ROMAI J., v.15, no.2(2019), 61–75

# MULTIDIMENSIONAL SCALING AS A TOOL TO ANALYZE THE STRUCTURE IN DATASETS

Markela Muça[1], Angjela Dhame[2]

[1] *Department of Applied Mathematics, Faculty of Natural Sciences, University of Tirana, Albania*

[2] *Promoter for Samsung Campaign, Tirana, Albania*

markela.muca@fshn.edu.al, angjeladhame@yahoo.com

**Abstract** Statistical methods are increasingly relevant for uncovering structure in different datasets. This paper aims to explain one of these methods that is known as Multidimensional Scaling (MDS) by using simple and intuitive examples. Multidimensional Scaling accepts quantitative, qualitative and mix data. There is a wide variety of methods to obtain appropriate data for MDS. The most direct way is to ask subjects to give pairwise ratings or to sort stimulus objects according to their similarity, relatedness and association. There are three types of this method: classical MDS, metric MDS and non-metric MDS. The purpose of each is to provide a visual representation of the pattern of proximities among a set of objects in a low number of dimensions (Borg, I. & Groenen, F. J. F., 2005).

**Keywords:** distance, similarity, scaling, classical, metric, nonmetric,multidimensional scaling.

**2010 MSC:** 62H99.

## 1. INTRODUCTION

Multidimensional Scaling (MDS) is one of the most used statistical techniques nowadays to analyze data. Its purpose is to create a visual representation of distances or differences between objects in a low-dimensional space. These 'objects' can be colors, faces, geographical coordinates, political beliefs, or any kind of real or conceptual stimulus. Objects that are more similar (or have shorter distances) are closer to each other on the graph than objects that are less similar (or have longer distances). MDS, in addition to interpreting differences between objects as distances in a graphical representation, also serves as a dimension reduction technique for high-dimensional data (Kruskal, J. B. & Wish, M., 1978).

The term ***scaling*** comes from psychometry, where abstract concepts (objects) are assigned numbers according to a rule. For example, a person's attitude toward global warming may be required to be quantified. The number '1' is assigned for the 'doesnt believe in global warming' attitude, the

'10' for 'firmly believes in global warming' and a scale of 2 to 9 for attitudes in between. Since the data can be qualitative (ordinal or nominal), Shepard (Shepard, R. N., 1962) and Kruskal (Kruskal, J. B., 1964) proposed a non-metric multidimensional scaling model. By using scaling, data can be scaled down, that means that they become simpler by creating data of a lower dimension. Although data can be reduced in size, they retain similar properties. For example, two data points that are close to each other in the space of a high dimension will also be close to a space of a lower dimension.

The ***multi-dimensional*** concept is due to the fact that we are not limited to two-dimensional graphs or data. Three-dimensional, four-dimensional or higher plots are also available. Nowadays, multidimensional scaling is used in a wide variety of disciplines. Its use is not limited to a specific matrix or dataset. In fact, any matrix can be analyzed with these methods as long as the matrix contains some type of relational data. Examples of relational data include correlations, distances, multiple rating scales or similarities. Data can be quantitative, qualitative or mixed. The ALSCAL (Young, F. W., et al., 1978) / PROXSCAL (Busing, F. M. T. A., et al., 1997) iterative algorithms provided by SPSS software will be used for data processing. An initial configuration is needed to start the algorithm of PROXSCAL. The procedure offers some options on how we want the initial setup to be. The first is the Simplex configuration which assumes that all objects have the same distance from each other in the maximum space we have selected. The Torgerson option offers a classic multidimensional scaling solution as initial setup. And finally, a Single random start solution can be obtained as the initial configuration.

In this paper we will discuss some basic aspects of classical, metric and non-metric multidimensional scaling analysis in order to visualize the data and then interpret the graphical representation. To implement this method and get the best solution, there are four points to focus on:
The ***input matrix*** in MDS method must be a symmetric square matrix whose values denote the relationships between a set of objects. The method accepts two forms of proximity values: dissimilarity or similarity. The difference is somewhat misleading, because similarity is not the only relationship between items that can be measured and analyzed using MDS. However, dissimilarity is still used as a means of indicating if larger values in our data mean that a given pair of items should be placed next to each other on the map, or farther apart. In a matrix of dissimilarities, the larger values show less similarity which means they have to be placed far apart on the map. If we say that data is 'similarities', it indicates a negative or subtractive relationship between the input values and the corresponding map distances. The similarities are converted to dissimilarities using the formula (Mardia, K. V., et al., 1979):

$$d_{ij} = \sqrt{s_{ii} + s_{jj} - 2s_{ij}} \qquad (1)$$

where $d_{ij}$ represents the dissimilarity between $i$ -th and $j$-th objects and $s_{ij}$ represents the similarity between $i$ -th and $j$-th objects. The similarity of an object to itself is equal 1. When the data consists of standard measurements instead of dissimilarities or similarities, we can create a matrix of dissimilarities by first creating the correlation matrix and then, using the formula (1), convert the correlations to dissimilarities. Certain programs automatically calculate pair-wise correlations for the variables we specify. The procedure starts after we have an initial configuration of the input matrix. As an initial configuration it can be assumed that all objects are equidistant from one another in the space we have chosen or a solution is obtained by the classical (metric) method. A third option, but not very usable, is to obtain a random dot configuration.

