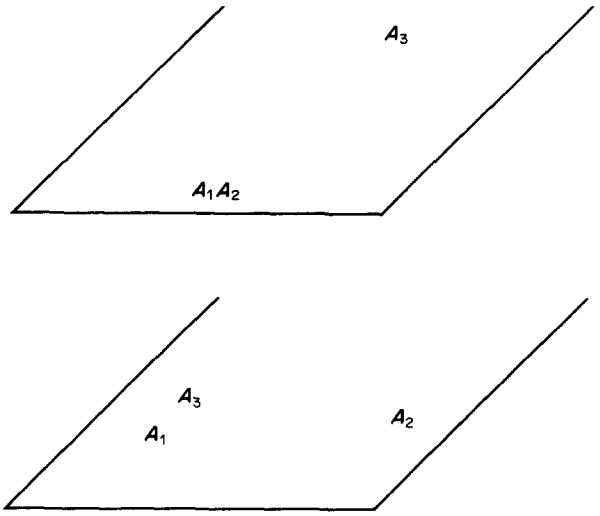

Fig. 3. MDS on two planes in the inconsistent case (Example 2).

Let e_{gi} be the value—1 or 0—of the variable W_i observed on the object y_g ($g = 1, 2, \ldots, M$; $i = 1, 2, \ldots, n$), and let

$$\mathbf{e}_g = [e_{g1}, e_{g2}, \ldots, e_{gn}]^T$$

($g = 1, 2, \ldots, M$). As objects with the same $\mathbf{e}_g$ are indistinguishable, let us assume that

$$\mathbf{e}_{g_1} = \mathbf{e}_{g_2} \Rightarrow g_1 = g_2,$$

and the objects have multiplicities. As the allocation of the various objects must obviously depend on the variables having value 1 on them, MMDS can't do anything with the objects on which at most one variable has value 1. Therefore, let us assume that on each object at least two variables have value 1. For arbitrary combination G_g of characters let the number of characters belonging to G_g be called the *side* of G_g, furthermore, let us denote by $O(G_g)$ the number of objects which have the characters belonging to G_g, but others not, and by N the number of objects different with respect to the characters and having at least two of them. Then each object y_g corresponds to a character combination G_g with side not less than 2, and has multiplicity $O(G_g)$ ($g = 1, 2, \ldots, N$).

Let

$$n_{ij}^{(m)} = \sum_{\substack{g: y_g \in Y_m \\ e_{gi}, e_{gj} = 1}} O(G_g) \tag{4}$$

(the number of those objects of the mth cluster in which A_i and A_j are present), and

$$E^{(m)} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} [n_{ij}^{(m)} \|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|^2 + (K - \|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|)^2],$$

then

$$E = \sum_{n=1}^{p} E^{(m)},$$

which is the MMDS function (K is an appropriate constant), is minimized by the alternate application of the following two procedures: (1) optimal classification of the objects among the actual clusters; (2) optimal allocation of the variables for the actual classification of the objects.

In the second procedure for $m = 1, 2, \ldots, p$ $E^{(m)}$ is (in the course of the various steps not minimized, but only) decreased similarly to the algorithm described in Section 1. Let the distances

$$d_{ij}^{(m)} = \frac{K}{n_{ij}^{(m)} + 1}$$

between the variables and the corresponding weights $c_{ij}^{(m)} = n_{ij}^{(m)} + 1$ be introduced, and let

$$V^{(m)} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} c_{ij}^{(m)} [d_{ij}^{(m)} - \|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|]^2,$$

then $[E^{(m)} - V^{(m)}]$ does not change in these, gradient steps (see Ref. [9]), thus MMDS is the generalization of MDS.

In the first procedure for $g = 1, 2, \ldots, N$

$$U_g^{(m)} = \sum_{\substack{i=1 \\ e_{gi}=1}}^{n-1} \sum_{\substack{j=i+1 \\ e_{gj}=1}}^{n} \|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|^2$$

is minimized in m. Ref. [9] proves that E is monotone decreasing in this step (in the gradient steps obviously).

