

The Choice of Initial Configurations in Multidimensional Scaling: Local Minima, Fit, and Interpretability

Ingwer Borg
WWU Münster

Patrick Mair
Harvard University

Abstract

Multidimensional scaling (MDS) algorithms can easily end up in local minima, depending on the starting configuration. This is particularly true for 2-dimensional ordinal MDS. A simulation study shows that there can be many local minima that all have an excellent model fit (i.e., small Stress) even if they do not recover a known latent configuration very well, and even if they differ substantially among each other. MDS programs give the user only one supposedly Stress-optimal solution. We here present a procedure for analyzing all MDS solutions resulting from using a variety of different starting configurations. The solutions are compared in terms of fit and configurational similarity. This allows the MDS user to identify different types of solutions with acceptable Stress, if they exist, and then pick the one that is best interpretable.

Keywords: MDS, initial configuration, local minima, procrustes, SMACOF, R.

1. Introduction

Multidimensional scaling (MDS) is a statistical method that optimally maps proximity data on pairs of objects (i.e., data expressing the similarity or the dissimilarity of pairs of objects) into distances among points in a multidimensional space. MDS is used for exploring or testing the structure of proximity data. There are many variants of MDS (see [Borg and Groenen 2005](#); [Cox and Cox 2000](#)). In applied research, two-dimensional ordinal MDS is probably the most popular model. Here, the proximity data—converted first to dissimilarity indices δ_{ij} in case the proximities are similarity measures—are optimally mapped into Euclidean distances $d_{ij}(\mathbf{X})$ among points of a two-dimensional Euclidean space (with the coordinate matrix \mathbf{X}). The order of the distances corresponds to the order of the data, and ties in the data can be broken in the distances (“primary approach to ties”).

The fit of an MDS model to the data is measured by the raw Stress coefficient

$$\text{Stress} = \sum_{i < j} e_{ij}^2 = \sum_{i < j} w_{ij} (f(\delta_{ij}) - d_{ij}(\mathbf{X}))^2, \quad (1)$$

where w_{ij} are non-negative fixed weights set by the user to weight the importance of error

(e.g., to handle missing data by setting $w_{ij} = 0$ if δ_{ij} is missing and $w_{ij} = 1$ otherwise). The fitted distances are defined by

$$d_{ij}(\mathbf{X}) = \left(\sum_{a=1}^m |x_{ia} - x_{ja}|^p \right)^{1/p}, \quad (2)$$

with $p = 2$ in the Euclidean case. The $f(\delta_{ij})$ are *disparities*, that is, dissimilarities optimally re-scaled (within the bounds set by the scale level assigned to the data) so that they approximate the distances as closely as possible. Expressed more technically, disparities are computed by a regression of the dissimilarities onto the distances so that $f(\delta_{ij}) = \hat{d}_{ij}$.

Since Equation (1) is minimized over both \mathbf{X} and \hat{d}_{ij} , an obvious but trivial solution is choosing $\mathbf{X} = \mathbf{0}$ and all $\hat{d}_{ij} = 0$. To avoid this solution, Equation (1) needs to be normalized which can be achieved by dividing by the sum of the squared distances. Doing so and taking the square root gives the usual *Stress.1* loss function of MDS:

$$\text{Stress.1} = \sqrt{\sum_{i<j} w_{ij} (\hat{d}_{ij} - d_{ij}(\mathbf{X}))^2 / \sum_{i<j} w_{ij} d_{ij}^2(\mathbf{X})}. \quad (3)$$

This normalization also has the advantage that the Stress value does not depend on the magnitude of configuration \mathbf{X} .

An optimal MDS solution is found by using iterative optimization algorithms. They all start with some initial configuration (IC), and then repeatedly move its points in space in small steps until the Stress has converged to a minimum. Modern algorithms guarantee convergence, but the response surface is bumpy and the optimization can easily end up in a local and not the global minimum (Groenen 1993). The choice of the initial configuration can be crucial, because the various local minima sometimes differ radically even if they represent the data almost equally well in terms of overall Stress. For most data, many local optima exist in most MDS models, in particular when using ordinal MDS and when the dimensionality of the MDS solution is low. Suboptimal local minima are particularly likely to occur in case of one-dimensional MDS, where standard MDS programs almost never find the global minimum (Mair and De Leeuw 2015). Conversely, the greater the dimensionality of the MDS space, the smaller the risk for suboptimal local minima (Borg, Groenen, and Mair 2013, p. 61ff.).