***Goodness of fit***: as in any data analysis problem, an expression is needed to express how well a particular set of data are represented by the model that the analysis imposes. In the case of MDS, the main goal is to model the distances.Therefore, the most obvious choice for a goodness-of-fit statistic should be based on the differences between actual distances and their predicted values. Such a measure is called Stress (Kruskal, J. B., 1964). For a given dissimilarity matrix $D = (d_{ij})$, MDS minimizes Stress over configurations $(x_1, x_2, .., x_k)^T$, thought a $k$x$q$ dimensional hypervectors of unknown parameters. The minimization can be carry out by straight forward gradient descent applied to stress viewed as a function on $R^{kxq}$:

$$Stress = \sqrt{\frac{\sum_i \sum_j \left(f(x_{ij}) - \hat{d}_{ij}\right)^2}{\sum_i \sum_j \hat{d}_{ij}^2}} \qquad i = 1:k, j = 1:(i-1) \qquad (2)$$

Here $\hat{d_{ij}}$ is the Euclidean squared distance between $i$ -th and $j$-th objects, $f(x_{ij})$ is the fitted distance when the $\hat{d_{ij}}$ are monotonically regressed on the $d_{ij}$. This means that $f(x_{ij})$ is monotonic relative to $d_{ij}$ and is obtained from $\hat{d_{ij}}$ with the smallest number of changes. The transformation of input values $f(x_{ij})$ depends on metric or non-metric scaling. In metric scaling, $f(x_{ij}) = x_{ij}$ . In other words, raw input data is directly compared to map distances. In non-metric scaling, $f(x_{ij})$ is a weak monotonic transformation of the input data that minimizes the stress function. Monotonic transformation is calculated through 'monotonic regression', also known as 'isotonic regression'. From a mathematical standpoint, nonzero *Stress* values occur for only one reason: insufficient dimensionality. So it may be impossible to perfectly represent the input data in two or a small number of dimensions. On the other hand, any data set can be perfectly represented using the k-1 dimensions, where $k$ is the number of objects. As the number of dimensions increases, the *Stress* must either decrease or stay at the same value. Of course, a 'map' obtained from MDS can be used even when *Stress* is not zero. The value of Stress is always

positive. Different people set different standards in terms of the amount of error this measurement has to tolerate. According to Kruskal (Kruskal, J. B., 1964) the value of *Stress* varies from 0 to 0.2 and at this interval is required the smallest value of Stress which corresponds to a two / three dimensional space.

One of the main tasks the analyst has is to determine the ***number of dimensions*** in the MDS model. Each dimension represents a different underlying factor. One of the aims of the MDS analysis is to keep the number of dimensions as small as possible. Usually, the analyst will choose two or three dimensions, at most. If more is required, MDS may not be appropriate for the data available. The usual technique is to solve the MDS problem for a different number of dimensions and to use the smallest number of dimensions that reach a small *Stress* value.

The last and most important step of MDS analysis is ***the interpretation*** of the map obtained from it. There are two important things to understand about the MDS map. The first is that the axes, by themselves, are meaningless and secondly that the orientation in the picture is arbitrary. Therefore it is said that what matters in this presentation is which points are close to each other. When looking at an MDS map that has a non-zero Stress value (imperfect), one should keep in mind that the distances between objects are an imperfect representation of the relationships given by the input matrix. Groups and dimensions must be looked at in order to interpret the graphical representation (Borg, T. & Lingoes, J., 1962). Objects belonging to one group are closer to each other than objects belonging to another group. It is often important to check if there is any type of object clustering or any special connections and configurations (e.g. annular). Dimensions, which are the attributes of objects, list the objects on the map in a continuous order. These dimensions should explain the perceived similarity between objects. To make this explanation, it is seen whether there is any difference between the objects positioned opposite side of each axis of the respective dimension.

## 2. TYPES OF MDS

MDS models are classified according to the type of proximities and the number of the similarity matrices. Type of proximities can be metric/quantitative (e.g. distance between cities) or non-metric/qualitative (e.g. the order of sorting objects). In the following examples, the number of input matrices is one.

