Visualizing big data via a mixture of PARAMAP and Isomap

Ulas Akkucuk

Assoc. Prof. of Quantitative Methods, Bogazici University, Department Management, Bebek, 34342, Istanbul, Turkey

A B S T R A C T

Dimension reduction strives to represent higher dimensional data by a lower-dimensional structure. A famous approach by Carroll called Parametric Mapping or PARAMAP (Shepard & Carroll, 1966) works by iterative minimization of a loss function measuring the smoothness or continuity of the mapping from the lower dimensional representation to the original data. The algorithm was revitalized with essential modifications (Akkucuk & Carroll, 2006). Even though the algorithm was modified, it still needed to make a large number of randomly generated starts. In this paper we discuss the use of a variant of the Isomap method (Tenenbaum et al., 2000) to obtain a starting framework to replace the random starts. The core set of landmark points are selected by a special procedure akin to selection of seeds for the k-means algorithm. These core set of landmark points are used to create a rational start for running the PARAMAP algorithm only once but effectively reach a global minimum. Since Isomap is faster and less inclined to local optimum problems than PARAMAP, and the iterative process involved in adding new points to the configuration will be less time consuming (since only one starting configuration is used), we believe the resulting method should be better suited to deal with large data sets, and more prone to obtain an acceptable solution in realistic time.

Introduction

The paradigm of “information explosion” or the exponential growth of available data is seen in many fields including business, genetics and other life sciences. In order to make sense of large amounts of data, techniques like MDS, which can reduce the dimensionality and create meaningful visual maps, can be very helpful in terms of analyzing and interpreting the data in a meaningful parsimonious fashion. The benefits of visualization have been demonstrated in the literature, Al-Kassab et al. stated “Information visualization can accelerate perception, provide insight and control, and harness this flood of valuable data to gain a competitive advantage in making business decisions” (2014). Visualization via dimension reduction algorithms may be one of the most important tools in dealing with Big Data problems. Algorithmic speed in such visualization tasks may be critical as some tasks may require immediate managerial attention. This paper proposes a clever starting configuration for a nonlinear mapping technique that can speed up the visualization process.

Dimension reduction techniques involve algorithms aiming to find a lower dimensional structure which preserves the characteristics of the higher dimensional input configuration. Applications of dimension reduction are found in a wide range of areas, including visualization, pattern analysis, data preprocessing, scale development, cybernetics, and localization (France & Carroll, 2011). Some applications of MDS in social sciences that could be found in the literature could be listed as follows:

i. Perceptual map of universities based on similarity judgments (Akkucuk & Kucukkancabas, 2007).
ii. Mapping countries based on Global Competitiveness index using derived distances rather than perceived similarities (Akkucuk, 2011; 2014).
iii. Using Supreme Court judges’ agreement data to locate judges on maps of different dimensionality (Akkucuk, Carroll & France, 2013).

© 2020 by the authors. Licensee SSBFNET, Istanbul, Turkey. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).
iv. Visualization of countries by using the patent applications in various technological categories (Akkucuk & Artemel, 2016).

Nonlinear mapping methods concentrate in structures where the relationship between the coordinates is highly nonlinear. Major examples to nonlinear surfaces are spheres, tori and spirals. PARAMAP (Shepard & Carroll, 1966) approach and Isomap (Tenenbaum et al., 2000) approach are two procedures that deal with such nonlinear surfaces. An evaluation of PARAMAP and Isomap algorithms was provided by Akkucuk and Carroll (2006). In this former article, solutions provided by Isomap and PARAMAP were compared using various performance criteria. The main conclusion of Akkucuk and Carroll was that Isomap was not designed to handle closed surfaces and also open surfaces with insufficient point density. PARAMAP could handle closed surfaces well but computation time was excessive. In the conclusion it was stated that:

“When there is prior knowledge of the structure of the data being examined, the decision of selecting the correct program to run would be trivial. However, this might not be so easy with real world data. Combining the ability of PARAMAP to deal with closed surfaces and the ability of Isomap to reach a global optimum in reasonable time, could result in even more effective and efficient algorithms for nonlinear mapping. It is apparent with increased computing power PARAMAP is more likely to become a choice in analyzing vast amounts of data because it makes minimal assumptions about the nature of the nonlinearity of the relationship between the higher dimensional data set and desired lower dimensional structure.”

