On some aspects of Multi-Dimensional Scaling: Insights and Applications

Abstract:
Multi-Dimensional Scaling (MDS) is primarily a data visualization method for identifying "clusters" of points, where points in a particular cluster are viewed as being "closer" to the other points in that cluster than to points in other clusters. The traditional way of performing MDS is referred to as classical scaling which is based on the assumption that the dissimilarities are precisely Euclidean distances without any additional transformation. The particular objective function (or loss function) to be minimised is a sum of squares, commonly called *"Stress"*. The theoretical concept of MDS, its various methods, elaboration of the analytical methods involved are discussed along with applications of the different types of MDS employed (Saeed *et al.*, 2018) and illustrated with the help of examples. To understand the undergoing procedure of MDS, a two-dimensional MDS map is obtained with an illustration by making use of the MS-Excel Solver Add-in utility. The use of scree-plots to determine the optimum number of dimensions is narrated. A recent approach called majorization (de Leeuw and Mair, 2009) is pursued and shown how to 'majorize' the raw Stress function, using the *Scaling by MAjorising a COmplicated Function* (SMACOF) theory for obtaining MDS maps. In addition, these MDS procedures have been demonstrated by analysing them in R software and hence the R codes and corresponding output with interpretation are also given.

Keywords: *Stress function, Proximity, Dissimilarities, Scree-plot.*

1. Introduction

~~1.1 Multidimensional Scaling (MDS)~~

MDS is a visual representation of distances or dissimilarities between sets of objects. "Objects" can be colors, faces, map coordinates, or any kind of real or conceptual stimuli (Kruskal and Wish, 1978). Objects that are more similar (or have shorter distances) are closer together on the graph than objects that are less similar (or have longer distances). The term scaling comes from psychometrics, where abstract concepts ("objects") are assigned numbers according to a rule . For example, you may want to quantify a person's attitude to global warming. You could assign a "1" to "doesn't believe in global warming", a 10 to "firmly believes in global warming" and a scale of 2 to 9 for attitudes in-between.

Specifically, MDS represents a family ofstatistical methods or models that portray the structure of the data in a spatial fashionso that it could easily be seen and understand what the data indicate. This may bethe reason that MDS tends to be viewed as a data visual technique, and sometimesit is considered with respect to mapping technique. The unifying theme of

MDS is the spatial representation of the data structure. Considering a map of particular geographical region consisting of certain number of cities and towns. It is easy to create a map that included distances between cities. The procedure would be relatively straightforward, involving nothing more complicated than taking a ruler and measuring the distance between each city. Generally, the map will be having a two-way table indicating how close a selected number of those towns and cities are to each other. The degree of "closeness" (or *proximity*) of the city in the row to the city in the column will be shown in the cells of such table. The proximity could have different meaning: for example, proximity could be defined as straight-line distance or as shortest travelling distance. The proximity of a pair of entities (entities mean an object, a brand-name product, a nation, a stimulus etc.) could be a measure of association (e.g. the absolute value of correlation coefficient), or some other measure of how alike (or how different) one perceives the entities.

Now consider the reverse problem, where you are given the table of distances between the cities, and are asked to reconstruct the original map as closely as possible. Geometric procedures are available for this purpose, but considerably more effort would be required. The general problem of MDS essentially reverses that relationship. The typical application of MDS, however, is much more complicated than this simple example would suggest. The number of dimensions in which the given entities are located is not known . i.e., it is seldom known in advance whether a simple two-dimensional map will be adequate, or whether a "map" using three, or four, or even more dimensions is needed. So, determining the number of dimensions is another major problem to be solved. For this purposescree-plots are used. Difficulties with increasing the number of dimensionsis that even three dimensions are difficult to display on paper and are significantly more difficult to comprehend. Four or more dimensions render MDS virtually useless as a method of making complex data more accessible to the human mind.

Some of the applications of MDS are mentioned below in brief: may be constructed in a paragraph

- Marketing: Derive "product maps" of consumer choice and product preference (e.g., automobiles, beer) so that relationships between products can be discerned
- Ecology: Provide "environmental impact maps" of pollution (e.g., oil spills, sewage pollution, drilling-mud dispersal) on local communities of animals, marine species, and insects.The complex correlations betweenglobal temperature time-series using MDS has been studiedin whichMDS provides a graphical representation of the pattern of climatic similarities between regionsaround the globe(Saeed*et al.*, 2018).
- Fisheries: To study the performance of 18 marine fishery resources of the state of Maharashtra in the western region of India (Adiga*et al.*, 2016).
- Molecular Biology: Reconstruct the spatial structures of molecules (e.g., amino acids) using biomolecular conformation (3D structure). Interpret their interrelations, similarities, and differences. Construct a 3D "protein map" as a global view of the protein structure universe.

- Social Networks: Develop "telephone-call graphs," where the vertices are telephone numbers and the edges correspond to calls between them. Recognize instances of credit card fraud and network intrusion detection.

A family of different algorithms that MDS uses are designed to arrive at an optimal low-dimensional configuration for a particular type of proximity data. MDS is primarily a data visualization method for identifying "clusters" of points, where points in a particular cluster are viewed as being "closer" to the other points in that cluster than to points in other clusters.A detailed discussion regarding various MDS techniques can be found in many books, to cite a few, Kruskal and Wish (1978), Coxon (1982), Hair*et al.* (1995), Cox and Cox (2001), Borg and Groenen (2005), Izenman (2008), de Leeuw and Heiser (1977), Ding (2018), Ramasubramanian (2019) etc. de Leeuw and Mair (2009) have given a good account on MDS discussing the various versions of MDS in a lucid manner also giving an R software MDS package named SMACOF in which all the known MDS procedures are embedded.