The various local minima may differ substantially even if they represent the data almost equally well in terms of overall Stress. Confirmatory MDS (Bentler and Weeks 1978; De Leeuw and Heiser 1980) is designed to check this. It puts additional theory-based restrictions onto the MDS solutions. For instance, allowing the user to specify the MDS dimensions (i.e., the columns of \mathbf{X}) with values that must remain fixed up to linear transformations, or requiring that all points of the MDS solution are located on a perfect circle. The MDS algorithm may still find configurations that satisfy these particular models with Stress values that are hardly worse than those produced by substantively blind exploratory MDS. Borg and Lingoes (1980) provide striking examples of cases where minimal-Stress exploratory MDS representations and minimal-Stress theory-compatible MDS representations of the same data are very different, but where both have acceptably low Stress values.

When running MDS in a confirmatory way, the researcher can almost always specify an IC that he or she derives from content theory or from structural relations established in previous empirical research. For example, in psychological research on personal values (such as power, hedonism, and self-direction) where hundreds of studies have led to robust structural expectations, one can predict that the MDS configurations representing correlations among items on the psychological importance of different personal values form a circle. The various personal values are represented as points on this circle in a particular order, and with certain values positioned in opposition to each other (Schwartz and Bilsky 1987; Schwartz, Cieciuch, Vecchione, Davidov, Fischer, Beierlein, Ramos, Verkasalo, Lönnquist, Demirutku, Dirilen-Gumus, and Konty 2014; Borg, Bardi, and Schwartz 2016). Hence, it is easy to formulate

a theoretically sound initial configuration when doing yet another MDS study on personal values.

In exploratory MDS, an initial configuration has to be set up by statistical reasoning. Most MDS programs offer a number of choices for such rational initial configurations. The one that is usually recommended as the best initial configuration—that is, as the one that is most likely to lead to a configuration with minimal Stress—is the “Torgerson” configuration obtained by “classical” MDS (Torgerson 1958; Davison 1978; Borg and Groenen 2005). It is computed by assuming that the dissimilarities are distances, then converting these data to scalar products by squaring and double-centering them, and finally taking the first m eigenvectors of the resulting matrix as the m -dimensional initial configuration.

A second (additional) recommended choice is to run MDS with many different random initial configurations and then pick the configuration with the smallest Stress as the optimal MDS solution. With today’s computing power this is a viable alternative, since it only takes seconds to run such an approach, and one can always compare the results of the random and the Torgerson method to check if they lead to the same results.

Unfortunately, applied researchers find themselves in a dilemma when using MDS. They can choose a particular type of initial configuration such as Torgerson. An MDS program will then deliver the supposedly optimal solution based on this choice. Or they can choose the random approach and then the program presents what appears to be the formally best solution based on the analyses of many different initial configurations. In either case, no information is provided on other local-minima solutions even though these solutions may have only marginally higher Stress values but may be theoretically much more meaningful and better interpretable. Yet, even if the solution given by the MDS program is indeed the global minimum solution, an MDS user may simply be interested to see what other local-minima solutions exist, what their Stress values are, and how similar they are. In the following, we study this issue with an artificial data set and describe a systematic approach that can be used by the applied MDS user to answer these questions for his or her data and the particular choice of MDS model.

2. Method

We now describe a procedure for generating local minima solutions in MDS beginning with different initial configurations, and for comparing the configurational similarity of these solutions so that the user can identify those solutions he or she wants to check for their interpretability. The procedure is then illustrated using a simulation study where the true configuration is known. Finally, we look at some applications in real empirical research.

2.1. The procedure

To avoid overlooking local minima solutions in MDS that have an acceptably good fit and that are also substantively interpretable, one must generate many such local minima (if they exist) and then systematically compare them rather than reporting only one Stress-optimal solution. We suggest to achieve this goal in the following way:

1. Run an MDS analysis with a set of different initial configurations (e.g., using many random configurations).
2. Save all resulting MDS solutions and their fit indices (Stress, p -values resulting from permutation tests, etc.).
3. Use generalized Procrustean fitting to eliminate all meaningless differences (i.e., differences not driven by the data) among the MDS solutions.
4. Compute the similarity of each pair of MDS configurations.