In this paper our purpose is to explain a hybrid technique that will combine vital features of Isomap and PARAMAP. The algorithm was also explained in Akkucuk and Carroll (2010) however some details about landmark selection from large data sets were not provided. In addition there is increased interest in visualization of big data so the algorithm may have a more thorough evaluation with the current advances in computing. Transformation of the program to user friendly open source statistical packages like “R” is also in progress. The algorithm uses Isomap as a predecessor to PARAMAP; therefore the resulting algorithm does not necessitate different random starts, but will entail a number of quick Isomap runs followed by one PARAMAP run. The primary Isomap run will end in either an incomplete or complete solution (in the sense of the lower dimensional embedding containing the same number of points as the higher dimensional input configuration). If the solution is incomplete, the points in the lower dimensional solution provided by Isomap will be viewed as the landmark points. The suggested algorithm creates the starting configuration from the incomplete solution containing the landmark points and uses this “rational” starting configuration as input to PARAMAP.

This paper will proceed as follows: The following section will review the salient technical features of PARAMAP and Isomap. The next section will try to explain why large data sets pose a problem for PARAMAP and detail a previously proposed method used to let PARAMAP deal with a large number of points. After that the two sections will present the new proposed hybrid approach combining Isomap. The final section will offer the conclusions and directions for future research.

PARAMAP and Isomap Algorithmic Details

The technical features of the two algorithms were deliberated broadly in a number of former papers (Akkucuk, 2004; Akkucuk & Carroll, 2006; Tenenbaum et al., 2000; Shepard & Carroll, 1966). Here we will provide a summary of the essential terminology for understanding the arguments we make in this article.

PARAMAP Algorithm Details

The PARAMAP algorithm tries to discover a lower dimensional configuration taking as input a higher dimensional configuration. In doing so PARAMAP minimizes a measure of continuity called “kappa” or abbreviated as \( \kappa \), that relates the lower dimensional configuration to the higher dimensional configuration. The kappa function is calculated by using the input dissimilarities (to be called \( \delta_{ij} \)) calculated from the higher dimensional input configuration and the distances computed from a lower dimensional configuration (to be called \( d_{ij} \)). The kappa function is computed as follows:

\[
\kappa = \sum_{i\neq j} \frac{\delta_{ij}^2}{d_{ij}^4} \left( \sum_{i\neq j} \frac{1}{d_{ij}^2} \right)^2
\]

This function measures the degree of continuity inversely; hence reduced values indicate that the higher dimensional structure is well embodied by the lower dimensional solution. This function is minimized using a gradient based approach with several parameters to be stated by the user. The most significant parameters being the number of random starts to use, and the maximum number of iterations allowed. The particular values of the parameters used in the comparative paper written by Akkucuk and Carroll (2006) were to use 300 random starts and a maximum of 2,000 iterations in an initial stage and then further running the finest solution from the 300 random starts by 100,000 repetitions.

It is sensible to make a note here about the values of kappa that are conveyed in this article. In order to standardize the kappa measure it is necessary to study the condition in which the \( \delta_{ij} \) and \( d_{ij} \) are perfectly correlated. If the kappa value is multiplied by a constant that is only a function of the \( \delta_{ij} \), in particular if this constant is chosen such that it is the inverse of the kappa value when \( \delta_{ij} = d_{ij} \), then kappa will be standardized such that the lowest possible value is one. This naturally has zero effect on the minimization progression, but results in an easier understanding of the kappa function.
Isomap Details

The Isomap algorithm (Tenenbaum et al., 2000) uses a shortest path technique to generate the matrix of the “geodesic” distances between the points lying on a nonlinear manifold, rather than the straight line Euclidean distances. Prior to running the shortest path algorithm, the nearby points are connected by using either of the two available options (a) by connecting the “k nearest neighbors” or (b) by connecting “neighbors closer than a small value ε”. After this connection step, the shortest path procedure calculates the shortest path distances between all the points and then a classical metric MDS step creates the lower dimensional solution. After the run, the algorithm computes the measure called “residual variance” which is abbreviated by RV, this is equal to 1 – (the squared correlation between the geodesic distances and the distances computed from the lower dimensional embedding).