The traditional way of performing MDS is referred to as classical scaling which is based on the assumption that the dissimilarities are precisely Euclidean distances without any additional transformation.However, the particular objective function (or loss function) used here is a sum of squares, commonly called "*Stress*". Majorization is used to minimize stress and this MDS solving strategy is known as SMACOF.In a strict sense, majorization is not an algorithm but rather a prescription for constructing optimization algorithms.

2. Definitions and examples

2.1. Proximity Matrices

The *proximity* measure gives the "closeness of two entities, which can be defined in a number of different ways. In many types of experiments, proximity data are obtained from a group of subjects, each of whom make similarity (or dissimilarity) judgements on all possible $m = \binom{n}{2} = \frac{1}{2}n(n-1)$ unordered pairs of n entities. It is irrelevant whether the similarities or dissimilarities are used as our measure of proximity between two entities. In other words, "closeness" of one entity to another could be measured by a small or large value. The only thing that matters when carrying out MDS is that there should be a monotonic relationship (either increasing or decreasing) between the "closeness" of two entities and the corresponding similarity or dissimilarity value. Anyway, usually similarities are converted into dissimilarities through a monotonically decreasing transformation. Consider a particular collection of *n* entities. Let δ_{ij} represent the dissimilarity of the *i*th entity to the *j*th entity. The *m* dissimilarities, $\{\delta_{ij}\}$, are arranged into $(m \times m)$ square matrix,

$$\Delta = (\delta_{ij}), \qquad \text{[define equation no]}$$

called a *proximity matrix*. In case of dissimilarities the proximity matrix is usually displayed as a lower-triangular array of non-negative entries, with the understanding that the diagonal entries are all zeroes and that the upper-triangular array is a mirror image of the given lower-triangle (i.e., matrix is symmetric). In other words, for all $i, j = 1, 2, \dots, n$,

$$\delta_{ij} \geq 0, \qquad \delta_{ii} = 0, \qquad \delta_{ji} = \delta_{ij}.$$ [define equation no]

2.2. Stress function

So far, the task of MDS was defined as finding a low-dimensional configuration of points representing objects such that the distance between any two points matches their dissimilarity as closely as possible. Of course, it is preferred that each dissimilarity should be mapped exactly into its corresponding distance in the MDS space. But empirical data always contain some component of error given by $f(\delta_{ij}) - d_{ij}$,where d_{ij}'s are the computed Euclidean distances between the objects in the arbitrarily constructed plot. Since positive and negative discrepancies are equally undesirable, the sum of squared errors for all proximities is taken, which yields the formula.

$$Raw\ stress = \sum_i \sum_j \left(\delta_{ij} - d_{ij}\right)^2, \text{ by taking } f(\delta_{ij}) = \delta_{ij}.$$ [define equation no]

To counter the effect of scale-dependency, the raw stress is normalised to have the general form,

$$\left\{\sum_{i<j} w_{ij}\left(\delta_{ij} - d_{ij}\right)^2\right\}^{1/2},$$ [define equation no]

where the $\{w_{ij}\}$ are weights chosen by the user. The most popular normalization is where $w_{ij} = \left(\sum_{i<j} {d_{ij}}^2\right)^{-1}$, so that the raw stress become the Stress1 i.e.,

$$Stress1 = S = \left\{\frac{\sum_{i<j}\left(\delta_{ij} - d_{ij}\right)^2}{\sum_{i<j} {d_{ij}}^2}\right\}^{\frac{1}{2}},$$ [define equation no]

where it is understood that the summations in both the numerator and denominator of *S*are computed for all $i, j = 1, 2, \dots, n$ such that $i < j$. The stress-1 value (*S*) lies between 0 and 1.The stress criterion *S* (more commonly known as *Kruskal's stress formula one* or Stress-1) can be interpreted as a loss function that depends upon the configuration points and the disparities and measures how well a particular configuration fits the given dissimilarities. It is worth noting that certain authors refer to the stress function asS^2.A variant, stress formula 2, differs only in that a different weights are used.

2.3. Scree-plot:

A scree-plot is a method for determining the optimal number of components useful to describe the data in the context of MDS. To create a Scree-plot, analysts scale the data several times (with higher dimensionality each time), and plot the stress values as a function of dimensions (Hout *et al*., 2013). Here, the stress values are plotted on y-axis and the number of dimensions are plotted on x axis as shown in fig (i). The aim is to evaluate the number of dimensions required to capture most information contained in the data. The 'elbow' of the plot determines the optimal number of dimensions to describe the data.

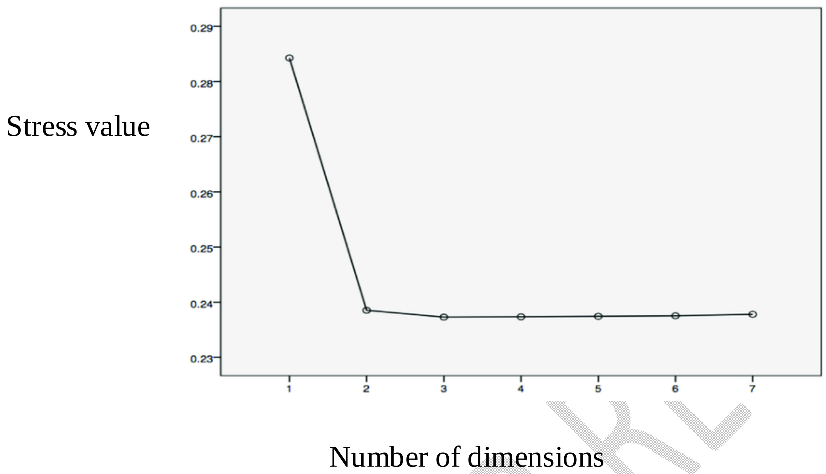


Fig (i) Scree-plot (figure to be constructed appropriately and NOT pasted as a screen shot; Fig no to be 1 not (i); *x* and *y* axis title fonts to be Times New Roman 10 to 12 size)

Normally, a complex set of relationships can be scanned at a glance with the aid of visual representation provided by MDS. Since maps on paper are two-dimensional objects, this translates technically to finding an optimal configuration of points in 2-dimensional space. However, limiting to a two dimensions may lead to a very poor, highly distorted, representation of the data. In order to overcome this limitation the number of dimensions may be increased(if needed) but there are difficulties in representing, comprehending and estimating the parameters for the higher dimensions. Four or more dimensions render MDS virtually useless as a method of making complex data more accessible to the human mind.