Let us denote by T the algorithm which minimizes $U_g^{(m)}$ in m in the first procedure and decreases $E^{(m)}$ in the second procedure. In the course of T some clusters can become almost or entirely empty. It is reasonable to cease such clusters (by which p naturally decreases) in such a way that some object y_g having belonged to them is put into that cluster of the remaining ones under which $U_g^{(m)}$ is minimal. For—among others—this reason it is worthwhile choosing the initial value of p large. Initial clusters can be obtained, e.g. by applying the k-means (see Ref. [10]), more precisely in our case p-means method to the objects y_g on the basis of the vectors $\mathbf{e}_g$. Initial point configurations under the various initial clusters can be chosen similarly to Section 1.

It would be good to evaluate MMDS of some data field from as many as two points of view. The mathematical evaluation of the goodness of MMDS compared with other clustering methods is problematic theoretically. Namely the various clustering methods differ from one another decisively just in the criteria they give to the goodness of clustering. According to its own criterion each method is the best, however the criteria can't be compared objectively. (For more thorough investigations of the comparison of clusterings see, e.g. Refs [11, 12].) One can evaluate how well the objects of the concrete data field can be clustered with respect to MMDS theoretically in the following way. Let us assume that at the end of MMDS of the data field under some K the number of clusters is p and the value of E is

$$E^* = E^*(n, \mu, p, K),$$

where

$$\mu = \frac{\sum_{m=1}^{p} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} n_{ij}^{(m)}}{C_n^2} = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} O_T(i,j)}{C_n^2},$$

$O_T(i,j)$ denotes the number of objects which have the characters A_i, A_j, and possibly others too. Let us consider the respective data fields which belong to the pair (n, μ) of values, and are optimal and pessimal with respect to MMDS. Let us denote by E_{opt} and E_{pess} the values of the term E at the end of MMDS of these data fields under the above values of p and K. One can characterize how good the concrete data field is with respect to MMDS in a natural way by the term

$$\frac{E^* - E_{\text{opt}}}{E_{\text{pess}} - E_{\text{opt}}}, \tag{5}$$

which can't be less than 0 and greater than 1. Its determination would need the knowledge of the values E_{opt} and E_{pess}. However they are unknown, because the optimal and pessimal data fields are unknown too. In the next section we will consider a data field which is though not pessimal, but from the point of view of clustering bad. Let us denote by E_{bad} the value of E at the end of MMDS of this data field under the above p and K. In a fortunate case $(E_{\text{bad}} - E_{\text{pess}})$ and E_{opt} are small enough, thus expression (5) can be substituted for (E^*/E_{bad}).

4. MMDS IN THE SYMMETRICAL CASE

In the operation of MMDS the function E characterizing MMDS has a decisive role. It was specified in such a way that it depends only on the joint occurrences of the various pairs of characters. It implies that if the probabilities of the occurrences of the various characters are different, the procedure doesn't become aware of independence: it brings two characters close to one another even if they occur often together only because both of them are frequent. (However, this is on purpose: in many cases—e.g. in the statistical investigation of congenital abnormalities, out of which MMDS has grown—it is an important requirement that typical character combinations should not remain unobserved.) One can easily accomplish that the procedure should become aware of independence: in definition (4) of $n_{ij}^{(m)}$ the sum is to be divided by the term

$$\sum_{\substack{g: y_g \in Y_m \\ e_{gi}=1}} O(G_g) \times \sum_{\substack{g: y_g \in Y_m \\ e_{gj}=1}} O(G_g) \Big/ \sum_{g: y_g \in Y_m} O(G_g).$$

However, if the probabilities of the occurrences of the various characters are equal to each other, it is not necessary. Therefore, independence was investigated in this, symmetrical case. Particularly, a data field with the property

$$O_T(i,j) \equiv \mu$$

$(1 \leqslant i < j \leqslant n)$ was considered. Then taking the number p of the clusters of objects constant, MMDS gave clusters of objects under which no cluster of characters became distinct and the points corresponding to the characters were situated symmetrically, according to Section 1. For the value $\tilde{E}(n, \mu, p, K)$ of the term E at the end of MMDS of the data field under given values of n, μ, p and K, the equation

$$\tilde{E}(n, \mu, p, K) = \frac{pK^2}{\mu + p} (p\tilde{V} + C_n^2 \mu)$$

was obtained [for $\tilde{V}$ see expression (1)].