The solutions and the RV values change naturally depending on the choice of the parameter k or ε. It is also clear that connected (the full set of points in the input configuration are represented in the solution) or disconnected (a subset of the full set of points in the input configuration is represented in the output configuration) solutions may occur depending on the selection of the value of the parameters of Isomap. When the k nearest neighbors alternative is used it is not possible to culminate in disconnected solutions but when the other option is used it is likely that below a threshold value some points will not be able to connect and will be left out of the solution.

The appropriate choice of the neighborhood size parameter (ε) has been discussed in the literature and one suggestion is increasing the neighborhood size at regular intervals, starting from a low value, and observing RV. It is reported that the RV value starts decreasing slowly than at a particular point experiences a sudden jump uphill (Balasubramanian et al., 2002). It may be thus sensible to stop and choose the solution with the smallest RV right before the abrupt jump. In our experiments we also observe a similar occurrence. Also we observe that in some of those cases the graph was not fully connected and the total number of points in the solution did not equal the number of points in the input configuration.

The main reason in the sudden worsening is that as ε increases some points on opposite sides of the manifold are connected (an act known as “short-circuiting” or “cross-branching”, similar to for example connecting two opposite points on a sphere by going inside the sphere rather than going over the surface of the sphere) and this perturbs the actual geodesic distances. The solution approach we propose in this paper is taking this partially complete solution (but otherwise geometrically sound) solution and use it as an input to PARAMAP program hence removing the need to do many random starts with the computationally expensive algorithm.

Assessment of the Results

Naturally the values of kappa and RV could be used to judge the quality of the mapping that takes place by either of the algorithms or the hybrid algorithm that we are proposing. Another possible measure is called the “Rate of Agreement in Local Structure” also abbreviated as A or “agreement rate” (Akkucuk & Carroll, 2006).

The agreement rate takes a certain neighborhood size (this is not related to the neighborhood size used in Isomap and in our applications we have taken this number as 5) and computes the percentage of points that are in the neighborhood of the same points both in the input configuration and in the output configuration (order is not important). The agreement rate has a maximum theoretical value of 100% and a minimum of 0%. In this paper kappa (or κ used interchangeably), RV and A will be used to evaluate the mapping performance of various results. The ideal (or best) values are respectively 1.00, 0.00 and 100.00 (the last one being expressed generally as a percentage).

The Problem of Large Data Sets: A Landmark Based Algorithm

Computationally intensive algorithms like PARAMAP sometimes make it difficult to deal with data sets of large size. Since the iterative procedure computes squared Euclidean distances at every step, as the size of the data grows, the computation time becomes prohibitively large. This fact, combined with the need to do random starts may pose practical difficulties for researchers. A common procedure to deal with large data sets is to choose a number of landmark points “wisely” (or randomly) and then use the minimization algorithm on the configuration defined by the landmark points. After this step, there must be a procedure to map the remaining points (the “holdout” points) on the landmark points.

Akkucuk (2004) presented one such method to let PARAMAP deal with a large number of points. The proposed solution strategy applies the PARAMAP algorithm to a subset of the original points (landmark points) and then extrapolates or interpolates the rest (holdout points). This is a general approach and is not restricted to the PARAMAP model. The general problem is broken down into three sub problems: (1) Determination of the landmark points and (2) running PARAMAP on the landmark points using the requisite amount of random starts, and finally (3) mapping the holdout points into the low dimensional space defined by the landmark points by one PARAMAP run.