Example 1. Application to Perceptions of Nations:

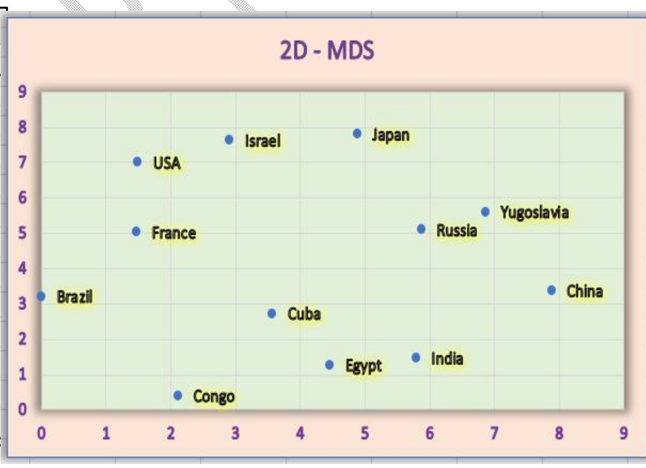
Now the procedure followed to obtain the two-dimensional MDS plots are discussed with an illustration using excel where the dissimilarity matrix is given. The data reflecting mean scores of 18 respondents' perceptions of overall dissimilarity between twelve nations on a scale ranging from 1 for "very familiar" to 9 for " very different" was ordered as a diagonal matrix of 66 pairs (Kruskal and Wish, 1978) as shown in fig (ii). Since it has been decided on a two dimensional representation of the data, a starting configuration for the n objects in the two dimensions has to be set up(i.e. Co-ordinates x_n,,y_n are arbitrarily selected for each object) represented in the left side of fig (iii).The next step involves calculating the Euclidean distances between the objects.Howeverthe data points arranged within the graph will always be a difference between the actual values in our original diagonal matrix and the inter-point distances reflected and measured in the graph. Even after trying thousands of different arrangements there will still be errors and the best option is to minimize the cumulative errors in an arrangement, i.e. minimise the stress and show that as the best representation made out of the data provided. In other words it would be a trial and error or iterative process of finding the best cumulative error minimising arrangement. Making use of the built in facility within Microsoft Excel(i.e. in Solver Add-in which is part of the Microsoft package)the solutions to such problems are found (Anonymous, 2007).Hence the optimised values of the co-ordinates (x and y, shown on right side of fig (iii)) are obtained by minimising the stress. And finally these points are plotted in a 2 dimensional scatter plot as shown in the fig (iii).

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		Brazil	Congo	Cuba	Egypt	France	India	Israel	Japan	China	Russia	USA	Yugoslavia
2	Brazil	0											
3	Congo	4.17	0										
4	Cuba	3.72	4.44	0									
5	Egypt	5.56	4	3.83	0								
6	France	4.28	5	4.89	4.22	0							
7	India	4.5	4.17	5	3.17	5.56	0						
8	Israel	5.17	5.67	5.39	4.33	5	4.89	0					
9	Japan	5.5	5.61	6.06	5.17	4.78	4.5	4.17	0				
10	China	6.61	5	3.5	4.61	5.33	4.89	6	4.83	0			
11	Russia	5.94	5.61	3.56	4.61	3.94	4.5	4.83	4.39	3.28	0		
12	USA	3.61	6.61	5.83	5.67	3.06	4.72	3.06	2.94	6.44	4	0	
13	Yugoslavia	5.83	5.5	3.89	4.72	4.28	5	4.56	4.72	3.94	2.23	5.44	0

Fig (ii). Representation of the dissimilarity matrix

(figure should be typed out in proper format as a TABLE)

S	T	U
	y	x
Brazil	0	0
Congo	4.17	0
Cuba	0	3.72
Egypt	5.56	0
France	0	4.28
India	4.5	0
Israel	0	5.17
Japan	5.5	0
China	0	6.61
Russia	5.94	0
USA	0	3.61
Yugoslavia	5.83	0



O	P	Q
	y	x
Brazil	0.001838	3.362963
Congo	2.060104	0.288286
Cuba	3.203355	2.524228
Egypt	4.439001	1.350378
France	1.642403	5.163494
India	5.714004	1.191912
Israel	3.297548	7.593489
Japan	5.302053	7.59025
China	7.910626	2.985486
Russia	5.99006	4.831061
USA	1.818794	7.085297
Yugoslavia	7.00531	5.269065

Fig (iii). Initial values of the co-ordinates (left), optimised values obtained after using SOLVER (right) and the MDS plot (middle).