## 2.1. CLASSICAL METRIC MULTIDIMENSIONAL SCALING (CMDS)

The *Classical Metric Multidimensional Scaling (cMDS)*, which is otherwise known as Principal Coordinates Analysis (PCoA) or Torgerson-Gower Scale (Gower, J. C., 1966; Torgerson, W. S., 1952), is used for quantitative similarities. It uses the fact that the coordinate matrix Y (which shows the scaling solution-called the stimulus space) can be obtained by decomposing eigenvalues from $B = YY'$.

1 The centering matrix $J_k = I_k - \frac{1}{k}11^T$ is calculated, where $I_k$ is the unit matrix of the order $k$, $1 = (1, 1, .., 1)'_{kx1}$ and $1^T$ the transposed one.

2 Then, matrix $B$ can be calculated as $B = -\frac{1}{2}J_k D^{(2)} J_k$, where D is matrix of distance which is given or can be calculated by using Euclidean distance.

3 The eigenvalues and eigenvectors of the matrix $B$ are calculated and from them are obtained the $q$ largest eigenvalues and the corresponding eigenvectors of $B$ (where $q$ is the number of the desired dimension).

4 Finally, the matrix $Y$ is defined by the formula: $Y_{kxq} = E_{kxq}\Lambda^{1/2}_{qxq}$ , where $E_{kxq}$ is the modal matrix which has as columns $q$ of the autoregressors and $\Lambda_{qxq}$ is the spectral matrix which has as its main diagonal elements the $q$ eigenvalues of matrix $B$.

***Mapping example of Classical Multidimensional Scaling***

Lets consider the air distances between the 15 cities of Albania which are presented in *Table 1.* In this case, it would be inappropriate to work with land distances as this would result in a representation not similar to the actual geographical position of the cities. So, we have to be very careful with the type of data we use. Starting from the data in Table 1, where the values of this table constitute the elements of the input matrix $X = (x_{ij})$ with dimensions *15x15* or *15*(15-1)/2=105* similarities, it is required to build a coordinates matrix of Y with dimensions *15x2* in order to create a map which approximates the real map of Albania. The dimensional space will be required two-dimensional so that the interpretation will be easy. As noted, the input matrix is equal to the real distance matrix. These distances (similarities) compose the basic data and the number of stimulus and the number of iterations are given preliminarily equal with 2 and 1 respectively. Therefore, in this case it would not be necessary to use the Euclidean distance to calculate the distance matrix $D = (d_{ij})$ *(i,j=1:15).*

By following the steps of the cMDS procedure, we can calculate matrix B and it's eigenvalues and eigenvector. As we want the number of dimension to

| Cities | Tir | Lez | Elb | Fier | Vl | Shk | Trop | Gjir | Kor | Kuk | Dib | Kav | Ber | Sar | Pog |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Tir | 0 | 57 | 40 | 46 | 100 | 87 | 123 | 142 | 113 | 94 | 64 | 24 | 74 | 181 | 78 |
| Lez | 57 | 0 | 94 | 109 | 159 | 24 | 69 | 206 | 170 | 71 | 63 | 72 | 131 | 233 | 130 |
| Elb | 40 | 94 | 0 | 38 | 78 | 126 | 153 | 105 | 74 | 119 | 80 | 47 | 42 | 131 | 49 |
| Fier | 46 | 109 | 38 | 0 | 49 | 129 | 170 | 101 | 100 | 166 | 131 | 51 | 37 | 119 | 88 |
| Vl | 100 | 159 | 78 | 49 | 0 | 180 | 224 | 68 | 103 | 197 | 160 | 82 | 49 | 97 | 107 |
| Shk | 87 | 24 | 126 | 129 | 180 | 0 | 65 | 228 | 194 | 75 | 86 | 98 | 161 | 262 | 161 |
| Trop | 123 | 69 | 153 | 170 | 224 | 65 | 0 | 258 | 205 | 42 | 80 | 144 | 193 | 293 | 165 |
| Gjir | 142 | 206 | 105 | 101 | 68 | 228 | 258 | 0 | 81 | 223 | 182 | 132 | 66 | 35 | 101 |
| Kor | 113 | 170 | 74 | 100 | 103 | 194 | 205 | 81 | 0 | 165 | 125 | 120 | 66 | 105 | 34 |
| Kuk | 94 | 71 | 119 | 166 | 197 | 75 | 42 | 223 | 165 | 0 | 41 | 122 | 161 | 257 | 126 |
| Dib | 64 | 63 | 80 | 131 | 160 | 86 | 80 | 182 | 125 | 41 | 0 | 93 | 121 | 208 | 85 |
| Kav | 24 | 72 | 47 | 51 | 82 | 98 | 144 | 132 | 120 | 122 | 93 | 0 | 68 | 166 | 97 |
| Ber | 74 | 131 | 42 | 37 | 49 | 161 | 193 | 66 | 66 | 161 | 121 | 68 | 0 | 89 | 61 |
| Sar | 181 | 233 | 131 | 119 | 97 | 262 | 293 | 35 | 105 | 257 | 208 | 166 | 89 | 0 | 132 |
| Pog | 78 | 130 | 49 | 88 | 107 | 161 | 165 | 101 | 34 | 126 | 85 | 97 | 61 | 132 | 0 |