Determination of the Landmark Points

The first crucial task in selecting the landmark points is to choose what percentage of the points to select as landmark points. There is no obvious straightforward answer to this question.
The precision required in the final solution of the overall structure would need to be considered along with other factors like computation time. The higher the number of landmark points selected the better the solution would be. Again the lower the number of landmark points used the algorithm speed would increase.

In work related to applying MDS to visualize higher dimensional data in two dimensions, \(\sqrt{n}\) landmark points have been used by Morrison, Chalmers, and Ross (2002, 2003). This number can be treated as a lower bound on the number of landmark points to select. If computational conditions permit even more landmark points could be selected. The square root selection would also have the nice theoretical property of making the algorithm scale by \(O(n)\) instead of \(O(n^2)\).

Once the number of landmark points to be selected is determined the second step is to determine which points will be the landmark points and which points will be the holdout points. The selection can be purely random or can be based on the structure in question. For example, if there is a sphere with 5 meridians and with 24 parallels all equally spaced, half of the parallels can be eliminated. But if the structure in question cannot be readily visualized or is not very regular, an algorithmic procedure would be most suitable for the landmark point selection process. A simple approach to the landmark selection problem is one based on a method originally used for determining an initial seed for the K-means procedure (DeSarbo, Carroll, Clark, & Green, 1984). The algorithm in effect tends to reach maximum “entropy” or in other words make the landmark points as spread out as possible, so that they “span” the full space in which the object points are embedded. This algorithm, to be called the “K-means landmark selection algorithm” would be relatively simple to implement as compared to, for example, running a full K-means algorithm. The algorithm can be described as follows:

i. Let \(L\) be the number of landmark points, among \(n\) points. Choose the first two points by finding the largest \(\delta_i\) in the matrix of Euclidean distances \(\Delta\). Designate \(i\) and \(j\) to be the first two points.

ii. Find another point \(k\) such that the minimum of Euclidean distances from the other chosen points is maximized; e.g., given \(i\) and \(j\) find \(k\) such that \(\min(\delta_i, \delta_j)\) is maximized (for the case when two points are selected).

iii. Repeat this procedure until \(L\) points are selected.

**Running PARAMAP on the Landmark Points**

Once the landmark points are determined the PARAMAP algorithm is run on these landmark points using the requisite number of random starts as explained before. Actually the most important difference between the landmark selection technique proposed previously and the hybrid technique we will propose in the next section is that the previous method only selects the landmark points it does not provide an initial lower dimensional solution.

Once the best solution is selected from the random starts (best solution in the sense of minimum kappa), this solution is fed to the next step, i.e., the step involving using this lower dimensional solution containing the landmark points and mapping in the rest of the points (the holdout points).

**Mapping in the Holdout Points**

The proposed solution to this problem is rather simple to implement, and makes use of the PARAMAP algorithm itself.

In the method PARAMAP is used conditionally minimizing \(\kappa\) to map the holdout points into the landmark configuration. The condition is that the locations of the landmark points are kept constant. The procedure is implemented as follows:

i. Run PARAMAP on the landmark points to get the lower dimensional embedding.

ii. Initially locate each holdout point on the landmark point that is closest to it in the lower dimensional embedding, thus obtaining a lower dimensional configuration for both the landmark and the holdout points.

iii. Use this lower dimensional configuration as a starting configuration for a conditional PARAMAP analysis and run the algorithm to convergence. In doing so, keep the locations of the landmark points constant throughout minimization. In other words in the minimization process regard the coordinate locations of the landmark points as constants rather than variables, thus conditionally minimizing \(\kappa\) as a function of coordinates of the holdout points only.