(the two overlapping figures should better be segregated for clarity)

Example 2. An Application of MDS to Morse Code Confusions Data:

Figure (iv) consists of confusions among 36 auditory Morse code signals (Kruskal and Wish, 1978). Each signal consists of a sequence of dots and dashes, such as · for K and · · for 2. Subjects who did not know Morse code listened to a pair of signals and were required to state whether the two signals they heard were the same or different. Each number in the table is the percentage of roughly 150 observers who responded "same" to the row signal followed by the column signal. Notice that the conventional Morse code letter names for the signals do not enter into the experiment in any way, and are used in the table simply as convenient names.The diagonal of the table corresponds to pairs which are truly the same, so it is expected the diagonal entries to be large. Off-diagonal entries correspond to pairs which are truly different, so they are expected to be smaller.Figure (v) shows the result of applying MDS to the proximities in Fig (iv).To this configuration we have added the letters, for convenience in reference, and the dot-dash description of the auditory signals which the subjects heard.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6	7	8	9	0	
A	82	4	5	13	3	14	10	15	46	5	22	3	25	34	6	0	9	35	23	6	37	13	17	12	7	3	2	7	5	3	6	6	5	6	2	3	A
B	8	84	37	31	5	28	17	21	5	19	34	40	6	10	12	22	25	16	18	2	18	34	8	84	30	42	12	17	14	40	32	74	43	17	4	4	B
C	4	38	87	17	4	29	13	7	11	19	24	35	14	3	9	51	34	24	14	6	6	11	14	32	2	30	13	15	31	14	10	30	28	24	15	12	C
D	6	42	17	88	7	23	40	36	9	13	81	56	8	7	9	27	9	45	29	6	17	20	27	40	15	33	3	9	6	11	9	19	8	10	5	6	D
E	6	13	14	6	97	2	4	4	17	1	5	6	4	4	5	1	5	10	7	67	3	3	2	5	6	5	4	3	5	3	5	2	4	2	3	3	E
F	4	1	33	19	2	90	10	29	5	33	16	50	7	6	10	42	12	35	14	2	21	27	25	19	27	13	6	16	41	25	26	24	21	5	5	5	F
G	9	10	27	38	1	14	90	6	5	22	33	16	14	13	82	52	23	21	5	3	15	14	32	21	23	39	15	14	5	10	4	10	13	23	20	11	G
H	0	45	23	25	9	2	6	7	10	10	9	29	5	0	0	14	6	17	37	4	36	59	8	33	14	11	3	9	15	43	70	35	13	4	3	3	H
I	4	7	7	13	10	8	6	12	93	3	5	16	13	30	7	3	5	19	35	16	10	5	8	2	5	7	2	4	6	9	6	0	5	2	4	5	I
J	7	7	34	9	2	24	10	5	4	5	22	31	0	5	21	63	47	11	2	7	9	9	9	22	32	28	67	66	33	15	7	11	28	29	26	25	J
K	5	24	38	73	1	17	25	11	5	27	91	33	10	12	31	14	31	22	2	2	23	17	33	63	16	18	5	9	17	8	8	18	14	13	3	6	K
L	2	69	43	45	10	24	12	26	9	30	27	56	6	2	9	37	36	28	12	5	16	19	20	31	25	59	12	13	17	15	26	29	36	16	7	3	L
M	24	12	5	14	7	17	27	4	6	11	23	6	96	62	11	10	15	20	7	9	13	4	21	9	10	6	5	7	6	5	5	7	11	7	10	4	M
N	31	4	13	30	0	12	10	16	13	3	16	8	59	83	5	9	5	28	12	10	16	4	12	4	6	11	5	2	3	4	4	6	2	2	10	2	N
O	7	7	20	6	5	9	76	7	2	39	26	10	4	8	86	37	35	10	3	4	11	14	25	35	27	27	19	17	7	7	6	18	14	11	20	12	O
P	6	22	33	12	5	36	22	12	3	78	14	46	9	6	21	3	43	23	9	4	12	19	19	19	41	30	34	44	24	11	15	17	24	25	25	13	P
Q		20	30	11	4	15	10	5	2	27	23	26	7	6	22	51	91	11	2	3	6	14	12	37	50	63	34	32	17	12	9	27	40	58	37	24	Q
R	13	14	16	23	5	34	26	15	7	12	21	37	14	12	12	29	8	87	16	2	23	23	62	14	12	13	7	10	13	4	4	12	7	9	0	2	R
S	17	24	3	30	11	26	5	38	16	3	13	10	5	17	6	6	3	18	96	9	6	24	12	10	6	7	8	2	2	15	26	9	5	5	5	2	S
T	13	10	1	5	46	3	6	6	14	6	14	7	6	5	6	11	4	4	7	96	6	5	4	2	2	6	0	5	3	3	3	8	7	6	14	6	T
U	14	29	12	32	4	32	11	34	21	7	44	32	11	13	6	20	12	40	1	6	93	7	34	17	9	11	6	6	16	34	10	9	9	7	4	3	U
V	6	17	24	16	0	29	6	39	5	11	26	43	4	1	9	17	10	17	11	6	32	92	17	57	35	10	10	14	25	79	44	30	25	10	1	5	V
W	9	21	30	27	9	36	25	15	4	25	29	18	15	6	26	20	25	61	12	4	19	20	86	22	25	22	10	22	19	16	5	9	11	6	3	7	W
X	7	64	45	19	3	28	11	6	1	35	50	42	10	6	24	32	61	10	12	3	12	17	21	91	48	26	12	20	24	27	16	37	29	16	17	6	X
Y	9	23	67	15	4	26	22	9	8	30	12	14	3	6	14	30	2	3	1	4	6	13	21	44	86	23	26	44	40	15	11	26	22	33	23	16	Y
Z	3	16	45	18	2	22	17	10	2	23	21	51	11	2	15	59	72	14	4	3	9	11	12	36	42	7	16	21	27	9	10	29	66	47	15	15	Z
1	2	5	10	3	3	5	13	4	2	29	5	14	9	7	14	30	28	9	4	2	3	12	14	17	19	22	4	63	13	0	10	0	19	32	57	55	1
2	7	14	22	5	4	20	13	3	25	26	9	14	2	0	17	37	28	6	5	3	6	10	11	17	30	13	62	89	34	20	5	14	20	21	16	11	2
3	3	6	21	2	4	32	6	12	8	23	6	13	5	2	5	37	19	9	7	6	4	16	6	22	25	12	18	64	6	31	23	41	16	17	8	10	3
4	6	19	10	12	6	25	14	16	7	21	13	18	3	3	2	17	29	11	9	3	17	55	8	37	24	3	5	26	44	9	42	44	32	10	3	3	4
5	6	45	15	14	2	45	4	47	7	14	4	41	2	0	4	13	1	9	27	2	14	45	7	45	10	10	14	10	30	69	90	42	24	10	6	5	5
6	7	80	30	13	4	23	4	14	8	11	11	27	6	2	7	16	30	11	14	3	12	30	9	58	38	39	15	14	26	24	17	6	68	14	5	14	6
7	6	33	22	14	5	25	6	4	6	24	13	32	7	6	7	36	39	12	6	2	0	13	9	30	30	50	22	29	14	15	12	41	45	70	20	13	7
8	3	23	40	6	3	15	15	0	2	33	10	14	0	6	14	12	45	2	6	4	6	7	5	24	35	50	42	29	16	16	9	30	60	69	41	26	8
9	3	14	23	3	1	4	14	5	2	30	5	7	16	11	10	31	32	0	6	7	6	3	8	11	21	24	57	32	9	12	4	11	42	6	91	18	9
0	2	3	1	2	0	7	14	4	5	30	8	0	2	3	25	21	29	2	3	4	5	3	2	12	15	20	50	26	0	11	5	22	13	52	81	94	0
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6	7	8	9	0	