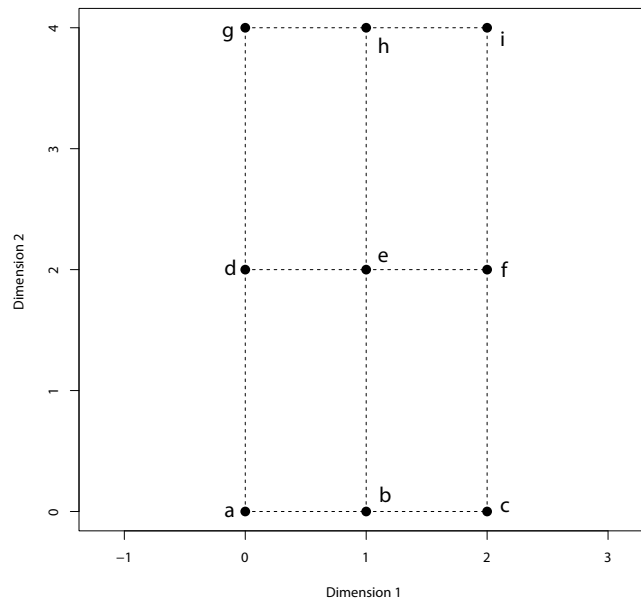


Figure 1: Configuration to be recovered by MDS (“true” configuration)

5. Analyze the similarity structure of the MDS configurations with two-dimensional MDS (to visualize the similarity structure) and cluster analysis (to identify types of MDS configurations).
6. For each type of MDS configuration with a reasonable Stress, plot one prototypical MDS solution and check its interpretability.
7. Pick the MDS solution that is both acceptable in terms of Stress and best interpretable as your MDS solution.

These steps can be easily programmed so that the user has to choose only the particular MDS model he/she wants to fit. Plots and statistics can be produced that allow the user to pick those MDS solutions that he/she wants to study further for their interpretability. A corresponding R (R Core Team 2016) code chunk is given in the supplementary materials. We use the **smacof** package (De Leeuw and Mair 2009) to fit the MDS models.

2.2. Simulation study

We illustrate our procedure by using an artificial case where the true MDS configuration is known. We begin by defining an MDS configuration \mathbf{X} which consists of nine points forming a rectangular grid (see Figure 1). The distances among the nine points are computed. Then, to introduce a somewhat non-linear mapping, their square roots are taken as dissimilarity data for MDS. Thus, there exists a true underlying configuration whose distances are monotonically (but not linearly) related to the dissimilarities.

We then ask if MDS succeeds to recover \mathbf{X} given the above dissimilarities. \mathbf{X} , of course, is the global minimum MDS solution with a Stress of zero. We use ordinal MDS because of the non-linear relation of the dissimilarities to the MDS distances in our case. We choose this setup also because ordinal MDS has been a very popular MDS model in applied research. In addition, most MDS programs run ordinal MDS with the additional default specification that ties in the data need not lead to the same distances in the MDS solution (the so called primary approach to ties).

Two types of initial configurations are used in the following, always employing the function `mds()` of the **smacof** package to compute MDS solutions:

1) The first type of IC is the Torgerson configuration. We then add successively more error (randomly sampled from a normal distribution with mean = 0 and $sd = 1, 2, \dots$) to the point coordinates of this IC and repeat the MDS analysis for each case. With $sd = 0$ we have the case that is the default IC in most modern MDS programs today, or the IC that is most often recommended as the best single IC based on extensive simulation studies (Borg and Groenen 2005). When adding more noise to a Torgerson IC, we test the case where other solutions with similar Stress exist but would not be found because the algorithm gets stuck in the neighborhood of the Torgerson IC.

2) To safeguard against sub-optimal local minimum solutions, we recommend repeating an MDS analysis with random initial configurations and then pick the resulting MDS solution with the smallest Stress as the best solution. We therefore also run random initial configurations (with coordinate values randomly sampled from a standard uniform distribution), but store each solution and not just the best one. We then compare the various MDS solutions so obtained by inspecting their Stress values and their configurations.

To compare many dozens configurations, a systematic approach is used: We first eliminate all meaningless differences of the various solutions (due to rotations, reflections, dilations, and translations) by Procrustean methods, then measure the overall similarity of each pair of configurations by computing the product-moment correlation of the coordinates of corresponding points, and finally use MDS and cluster analysis to visualize the similarity structure of these correlations to detect possible classes of acceptable MDS solutions.