*Table 1 Distances between the 15 cities of Albania expressed in kilometers (km). List of abbreviations: Tir=Tirana, Lezha=Lez, Elb=Elbasan, Vl=Vlora, Shk=Shkodra, Trop=Tropoja, Gjir=Gjirokastra, Kor=Korca, Kuk=Kuksi, Dib=Dibra, Kav=Kavaja, Ber=Berat, Sar=Saranda, Pog=Pogradec.*

be two, we will select the two largest eigenvalues and it's corresponding eigenvecor. Modal matrix $E_{15x2}$ and the spectral one $\Lambda_{2x2}$ are shown as follows:

$$\Lambda_{2x2} = \begin{pmatrix} 103314.6 & 0.00 \\ 0.00 & 17803.13 \end{pmatrix}$$

$$E_{15x2} = \begin{pmatrix} -0.08 & -0.26 & 0.04 & 0.10 & 0.23 & -0.33 & -0.44 & 0.37 & 0.19 & -0.32 & -0.19 & -0.03 & 0.16 & 0.47 & 0.08 \\ 0.13 & 0.21 & -0.02 & 0.34 & 0.33 & 0.31 & -0.11 & -0.05 & -0.43 & -0.32 & -0.30 & 0.28 & 0.05 & -0.05 & -0.38 \end{pmatrix}^T$$

The coordinate matrix Y, the corresponding plot and the map of Albania are shown below:

$$Y_{15x2} = \begin{pmatrix} -24.14 & -82.34 & 13.37 & 30.89 & 73.96 & -106.35 & -140.20 & 117.64 & 61.15 & -103.83 & -60.39 & -8.33 & 52.52 & 151.48 & 24.58 \\ 17.04 & 28.34 & -2.19 & 45.05 & 44.20 & 41.22 & -14.54 & -6.23 & -56.90 & -42.96 & -40.06 & 37.74 & 6.57 & -6.56 & -50.73 \end{pmatrix}^T$$

To make this representation easier to perceive, an interpretation of its dimensions is made. It is noted that the points are scattered mainly along the horizontal axis which, given the real position of the cities on the map of Albania, might otherwise be interpreted as the "South-North" axis. The second dimension would be appropriate if interpreted as the "East-West" axis. Based on formula (2) we can calculate Stress value which in this case is equal to 0.994. Of course this is not a satisfactory value, but such an outcome was expected as this method consists of only one iteration.Therefore it is not very usable nowadays. Mostly it is used as an initial configuration of objects in

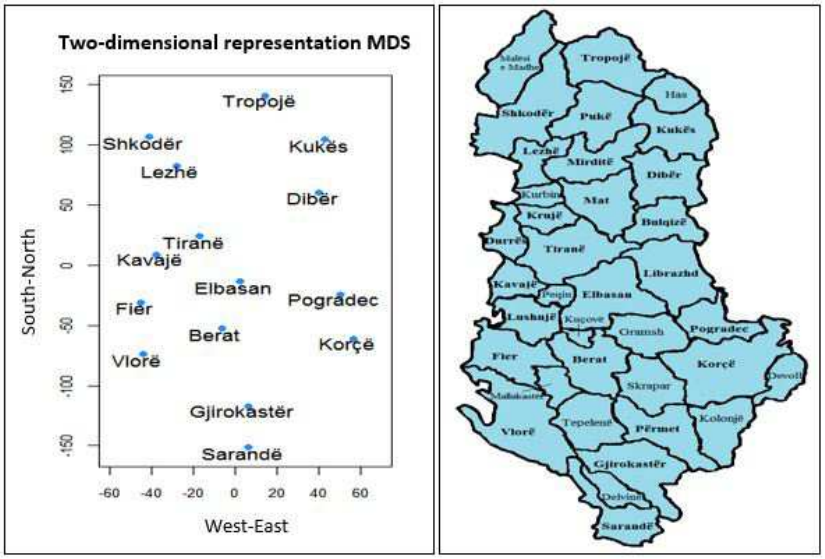


*Fig. 1. (a) Two-dimensional representation of cities and (b) Map of Albania.*

the other two types of multidimensional scaling, metric and non-metric, which will be discussed respectively in section 2.2 and 2.3.