Fig (iv) Representation of Morse code data.

In this application the proximities are similarities, that is, a large value means that the two signals are very much alike. Consider two signals, for example, B · · · and X · · , which have largesimilarity values, 84% and 64%. In the geometric configuration the points for B and X are very close together. Likewise consider two signals, for example E · and 0 , which have very smallsimilarity values, 3% and 5%. In the geometric configuration the points for E and 0 are very far apart. For other signals, the same thing holds true: a large similarity value corresponds to asmall distance, and a small similarity value corresponds to a large distance. Despite a few exceptions, there is a clear relationship which we display more fully later. This is what we mean bysaying that the geometric configuration reflects the proximity values.The picture on the right side of the Fig. (v) indicates two things:

(i) The number of components (dots and dashes) increases from bottom to top and

(ii) Dots are more on left and Dashes are more on the right.

And also when it is compared with the three-dimensional MDS plot, it can be seen that the 'O' appears to be closer to 'I' than '9' in two-dimensional MDS plot but in the three-dimensional MDS plot '9' appears to be closer to 'I' than 'O' (marked by small circles in fig (vi)).

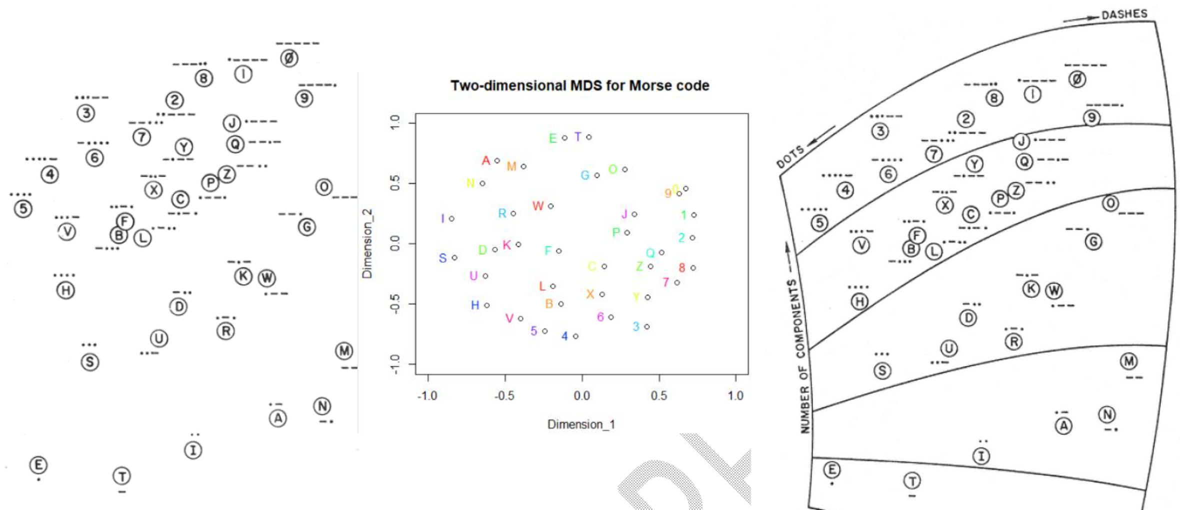


Fig (v) Result of applying MDS to the proximities of Morse code data.

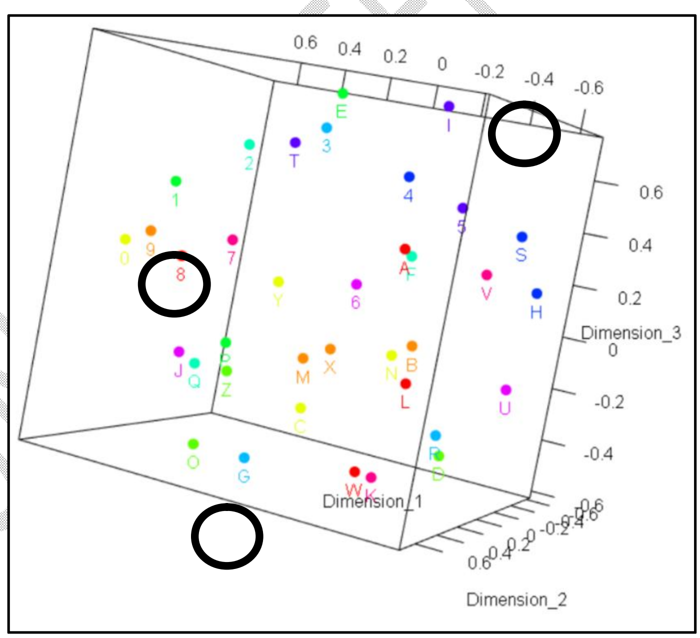


Fig (vi) Three-dimensional MDS plot for Morse code data.