This method has two important aspects. The first is, since the starting configuration is a rational starting configuration (i.e., the objective function would already be at a low value only requiring some fine tuning), the need for different random starts would be eliminated altogether. This would save considerable time since running PARAMAP from randomly generated configurations requires many different random starts. Second, even if the number of points is prohibitively large, so that even running the algorithm to convergence once would be very time consuming, the holdout points can be broken down into batches and then separate runs of PARAMAP (conditioned on holding the same set of \(L\) landmark points fixed) can be performed, and the results later put together. Since the landmark points are constrained to be fixed and stable, when adding together the resulting solutions, there would be no need to find rotation, translation or other transformation parameters. It would be possible even to map one point at a time, but this would definitely not be wise as far as computational time is concerned.
However, there is one disadvantage in that after the landmark points are selected, random starts still need to be performed on these landmark points. Another approach may yield a lower dimensional configuration containing the landmark points. Isomap could be such an approach providing an initial rational configuration. These observations were already made for the landmark selection method presented in Akkucuk (2004) and lead the way into development of the hybrid PARAMAP and Isomap algorithm. This algorithm will be explained in the next section.

**Combining PARAMAP and Isomap Features**

The core dissimilarity between the Isomap approach and PARAMAP is Isomap’s ability to converge in practical time. The mixture procedure takes advantage of this piece of Isomap and uses a configuration found by Isomap as an early starting configuration which will be further fine-tuned by PARAMAP.

The PARAMAP procedure will only be run once such that the additional costs of running many random starts will be removed. In some Isomap solutions the number of data points may be less than the number of data points in the input configuration (an incomplete start). On the other hand, in some solutions the number of data points in the solution will be equal to that of the input configuration (a complete start). If the start is incomplete, the so called “core” set of landmark points will define the “interior” of the ultimate total configuration we wish to fit, and then new points will be added based on their proximity to landmark points in this initial structure. As new points are added to this principal set of landmark points, the complete framework is gradually built via an evolutionary progression by which points farther and farther out in the periphery of the ultimately defined structure are incorporated into what will be this ultimate structure.

The iterative process underlying PARAMAP also will be likely to change the interrelationships among points in the initial framework, so that the final configuration may bear relatively slight relationship to this initial structure. This latter point is also another one of the major modifications between the methods explained in before (landmark based procedure) and the hybrid algorithm proposed here and by Akkucuk and Carroll (2010). The proposed algorithm will be composed of two phases, the first phase will be identified as “Prior Isomap Step” and the second and final step is identified as “Holdout Point Mapping”. The proposed mixture procedure will work as follows.

**Prior Isomap Step**

In this step the input configuration will be subject to analysis via the Isomap algorithm. It is essential here to try different values of the parameter $\varepsilon$ and compare the residual variance ($RV$) values reported by Isomap software. The details of the Isomap algorithm are given in Akkucuk and Carroll (2006) and Tenenbaum et al. (2000). We will not repeat all the details here, yet it is worth explaining the procedure used by Isomap to construct the neighborhood graph. Isomap takes as input a user specified parameter which can be either one of the options “$k$” or “$\varepsilon$”.

The first option connect to each point the $k$ nearest neighbors hence constructs the neighborhood graph.

The second option ties the points that are less than distance of $\varepsilon$ to each point and hence builds the neighborhood graph. It is evident that with the first option disconnected solutions are not possible while with the second option we might have some disconnected solutions and this actually forms the basis of our procedure.

When selecting different values of $\varepsilon$ one sensible technique is to start from zero and rise in fixed step sizes. The fixed step size should be small enough to permit disconnected solutions. In some data sets the solutions will not be disconnected due to adequate point density. The step size can also be expressed as a function of the interpoint distances. In this article we use a step size as small as 0.05. This is sensible with the data sets we have but may not work with other data sets.

When the $RV$ values for diverse step sizes are recorded at some point the $RV$ values should degenerate. This is due to an effect known as “cross branching”, i.e., the step size is too much and the nearest neighbor turns out to be a point on the other side of the manifold. The starting configuration for the PARAMAP algorithm will be selected to be the one just before the degeneration occurs.