## 2.2. METRIC MULTIDIMENSIONAL SCALING

Metric Multidimensional Scaling is used for quantitative similarities. He is an extension of Classic MDS that generalizes the optimization procedure for a variety of loss functions and initial matrix of known distances. A useful loss function in this context is Stress, which is often minimized using a procedure called majorization of stress value. This procedure consists of the steps of the cMDS algorithm, where after a distance matrix is obtained with the first iteration, the same steps are applied again on this matrix. The second configuration has a smaller value of the loss function. The procedure continues until an acceptable Stress value is obtained. In multidimensional metric scaling, it is assumed that distances can be transformed into Euclidean distances by means of several other transformations parametric.

### *Mapping example of Metric Multidimensional Scaling*

What should be achieved by the Metric MDS method is the creation of a 'map' that approximates the real map of Albania. The dimensional space will be required two-dimensional so that the interpretation will be easy. The PROXSCAL procedure was used in SPSS software to obtain this representa-

tion. In this example as an initial configuration is used the 'Simplex' configuration. The result is shown in *Figure 2*.

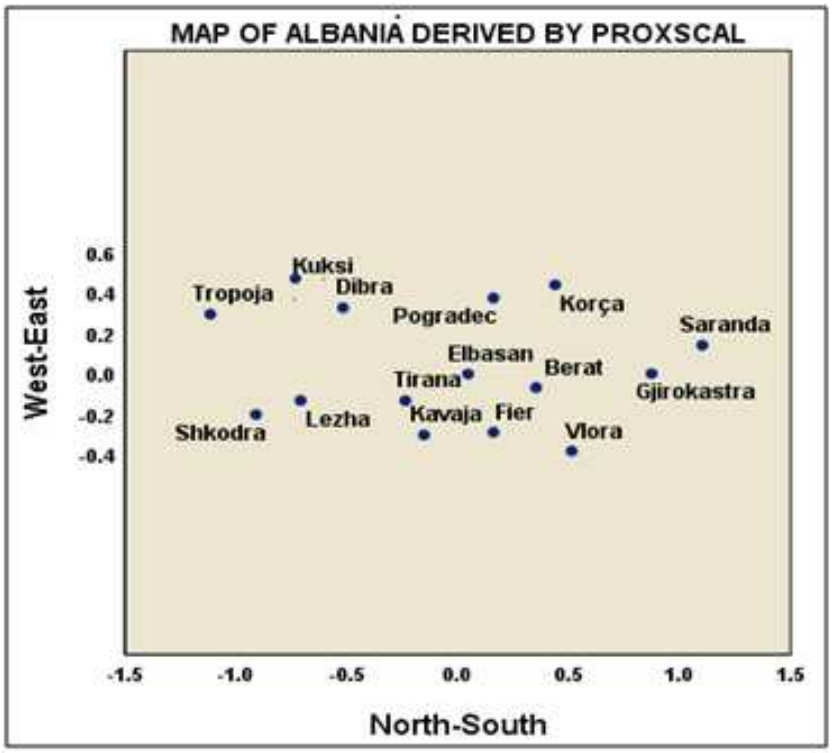


*Fig. 2. Two-dimensional representation of cities' distances.*

The quality of this solution can also be checked with the naked eye. What needs to be done is comparing the distance cities have in this layout with the data in *Table 1*. If distances and data are highly correlated, then distances represent data very well in a linear sense. This is obviously true in the given case, and it can therefore be said that the distances in *Figure 2* represent the distances in *Table 1* quite well.

To numerically confirm the goodness of this solution we interpret *Table 2* obtained from SPSS. This table shows that there are two types of indicators: those for which zero represent a perfect fit and those for which the perfect fit is represented by 1. Indicators of the first type are *Normalized Raw Stress, Stress-I, Stress-II* and *S-Stress*. It is noted that these values are close to zero. The *Dispersion Accounted For (D.A.F.)* and *Tucker's Coefficient of Congruence*, which are indicators of the second kind, are very close to one. What we care about most from this chart is the Normalized Raw Stress, which is obviously of little value. So, based on Kruskal's estimation that a Stress value

*Table 2 Stress and Fit Measures.*

| | |
|---|---|
| **Normalized Raw Stress** | .00116 |
| **Stress-I** | .03400[a] |
| **Stress-II** | .07505[a] |
| **S-Stress** | .00239[b] |
| **Dispersion Accounted For (D.A.F.)** | .99884 |
| **Tucker's Coefficient of Congruence** | .99942 |

close to 0 is nearly perfect, it can be said that the fit of the model in this case is quite good.