2.4SMACOF :

The Stress function that measures the deviance of the distances between points in a geometric space and their corresponding dissimilarities is to be minimised.An easy and powerful minimization strategy is the principle of minimizing a function by iterative majorization. Because for finding the minimum of a function $f(x)$, it is not always enough to compute the derivative $f'(x)$, set it equal to zero, and solve for x. Sometimes the derivative is not defined everywhere, or solving the equation $f'(x) = 0$ is simply impossible. For such cases, other mathematical techniques are referred. A useful method consists of trying to get increasingly better estimates of the minimum. It consists of a set of computational rules that are usually applied repeatedly, where the previous estimate is used as input for the next cycle of computations which outputs a better estimate. An elegant method is called iterative majorization. The main principles of iterative majorization are:

Principles of Majorization

The central idea of the majorization method is to replace iteratively theoriginal complicated function f(x) by an auxiliary function g(x, z), where z in g(x, z) is some fixed value. Thefunction g hastomeet thefollowing requirements to call g(x, z)a majorizing function of f(x).The auxiliary function g(x, z) should be simpler to minimize than f(x). For example, if g(x, z) is a quadratic function in x, then the minimum of g(x, z)over x can be computed in one step.The original function must always be smaller than or at most equal to the auxiliary function; that is, f(x) ≤ g(x, z).The auxiliary function should touch the surface at the so-called supporting point z;thatis, f(z)= g(z, z).

Hence, the iterative majorization algorithm is given by

1. Set $z = z^0$, where z^0 is a starting value.
2. Find update x_u for which g(x_u, z) ≤ g(z, z).
3. If $f(z) - f(x_u) < \varepsilon$,thenstop.(ε is a small positive constant)
4. Set $z = x_u$ andgoto step 2.

3.Illustration using a practical dataset related to agriculture:

In a study, information from experts wasobtained through questionnaires for identification of specific technologies/ scientific development that need major attention for increasing the productivity of cereals, pulses and oilseeds in India which was then statistically analyzed for prioritizingfuture technological needs (Ramasubramanian*et al.*, 2014). Attempts are made to analyze theavailable information using MDS approach. A total of 35experts respondedfor ranking the factors responsible for enhancing agriculturalproductivity. The data is represented in fig (vii).

In order to study the experts' perceptions of important factors attributable to agricultural growth, the responses (differing in their levels of importance as viewed by the experts) were considered two at a time ("all-pairs design").Thus the responses (on a five point score from 0 to 4)of experts for the possible ${}^{10}C_2$ = 45 pairs of factors were collated. The rating for each pair of factors was averaged over all respondents and the result divided by 4 to bring the similarity ratings into the interval (0,1). These mean similarity values were then collected into a (10 x 10) table, which can then be treated as a correlation-like matrix. The similarities were converted into dissimilarities which are tabulated below.

Factors	1 (1.00)	2 (0.75)	3 (0.50)	4 (0.25)	5 (0.00)	Score
Quality seed availability	25	6	2	0	0	30.5
Better varieties	22	8	3	0	0	29.5
Timely availability of inputs	11	20	2	0	0	27.0
Proper research infrastructure	16	11	4	2	0	26.8
Better agronomic practices	11	15	7	0	0	25.8
Adaptation to changing climatic and environmental scenario	12	12	7	2	0	25.0
Marketing facilities	11	11	8	3	0	25.0
Minimum Support Price (MSP)	11	10	10	2	0	24.0
Development of location specific technologies	9	13	8	3	0	23.5
Better extension services	11	8	12	2	0	23.5
Genetically Modified (GM) crops	3	18	10	2	0	22.0
Technology to fill the gap between nutrient depletion and supply to soil	3	18	10	2	0	22.0
Nutrient management	5	15	10	3	0	22.0
Plant Protection measures including IPM	4	14	13	1	1	21.3
Use of Information and Communication Technology (ICT)	4	13	14	2	0	21.3
Post harvest management	3	15	11	4	0	20.8
Domestic / International trade	5	11	13	4	0	20.8
Farm mechanization	5	9	15	4	0	20.3

Fig (vii) Representation of dataset related to agriculture. [convert to table]

[no table no below]

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
F1	0.00	0.23	0.27	0.15	0.25	0.24	0.22	0.23	0.21	0.22
F2	0.23	0.00	0.09	0.18	0.25	0.26	0.22	0.10	0.19	0.22
F3	0.27	0.09	0.00	0.21	0.26	0.28	0.25	0.13	0.16	0.19
F4	0.15	0.18	0.21	0.00	0.21	0.20	0.14	0.16	0.15	0.16
F5	0.25	0.25	0.26	0.21	0.00	0.23	0.24	0.25	0.25	0.21
F6	0.24	0.26	0.28	0.2	0.23	0.00	0.10	0.22	0.25	0.20
F7	0.22	0.22	0.25	0.14	0.24	0.10	0.00	0.21	0.19	0.20
F8	0.23	0.10	0.13	0.16	0.25	0.22	0.21	0.00	0.16	0.21
F9	0.21	0.19	0.16	0.15	0.25	0.25	0.19	0.16	0.00	0.16
F10	0.22	0.22	0.19	0.16	0.21	0.20	0.20	0.21	0.16	0.00

where

F1 - Quality seed availability

F2 - Better varieties

F3 - Timely availability of inputs

F4 - Proper research infrastructure
F5 - Better agronomic practices
F6 - Adaptation to changing climatic and environmental scenario
F7 - Marketing facilities
F8 - Minimum Support Price (MSP)
F9 - Development of location specific technologies
F10 - Better extension services