Specifying the function f in Equation (1) differently, we use the same procedure to study the effect of different ICs for other MDS models, namely ordinal MDS with the so-called secondary approach to ties where tied dissimilarities must be mapped into tied distances (“keep ties”); interval MDS, where the dissimilarities can be linearly transformed; and ratio MDS where the dissimilarities are fixed up to a multiplicative constant.

2.3. Real data applications

Simulations can be illuminating to illustrate a method but they may also be too contrived for the applied researcher, showing applications and solutions for cases that almost never become relevant in empirical research. We therefore use our procedure on some real data sets that have been used before in the applied MDS literature. One such data set is a study by Wish (1971) who asked students about the subjective similarity of 12 different nations. These data are one of the oldest cases of using MDS in attitude research. They were analyzed, for example, by Kruskal and Wish (1978) in their classic introductory textbook on MDS by using ordinal MDS with a Torgerson initial configuration.

A second application uses ratings of 327 psychology students on the importance of personal values as guiding principles in their life (Borg *et al.* 2016). What is scaled here are the inter-correlations of indexes for the so called 10 basic values of the Schwartz theory on values (Schwartz *et al.* 2014). The theory predicts a circular scale, with the points ordered on the circle. This study replicates what was done in literally hundreds of related studies in research on personal values, where MDS has been the cardinal method for testing and refining what is called the Theory of Universals in Values (TUV). Almost always ordinal MDS has been used, but systematic studies of using different initial configurations are not reported.

Finally, a classic MDS application is on data on the subjective similarity of different colors by Ekman (1954). These data were used in the first papers on ordinal MDS by Shepard (1962) and Kruskal (1964) where it was shown that the structure of the observed similarities is an almost perfect color circle with very low Stress.

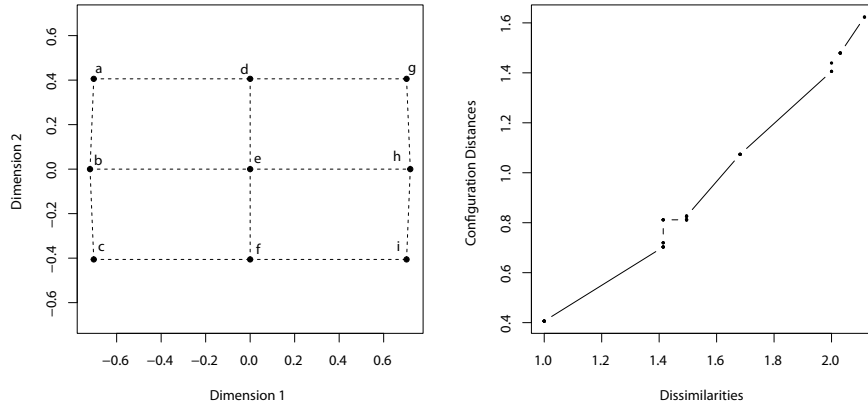


Figure 2: Zero-Stress MDS solution for dissimilarities based on the grid in Figure 1, using Torgerson initial configuration (left panel); Shepard diagram (right panel).

3. Results

3.1. Results: simulation study

A standard ordinal MDS analysis of the dissimilarity data derived from Figure 1 with a Torgerson IC leads to an almost perfect recovery of the underlying configuration, with zero Stress (Figure 2, left panel). The Shepard diagram (right panel of Figure 2) for this configuration shows that the square-root relation of dissimilarities to distances is also closely recovered.

Adding error to the Torgerson IC sometimes leads to MDS solutions with zero Stress but sometimes to radically different solutions that also have zero Stress. Figure 3 gives an example where the points do not form the rectangular grid but are collapsed in one direction to almost a single straight line. The Shepard diagram in the right-hand panel of Figure 3 shows, for example, that **smacof** maps all dissimilarities in the interval $[1.0, 1.4]$ into line segments with essentially zero length.

An even more pronounced step function comes with the solution shown in Figure 4. In this case, MDS maps the dissimilarities into only two types of distances, large ones and small ones. The configuration, therefore, exhibits the shape of an equilateral triangle. Its lines represents