## 2.3. NON- METRIC MULTIDIMENSIONAL SCALING

Unlike metric multidimensional scaling, non-metric scaling uses qualitative disimilarites (similarities). Implicit from the above is the assumption that there is a real configuration in k dimensions, that $D = [d_{ij}]$ is a distance matrix. Often, it is more realistic to assume a less stringent relationship between observed distances and real distances. The idea of non-metric MDS is to require a less rigid relation between the disimilarities $d_{ij}$ and the distances $\delta_{ij}$. Suppose that f(x) is an unknown monotonically increasing function:

$$\delta_{ij} = f(d_{ij})$$

which is used to generate the known distances as a function of the given dissimilarities. Here $f$ has the property that

$$if \quad \delta_{ij} < \delta_{rs}, \;\; then \;\; f(\delta_{ij}) < f(\delta_{rs}).$$

The scaling is based on the rank order of the dissimilarities, so non-metric MDS is ordinal in character. The most common approach used to determine the elements $\delta_{ij}$ and to obtain coordinates of the objects $x_1, x_2 ... x_k$ given only rank order information is an iterative process called the Shepard-Kruskal algorithm (Shepard, R. N., 1962; Kruskal, J. B. 1964).

1 If all objects have different coordinates, Euclidean distances $\delta_{ij}^{(0)}$ are calculated from an arbitrarily chosen initial configuration $X^{(0)}$ in the dimension $q$. One might use metric MDS to obtain these initial coordinates.

2 The non-metric phase. Determines disparities (fitted distances) $\hat{\delta}_{ij}^{(0)}$ from the distances $\delta_{ij}^{(0)}$. They are obtained by constructing a monotonic regression relationship between the $\delta_{ij}^{(0)}$s and $d_{ij}$ s, under the requirement that:

$$if \quad d_{ij} < d_{rs} \ , \ then \quad \hat{\delta}_{ij}^{(0)} \leq \hat{\delta}_{rs}^{(0)} .$$

This is called weak monotonicy requirement. To obtain the disparities (fitted distances) $\hat{\delta}_{ij}^{(0)}$, a useful approximation method that is the PAV algorithm (*pool-adjacent violators*) is used.

The PAV algorithm is described as follows:
Beginning with the lowest ranked order value of $d_{ij}$, the adjacent $\delta_{ij}^{(0)}$ values are compared for each $d_{ij}$ to determine if they are monotonically related to the $d_{ij}$s. As long as the required monotonic property is true, we assign $\hat{\delta}_{ij}^{(0)} = \delta_{ij}^{(0)}$ . Whenever a block of consecutive value of $\delta_{ij}^{(0)}$ are encountered that violate the required monotonic property the $\delta_{ij}^{(0)}$ values are averaged together with the most recent non-violator $\delta_{ij}^{(0)}$ value to obtain an estimator $\hat{\delta}_{ij}^{(0)}$. Finally this value $\hat{\delta}_{ij}^{(0)}$ is assigned to all points in the particular block.

3 The metric phase. The spatial configuration of $X^{(0)}$ is altered to obtain $X^{(1)}$. $X^{(1)}$ is selected in such a way that the Stress value is minimized. From $X^{(1)}$ the new distances $\delta_{ij}^{(1)}$ can be obtained, which are more closely related to the disparities $\hat{\delta}_{ij}^{(0)}$ from step two. If the Stress value is not the desired one, go back to step two and re-calculate the distances $\delta_{ij}^{(1)}$ .

4 Terminate procedure. The Stress value is used to evaluate whether or not its change as a result of the last iteration is sufficiently small that the procedure is terminated.

**Mapping example of Non- Metric Multidimensional Scaling**

As computational methods are becoming increasingly important for detecting and visualizing structure within the vertebrate brain, it will be seen how appropriate is for data analysis / visualization the non-metric multidimensional scaling method. The anatomy of the primate visual system has been extensively studied in the macaque monkey, whose visual system is in many aspects similar to that of humans. The analysis focuses on the visual system, which includes 25 neocortical domains that are primarily or exclusively visual in function, plus 7 complementary areas that are considered visual-associative