Using the below mentioned R codes the MDS of 1,2,3,4,5 dimensions were fitted and the respective stress value vs dimension were plotted to obtain a scree-plot as shown in Fig. (viii).

```
ag=read.csv(file.choose())
agg=ag[-1]
head(agg)
rownames(agg)=colnames(agg)
aggmds1 = smacofSym(delta = agg,ndim = 1, type = "ratio" )
aggmds2 = smacofSym(delta = agg,ndim = 2, type = "ratio" )
aggmds3 = smacofSym(delta = agg,ndim = 3, type = "ratio" )
aggmds4= smacofSym(delta = agg,ndim = 4, type = "ratio" )
aggmds5 = smacofSym(delta = agg,ndim = 5, type = "ratio" )
slotNames(summary(aggmds3))
#######______screeplot--------
stress=c(aggmds1$stress,aggmds2$stress,aggmds3$stress,
         aggmds4$stress,aggmds5$stress)
dimensions=c(1:5)
screeplot= plot(dimensions,stress,type = "b")
```

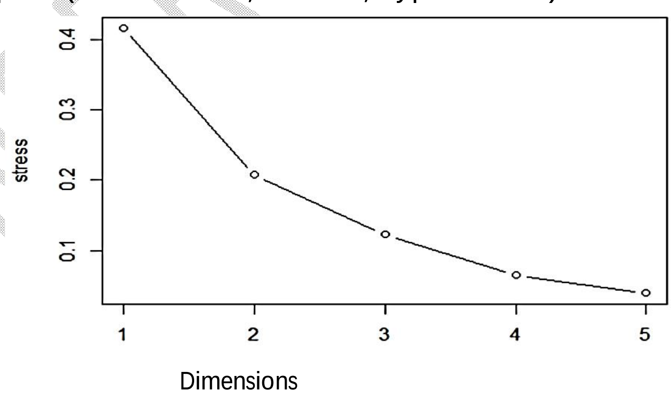


Fig (viii) Scree-plot obtained for the practical dataset.

With the aid of the scree-plot it is found that 3 dimensional MDS would be more appropriate. For the comparison sake, both the 2 dimensional and 3 dimensional MDS plots are obtained. The 2 dimensional plot is obtained using the following R codes

```
zz=matrix(c(  -0.1721, -0.7945,
```

```
0.6601,-0.1938,
           0.7472,  0.1648,
           -0.0980, -0.2118,
            -0.4939,  0.7607,
          -0.8062, -0.0477,
            -0.5458, -0.1848,
              0.4331, -0.3045,
0.3099,  0.3091,
            -0.0344,  0.5025),10,byrow = T)
x <- zz[, 1]
y <- zz[, 2]
plot(x, y,xlab = "Dimension_1",
     ylab = "Dimension_2",
main = "Two-dimensional MDS ",xlim = c(-1,1),ylim = c(-0.6,0.8))
text(x,y,labels = rownames(agg),col =rainbow(11),pos = 2)
```

The 2 dimensional plot obtained is shown in fig (ix)

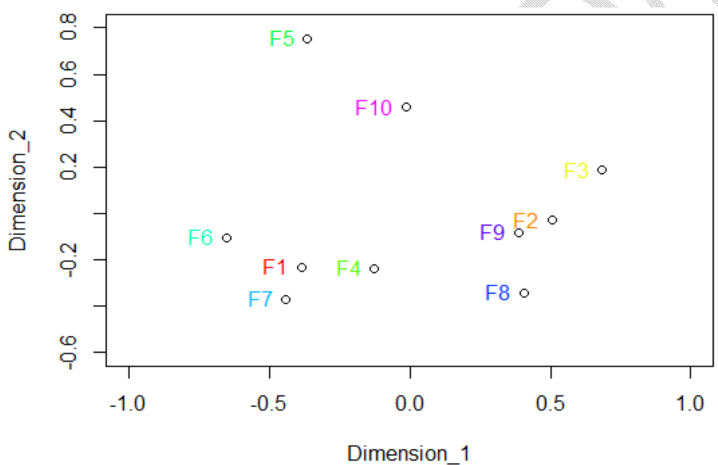


Fig (ix) Two-dimensional MDS plot.

Similarly the 3 dimensional MDS was obtained by using the following R code:

```
zzz=matrix(c( -0.3870 ,-0.2334 , 0.6767
            , 0.5048 ,-0.0301 ,-0.3832
             ,    0.6853,  0.1862, -0.1379
             ,   -0.1254, -0.2381,  0.1989
             ,   -0.3648,  0.7538, -0.2317
             ,   -0.6523, -0.1045, -0.4210
             ,   -0.4402, -0.3687, -0.2364
             ,    0.4051, -0.3406, -0.2151
             ,    0.3870, -0.0817,  0.4253
             , -0.0124,   0.4571,  0.3244),10,byrow = T)

x <-zzz[, 1]
y <- zzz[, 2]
z= zzz[,3]
```

```
#library(rgl)
plot3d(x,y,z,xlab = "Dimension_1",
       ylab = "Dimension_2",
zlab="Dimension_3",
       col = rainbow(11),size = "10")
text3d(x,y,z,row.names(agg),pos=1,col=rainbow(11))
```

The plots are depicted in fig (x).