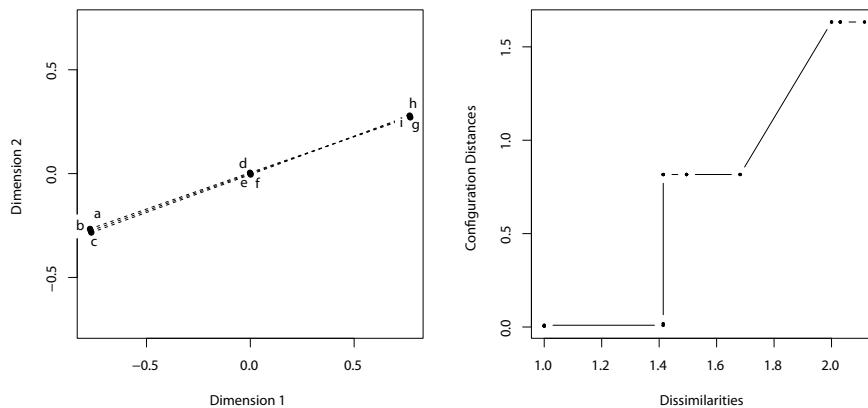


Figure 3: Zero-Stress MDS solution for dissimilarities based on the grid in Figure 1, using Torgerson plus small error initial configuration (left panel); Shepard diagram (right panel).

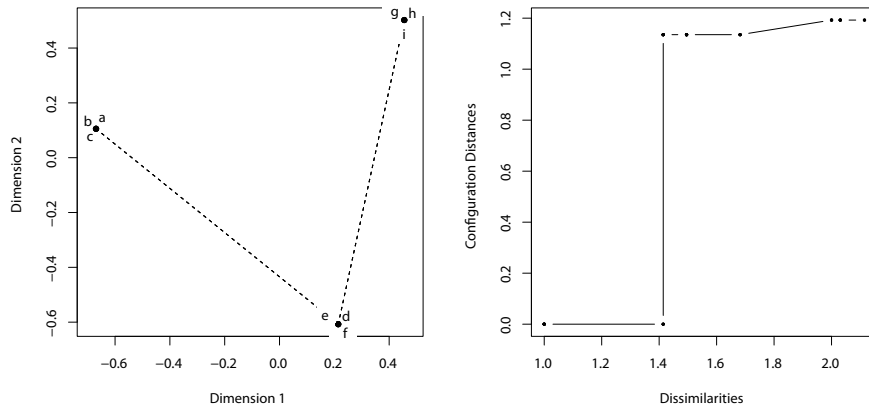


Figure 4: Zero-Stress MDS solution for dissimilarities based on the grid in Figure 1, using Torgerson plus larger error initial configuration (left panel); Shepard diagram (right panel).

the vertical grid lines of the design configuration, collapsed and wired around the triangle’s center. Its corner points each summarize the three points of one horizontal grid line of the design configuration.

In a second simulation setting the procedure from Section 2.1 is illustrated on the artificial grid data. 100 different random ICs (drawn from a $U(0, 1)$ distribution) are generated and for each of them an ordinal MDS is fitted. Analyzing the similarities of the resulting MDS solutions shows how bumpy the response surface of ordinal MDS is for this set of data. The structure of the 100 MDS solutions is visualized in Figure 5. This plot is a two-dimensional interval MDS representation of the similarity structure of the 100 ordinal MDS solutions where the

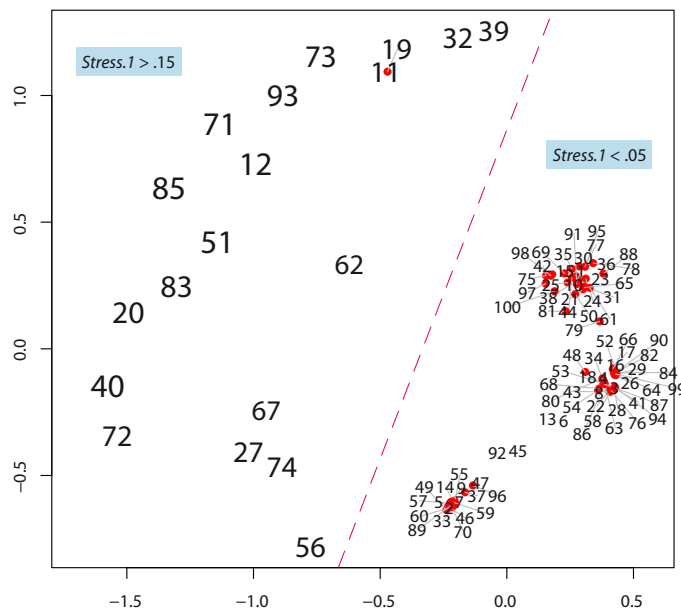


Figure 5: MDS configuration of 100 MDS solutions based on random initial configurations; plotted with **wordcloud** to unclutter point labels; label size corresponds to Stress of respective MDS solution; dashed line partitions plane into high- and low-Stress solutions, resp.