*Table 3 Visual areas in macaque monkey.*

| Lobe | Acronym | Label |
|---|---|---|
| *Occipital* | V1 | Visual area 1 |
| | V2 | Visual area 2 |
| | V3 | Visual area 3 |
| | VP | Ventral posterior |
| | V3A | Visual area 3A |
| | V4 | Visual area 4 |
| | VOT | Ventral occipitotemporal |
| | V4t | V4 transitional area |
| | Mt | Middle temporal |
| *Temporal* | FST | Floor of superior temporal area sulcus |
| | PITd | Posterior inferotemporal (dorsal) |
| | PITv | Posterior inferotemporal (ventral) |
| | CITd | Central inferotemporal ( dorsal) |
| | CITv | Central inferotemporal ( ventral) |
| | AITd | Anterior inferotemporal (dorsal) |
| | AITv | Anterior inferotemporal (ventral) |
| | STPp | Superior temporal polysensory (posterior) |
| | STPa | Superior temporal polysensory (anterior) |
| | TF | Temporal area TF |
| | TH | Temporal area TH |
| *Parietal* | MDTd | Medial superior temporal (dorsal) |
| | MSTl | Medial superior temporal (lateral) |
| | PO | Paretio-occipital |
| | PIP | Posterior intraparietal |
| | LIP | Vateral intraparietal |
| | VIP | Ventral intraparietal |
| | MIP | Medial intraparietal |
| | MDP | Medial dorsal parietal |
| | DP | Dorsal prelunate gyrus |
| | A7a | Area 7a |
| *Frontal* | FEF | Frontal eye field |
| | A46 | Area 46 |

based on their broad visual inputs. *Table 3* shows the visual areas in the macaque monkey grouped by physical position (the lobe they are part of) (Felleman, D. J., & Van Essen, D. C, 1991). A total of 305 links have been reported between these 32 visual and visual-associative areas. We focus on the dataset of connections between areas in the primary visual cortex that have been identified by anatomical and physiological criteria.

The possibility that the visual cortex can operate through a perfect serial scheme can be ruled out only by the recognition of the diversity of connections for each area and the overall proximity of reciprocal connections. On the other hand, it seems highly unlikely that the visual cortex is a network that does not have any difference between processing levels. Many studies, based on electrophysiology, show that some visual areas, such as those of the temporal and parietal lobes, are involved in a higher degree of information processing than occipital areas such as VI and V2 (Van Essen, D. C., 1985). Between these two extremes (a perfect serial scheme on the one hand and a completely irregular network on the other), there are many intermediate possibilities. Felleman and Van Essen (Felleman, D. J., & Van Essen, D. C, 1991), published an extremely complex hierarchical scheme that includes all 32 visual cortical areas, organized into 10 hierarchical levels. There have been many at-

*Table 4 Symmetric binary connectivity matrix. Part I.*

| AREA | V1 | V2 | V3 | VP | V3A | V4 | VOT | V4t | Mt | FST | PITd | PITv | CITd | CITv | AITd |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| V1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| V2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| V3 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| VP | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| V3A | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| V4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| VOT | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| V4t | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| Mt | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| FST | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| PITd | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| PITv | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| CITd | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |
| CITv | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| AITd | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| AITv | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| STPp | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| STPa | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| TF | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| TH | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| MSTd | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| MSTl | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| PO | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| PIP | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| LIP | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 |
| VIP | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| DP | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| A7a | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| FEF | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 |
| A46 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |

tempts to draw connectivity 'maps' for different parts of the brain, which have led to connectivity diagrams of increasing complexity. Young (Young, M. P., 1992) was the first to apply MDS to this problem for the primary visual cortex.

As mentioned above, in the simplest case, the MDS takes as input a symmetric M x M matrix which describes the "dissimilarities" between a set of M objects. In this case these objects will be the 32 brain areas and the "dis-

*Table 5 Symmetric binary connectivity matrix. Part II.*

| AREA | AITv | STPp | STPa | TF | TH | MSTd | MST1 | PO | PIP | LIP | VIP | DP | A7a | FEF | A46 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| V1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| V1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| V3 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| VP | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| V3A | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| V4 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| VOT | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| V4t | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Mt | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 |
| FST | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| PITd | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| PITv | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| CITd | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| CITv | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| AITd | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| AITv | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| STPp | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| STPa | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| TF | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| TH | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 |
| MSTd | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| MSTl | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| PO | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| PIP | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| LIP | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| VIP | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| DP | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| A7a | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| FEF | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| A46 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |

similarities" will be the information whether these areas are connected or not. Based on the FVE matrix (Felleman, D. J., & Van Essen, D. C, 1991) a symmetric binary connectivity matrix is constructed. The matrix is filled with 1 and 0, where '1' means that there is no connection in any direction and '0' means there is a connection at least in one direction. As the MDP and MIP areas, in the parietal lobe, each have two exits and no entrances, they will not be included in the binary matrix as they do not provide valuable information.