**Holdout Point Mapping**

In this stage of the algorithm the holdout points are plotted in a single step. In this step the entire set of holdout points are first located at the points they are closest to in the input configuration. Then one PARAMAP run is taken to find the solution. This process is similar to the algorithm proposed previously and explained in great detail in a previous section, but during the PARAMAP run the positions of the landmark points are free to fluctuate as the holdout points.

**Conclusion**

We consider this mixture method to be a very beneficial data analytic device that can deal with either open surfaces with inadequate point density or with closed surfaces with a big number of points. The former case is problematic for Isomap as the resulting solution will not be whole. In this case the proposed hybrid algorithm can be used to map in the left over points. In the second case running different random starts with PARAMAP may be expensive, so the hybrid algorithm can be used and only one PARAMAP run is needed. In cases where Isomap yields a complete solution with a superior residual variance, then PARAMAP can be used to test the
optimality of the solution. PARAMAP generally results in no further enhancements in such a case since Isomap apparently produces the globally optimal solution using the PARAMAP criterion (kappa).

There are a number of further research issues that need to be explored concerning this hybrid algorithm proposal. One possible opportunity of research is exploring the use of mapping the holdout points in batch wise instead of mapping in all points at the same time. In this situation the optimal selection of the batch size is a major concern.

One more vital question for auxiliary research opportunities is the study of using the agreement rate as an ending criterion in the initial Isomap procedure rather than the RV termination criterion. One other useful enhancement to the algorithm would involve integrating the two approaches into a distinct program (perhaps coded in Matlab or R), thus creating a single suite with numerous alternatives.

Lastly, a vital problem with the mixture procedure is the incapacity to test different random starts. This problem can be partially resolved by adding relatively small random perturbations to the final solution and running PARAMAP a number of times from such differently perturbed configurations. This may have the effect of permitting discovery of a nearby globally optimal solution.

Visualization becomes ever important in the area of marketing and other social sciences where big data sets are encountered. France and Ghose (2016) have introduced a method for identifying, analyzing, and visualizing submarkets in product categories using visualization. They state “In the era of big data and with the increasing availability of large-scale consumer purchase data, there is a need for techniques that can interpret these data and use them to help make managerial decisions.” Another further research area is the development of the fit measures used in dimension reduction techniques. These measures are important in terms of comparing performance of different algorithms. The agreement rate (Akkucuk & Carroll, 2006; France & Carroll, 2007; Chen & Buja, 2009) is one of these measures and can be used to compare a low dimensional solution to a higher dimensional input data in terms of the proximity of the objects to other objects. Further studies on measures of fit, agreement or loss of quality (Gracia et al., 2014) will lead to better evaluation of visualization algorithms and will contribute to the improvement of visualization methods for data mining tasks.

Acknowledgement

Initial version of this paper has been published in IJDSCA Draft issue in 2016. Article has been re-edited and published in IJDSCA.

References

Al-Kassab, J., Ouertani, Z. M., Schiuma, G., & Neely, A. (2014). Information visualization to support management decisions, *International Journal of Information Technology & Decision Making*, 13(2), 407-428. https://doi.org/10.1142/S0219622014500497

Akkucuk, U. (2004). Nonlinear Mapping: Approaches Based on Optimizing an Index of Continuity and Applying Classical Metric MDS to Revised Distances. Doctoral Dissertation, Rutgers University, Newark, NJ. http://search.proquest.com/docview/305116738?accountid=9645

Akkucuk, U., (2011). A Study on the Competitive Positions of Countries Using Cluster Analysis and Multidimensional Scaling, *European Journal of Economics Finance and Administrative Sciences*, 37, 17-26. https://www.researchgate.net/publication/286042216_A_Study_on_the_Competitive_Positions_of_Countries_Using_Cluster_Analysis_and_Multidimensional_Scaling?ev=prf_pub

Akkucuk, U. (2014). Application of Statistical Visualization Tools on Global Competitiveness Data. In Hasan Dincer and Umit Hacoglu (Eds.), *Global Strategies in Banking and Finance* (14-27). IGI-Global: Hershey. https://doi.org/10.4018/978-1-4666-4635-3.ch002