5. MMDS OF CONGENITAL ABNORMALITIES

From the conception until the birth structural defects may develop in the embryo and fetus. Such a defect is called *congenital abnormality* (CA). Within the CAs the *multiple congenital abnormalities* (MCAs), which are the concurrences of two or more different CAs in the same person, have a special importance. A major purpose of their statistical analysis, reported in details in Ref. [13], was to explain the possible cause(s) of combination of CAs within MCAs. The analysis was based on children born in Hungary 1970–1976 (Data field 1) and 1977–1982 (Data field 2). Under Data field 1 $n = 40$, $M = 1{,}186{,}776$ and $N = 881$, under Data field 2 $n = 45$, $M = 937{,}320$ and $N = 867$.

The statistical investigation of MCAs can be based on different alternative hypotheses. There are two models which are reasonable, general and effective. The *Gaussian threshold model* assumes that for any member of the population there is a measure of any CA which can be expressed in a real number. In other words, a background variable L_i, the so-called liability is assigned to A_i. According to the model, the joint distribution of the L_is are multidimensional Gaussian. The fact that somebody has A_i mean that his (or her) liability exceeds a threshold T_i characteristic for the population. The *mixture model* assumes that the probability distribution of MCAs is the mixture of distributions having the property that in each of them the CAs are independent.

One of the main aims of the statistical investigation of MCAs was the classification of children with CAs on the basis of the occurrences of the CAs, and by this the detection of characteristic CA combinations. MMDS was elaborated for solving this task. We can ascertain its adequacy in such a way that to clusters of CAs generated randomly or produced systematically we generate a random data field, and examine whether the classification of the generated children by MMDS has the property that under the various clusters of children the clusters of CAs become distinct. For generating the random children we must first, of course, specify the hypothesis and model

(describing MCAs) on the basis of which the generation will be performed. The problem of the inconsistency of the variables W_i requiring MMDS doesn't arise under the Gaussian threshold model, therefore we generated data fields on the basis of the mixture model. Their MMDS gave favourable results, ascertaining the adequacy of MMDS of the "real" data fields [13].

REFERENCES

1. L. Fejes Tóth, Symmetry induced by economy. *Comput. Math. Applic.* **12B,** 83–91 (1986). Reprinted in *Symmetry: Unifying Human Understanding* (Ed. I. Hargittai). Pergamon Press, Oxford (1986).
2. K. V. Mardia, J. T. Kent and J. M. Bibby, *Multivariate Analysis.* Academic Press, New York (1979).
3. J. B. Kruskal and M. Wish, *Multidimensional Scaling.* Sage Publications, Beverly Hills (1978).
4. S. S. Schiffman, M. L. Reynolds and F. W. Young, *Introduction to Multidimensional Scaling.* Academic Press, New York (1981).
5. A. D. Gordon, *Classification.* Chapman and Hall, London (1981).
6. J. De Leeuw and W. Heiser, Theory of multidimensional scaling, In *Handbook of Statistics, Vol. 2: Classification, Pattern Recognition and Reduction of Dimensionality* (Eds P. R. Krishnaiah and L. N. Kanal), pp. 285–316. North-Holland, Amsterdam (1982).
7. M. Wish and J. D. Carroll, Multidimensional scaling and its applications, In *Handbook of Statistics, Vol. 2: Classification, Pattern Recognition and Reduction of Dimensionality* (Eds P. R. Krishnaiah and L. N. Kanal), pp. 317–345. North-Holland, Amsterdam (1982).
8. J. Gower, Multivariate analysis: ordination, multidimensional scaling and allied topics, In *Handbook of Applicable Mathematics, Vol. VI: Statistics,* Part B (Ed. E. Lloyd), pp. 727–781. Wiley-Interscience, Chicester (1984).
9. L. Telegdi, Multiple multidimensional scaling: a new approach to the analysis of multidimensional contingency tables with application to congenital abnormalities. *Metron* **40,** 277–288 (1982).
10. J. A. Hartigan, *Clustering Algorithms.* Wiley, New York (1975).
11. L. J. Hubert and P. Arabie, Comparing partitions. *J. Classification* **2,** 193–218 (1985).
12. J. P. Barthelemy, B. Leclerc and B. Monjardet, On the use of ordered sets in problems of comparison and consensus of classifications. *J. Classification* **3,** 187–224 (1986).
13. A. Czeizel, L. Telegdi and G. Tusnády, *Multiple Congenital Abnormalities.* Publishing House of the Hungarian Academy of Sciences, Budapest (1988).