So the binary matrix will be 30x30 and we obtain it by joining Table 4 and Table 5 (the whole matrix could not be contained in a single page). It has 0 in the main diagonal and is symmetric.
To solve the problem we used ALSCAL algorithm and the presentation obtained is shown at *Figure 2*. In this configuration we will not focus on the interpretation of the axes, but on the position of the dots in the two-dimensional space. It is easy to notice that this configuration has a strong annular shape. He suggests that, starting from V1, the visual information flows into two distinct and hierarchically organized streams as determined by the division and sequence of zones on the both sides of circle and then reconfirmed in A46 and TF zones. This fits in neatly with the view of the two streams in the hierarchy mentioned above published by Felleman and Van Essen. Two of the indicators that show the goodness of this model are ***Stress*** and ***RSQ*** (squared correlation between distances and deviations). A perfect solution is when the Stress is 0 and the RSQ is 1. In this procedure the Stress is 0.30299, the RSQ is 0.49554. this means that we have received a satisfactory configuration.

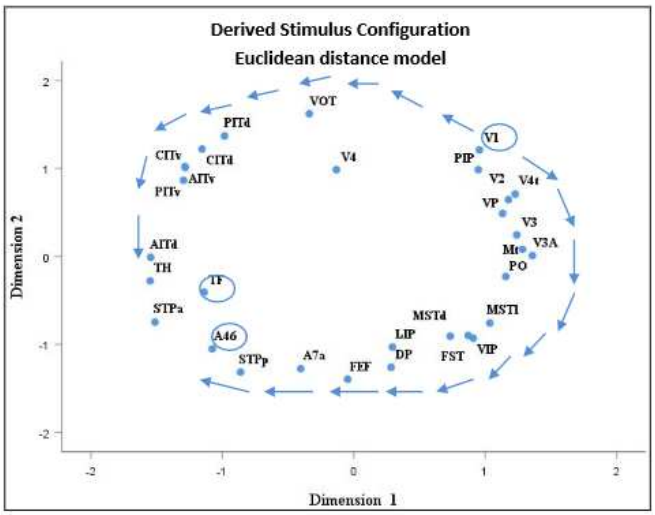


*Fig. 3. NMDS configuration for binary version of FVE matrix.*

## 3. CONCLUSIONS

In this paper, multidimensional scaling method are treated, for analysing and visualizing data, in a space as small as possible. Classical MDS, metric MDS and non-metric MDS types are treated which are part of the statistical package SPSS.
CMDS is sensitive to the type of data and the number of stimulus.Metric

scaling uses the actual values of the dissimilarities, while nonmetric scaling effectively uses only their ranks (Shepard 1962; Kruskal 1964a). Nonmetric MDS is realized by estimating an optimal monotonic transformation $f(x_{ij})$ of the dissimilarities simultaneously with the configuration. The value of stress decreases when the number of dimensionas increases.
From the examples, we can understand the usefulness of the MS method for graphical representation of data in 2-3 dimensional spaces so that the structure of distances between points is maintained and visible. In this way we can locate groupings or even extreme points.

## References

[1] Borg, I., Lingoes, J., *Multidimensional Similarity Structure Analysis*, September 1962, Volume 27.

[2] Borg, I., Groenen, P. J. F., *Modern Multidimensional Scaling*,New York: Springer, 2005.

[3] Busing F.M.T.A, Commandeur J.J.F, Heiser W.J, Bandilla W, Faulbaum F., PROXSCAL, *A multidimensional scaling program for individual differences scaling with constraints. Advances in Statistical Software*, 6:6773, 1997.

[4] Felleman, D. J., Van Essen, D. C., *Distributed hierarchical processing in the primate cerebral cortex*, Cereb Cortex, Jan-Feb 1991 .

[5] Gower, J. C., *Some distance properties of latent root and vector methods used in multivariate analysis*, Biometrika, Volume 53, Issue 3-4, December 1966.

[6] Kruskal, J. B., *Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis*, Volume 29, March 1964.

[7] Kruskal, J. B., Wish,M., *Multidimensional Scaling*, Bell Laboratories, Murray Hill, N.J. USA 1978.

[8] Mardia, K.V., J.T.Kent, J.M.Bibby, *Multivariate Analysis*, London: Academic Press, 1979

[9] Torgerson, W. S., *Theory and methods of scaling*, Volume 17, December 1952.

[10] Shepard, R. N., *The analysis of proximities: Multidimensional scaling with an unknown distance function. II*, September 1962, Volume 27.

[11] Van Essen, D. C., Maunsell, J. H. R., *Hierarchical organization and functional streams in the visual cortex*, Volume 6, 1983.

[12] Van Essen, D. C., *Functional Organization of Primate Visual Cortex*, Cerebral Cortex, Volumi 3 New York: Plenum, 1985.

[13] Young .FW, Takane Y, Lewyckyj R, *ALSCAL: A nonmetric multidimensional scaling program with several individual-differences options. Behavior Research Methods.*, 10:451453,1978.

[14] Young, M. P., *Objective analysis of the topological organization of the primate cortical visual system*, Nature, 358, 152155,1992.