Akkucuk, U. & Carroll J. D. (2006). PARAMAP vs. Isomap: A Comparison of Two Nonlinear Mapping Algorithms. *Journal of Classification*, 23, 221-254. https://doi.org/10.1007/s00357-006-0014-2

Akkucuk, U. & Carroll, J. D. (2010). Nonlinear Mapping Using a Hybrid of PARAMAP and Isomap Approaches. In Hermann Locarek-Junge and Claus Weihs (Eds.), *Classification as a Tool for Research* (371-380). Springer: Berlin-Heidelberg-New York. https://doi.org/10.1007/978-3-642-10745-0_40

Akkucuk, U., Carroll, J. D., & France, S. (2013). Visualizing Data in Social and Behavioral Sciences: An Application of PARAMAP on Judicial Statistics. In Berthold Lausen, Dirk Van den Poel and Alfred Ultsch (Eds.), *Algorithms from and for Nature and Life* (147-154). Springer: Berlin-Heidelberg-New York. https://doi.org/10.1007/978-3-319-00035-0_14

Akkucuk, U. & Artemel, M. N. (2016). Patent Data Visualization: A Regional Study, *International Journal of Research in Business and Social Science*, 5(3), 66-79. http://dx.doi.org/10.20525/ijrbs.v5i3.358

Akkucuk, U., & Kucukkancabas, S. (2007). Analyzing the Perceptions of Turkish Universities Using Multidimensional Scaling (MDS) Analysis, Bogazici Journal, 21, 125-141.

Balasubramanian, M., Schwartz, E. L., Tenenbaum, J. B., De Silva, V., & Langford, J. C. (2002), The Isomap Algorithm and Topological Stability, *Science*, 295, 7a. https://doi.org/10.1126/science.295.5552.7a

Desarbo, W. S., Carroll, J. D., Clark, L. A., & Green, P. E. (1984). Synthesized Clustering: A method for amalgamating alternative clustering bases with differential weighting of variables. *Psychometrika*, 49, 57-78. https://doi.org/10.1017/0BF02294206
Dincer, H. and Hacioglu, U. (2013). Global Strategies in Banking and Finance, Publisher: IGI-Global, USA: https://doi.org/0.4018/978-1-4666-4635-3
France, S. L. & Carroll, J. D. (2011). Two-way multidimensional scaling: A review, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 41(5), 644-661. https://doi.org/10.1109/TSMCC.2010.2078502
France, S. L. & Ghose, S. (2016). An Analysis and Visualization Methodology for Identifying and Testing Market Structure, Marketing Science, 35(1), 182-197. https://doi.org/10.1287/mksc.2015.0958
Gracia, A., González, S., Robles, V., & Menasalvas, E. (2014). A methodology to compare Dimensionality Reduction algorithms in terms of loss of quality, Information Sciences, 270, 1 – 27. https://doi.org/10.1016/j.ins.2014.02.068
Morrison, A., Ross, G., & Chalmers, M. (2002). A hybrid layout algorithm for sub-quadratic multidimensional scaling. In P. C. Wong & K. Andrews (Eds.), IEEE Symposium of Information Visualization. (pp. 152-158). Silver Spring, MD: IEEE Computer Society. https://doi.org/10.1109/INFVIS.2002.1173161
Morrison, A., Ross, G., & Chalmers, M. (2003). Fast multidimensional scaling through sampling, springs and interpolation. Journal of Information Visualization, 2, 68-77. https://doi.org/10.1057/palgrave.ivs.9500040
Shepard, R. N. & Carroll, J. D. (1966). Parametric representation of nonlinear data structures. In P. R. Krishnaiah (Ed.). Multivariate Analysis, (pp. 561-592). New York, NY: Academic Press.
Tenenbaum, J. B., De Silva, V., & Langford, J. C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction, Science, 290, 2319-2323. https://doi.org/10.1126/science.290.5500.2319

Publisher's Note: SSBFNET stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee SSBFNET, Istanbul, Turkey. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).