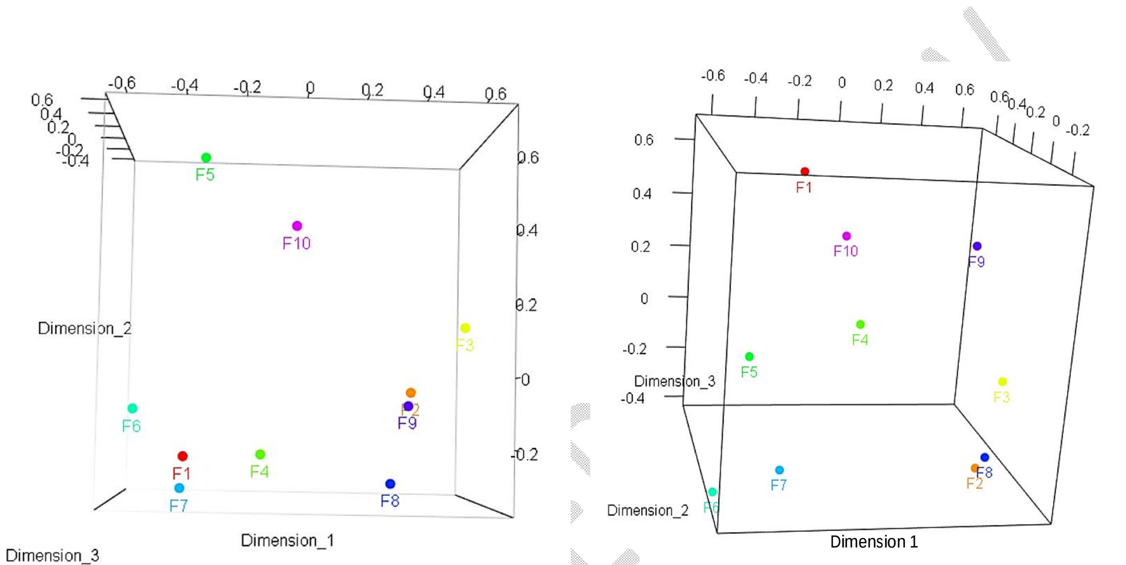


Fig (x) Three-dimensional MDS plots.

The three dimensional plot provides more information as in the two dimensional plot it can be seen that F1 and F7 and close together which is also represented in the left three dimensional plot, but the actual distance in three dimension between F1 and F7 can be seen in the right three dimensional plot. As said earlier it is difficult to analyse and interpret the higher dimensional MDS. [Specific reasons may be stated]

4. Concluding remarks: **[little more elaboration will be more meaningful]**

MDS is used as a data visualisation technique as the data is made simple and easily understandable by portraying the structure of the data in a spatial fashion. MDS is a useful tool to quantify the similarity / closeness between the things which is empirically difficult to measure. By subjecting similarity estimates to MDS, researchers acquire maps of the relationships among a set of stimuli. This map reduces the complexity inherent to a large table of proximities. Even though it is difficult to choose the number of dimensions to be employed, scree-plot guide for the appropriate selection. Basic idea on MDS is well illustrated using examples and software such as MS-Excel, R. Hence, MDS has a broad applicability including in the areas of Marketing, Ecology, Molecular biology, Network and systems, etc.

5. References:

Adiga, M. S., Ananthan, P. S., Kumari, H. D., and Ramasubramanian, V. (2016). Multidimensional analysis of marine fishery resources of Maharashtra, India. *Ocean and Coastal Management*, *130*, 13-20.

Anonymous (2007). Constructing perceptual maps with the aid of SOLVER in Office 2007, http://cw.routledge.com/textbooks/9780415458160/instructorresources/MDS%20with%20spreadsheet.docx, accessed on 21 December, 2019.

Borg, I. and Groenen, P.J.F. (2005). *Modern Multidimensional Scaling: Theory and Applications*, Second edition, Springer-Verlag, New York.

Cox, T.F. and Cox, M.A.A. (2001). *Multidimensional Scaling*, Second edition. Chapman and Hall/CRC, Boca Raton.

Coxon, A.P.M. (1982). *The User's guide to Multidimensional Scaling*, Heinemann Educational Books, Great Britain.

de Leeuw, J., and Heiser, W. J. (1977). Convergence of correction matrix algorithms for multidimensional scaling. *Geometric representations of relational data*, 735-752.

de Leeuw, J., and Mair, P. (2009). Multidimensional scaling using majorization: SMACOF in R, *Journal of Statistical Software*, **31(3)**, 1-30. https://escholarship.org/uc/item/9z64v481

Ding, C. S. (2018). *Fundamentals of applied multidimensional scaling for educational and psychological research*. Springer.

Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1995). *Multivariate data analysis*, 4th Edition, Prentice Hall, New Jersey.

Hout, M. C., Papesh, M. H., and Goldinger, S. D. (2013). Multidimensional scaling. *Wiley Interdisciplinary Reviews: Cognitive Science*, **4(1)**, 93-103.https://doi.org/10.1002/wcs.1203

Izenman, A.J. (2008). *Modern multivariate statistical techniques: Regression, classification and manifold learning*. Springer, New York.

Kruskal, J.B. and Wish,M. (1978), *Multidimensional Scaling*, Series: Quantitative applications in the social sciences, Sage University Press, , California.

Ramasubramanian V. (2019). Multi-Dimensional Scaling, Reference Book on *Quantitative Methods for Social Sciences*, Ed. Nikam, V., Jhajhria, A. and Pal, S., ICAR-National Institute for Agricultural Economics and Policy Research, New Delhi, pages 31-45, ISBN 978-81-940080-2-6.

Ramasubramanian, V., Kumar, A., Prabhu, K. V., Bhatia, V. K., andRamasundaram, P. (2014). Forecasting technological needs and prioritizing factors in agriculture from a plant breeding and genetics domain perspective: a review. *Indian Journal of Agricultural Sciences*, *84*(3), 311-316.

Saeed, N., Nam, H., Haq, M. I. U., and Muhammad Saqib, D. B. (2018). A survey on multidimensional scaling. *ACM Computing Surveys (CSUR)*, **51(3)**, 1-25.