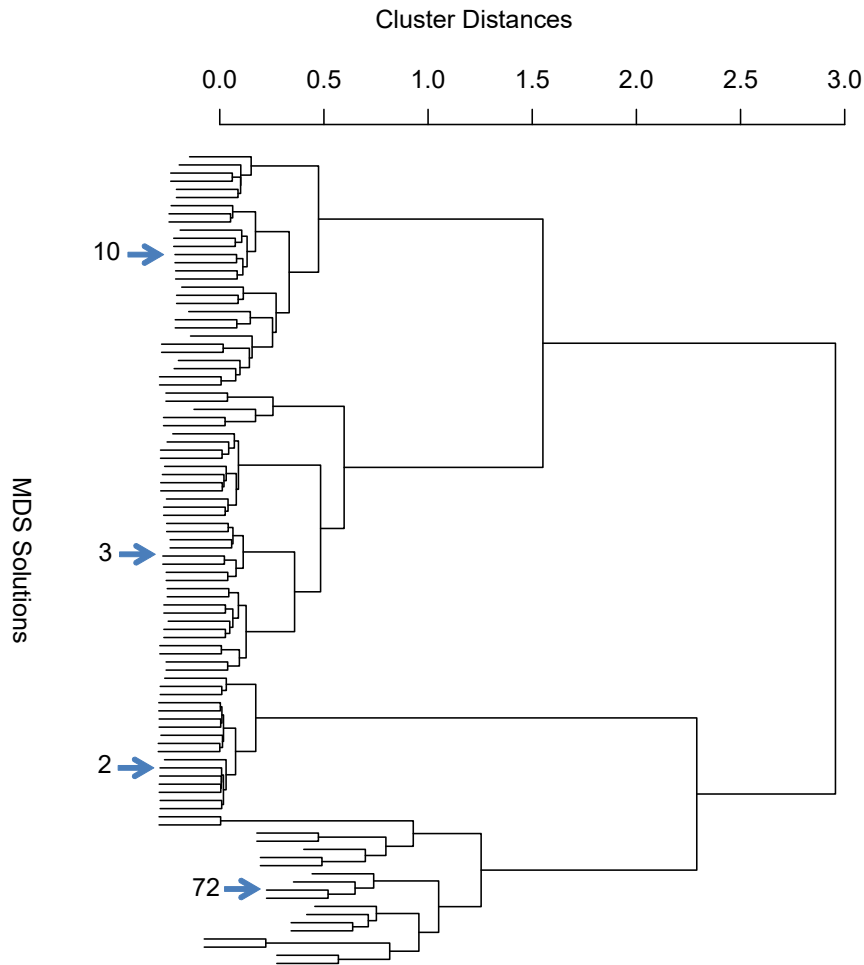


Figure 6: Dendrogram (hierarchical cluster analysis; Ward criterion) of similarity indexes of the 100 MDS solutions based on random initial configurations; arrows point to four prototypes of the major clusters.

similarity of two configurations is measured by intercorrelating (using the Pearson correlation coefficient) for each pair of solutions the coordinates of corresponding points after Procrustean fitting (cf. Borg and Leutner 1985). To unclutter the plot we used the R package **wordcloud** (Fellows 2014). Figure 5 represents each point by its label, except in cases where the point labels would be overlapping. In that case, it plots the points as red dots and connects the point label to the point with a straight line so that the various labels do not overlap (see the point clouds on the right-hand side of the plot).

The figure shows three rather dense clusters of low Stress solutions on the right-hand side, and various widely scattered solutions with high Stress on the left side of the plot (large labels). The plot can be partitioned by a straight line that separates all MDS solutions with poor Stress (left-hand side) from those with acceptably low Stress (right-hand side).

One can also use cluster analysis to study the similarity of the 100 MDS solutions: Figure 6 shows that a hierarchical cluster analysis identifies four major clusters. For each cluster, one configuration is marked by an arrow in the dendrogram as a prototype (the configurations #10, #3, #2, and #72). Figure 7 exhibits these configurations. Only one of them, #72 has a poor fit ($Stress.1 = .236$). The other three configurations correspond closely to the configurations in Figure 2 (“grid”, #10), Figure 3 (“collapsed horizontals”, #3), and Figure 4 (“triangle”, #2), respectively. They all have Stress values of almost zero. Thus, formally,

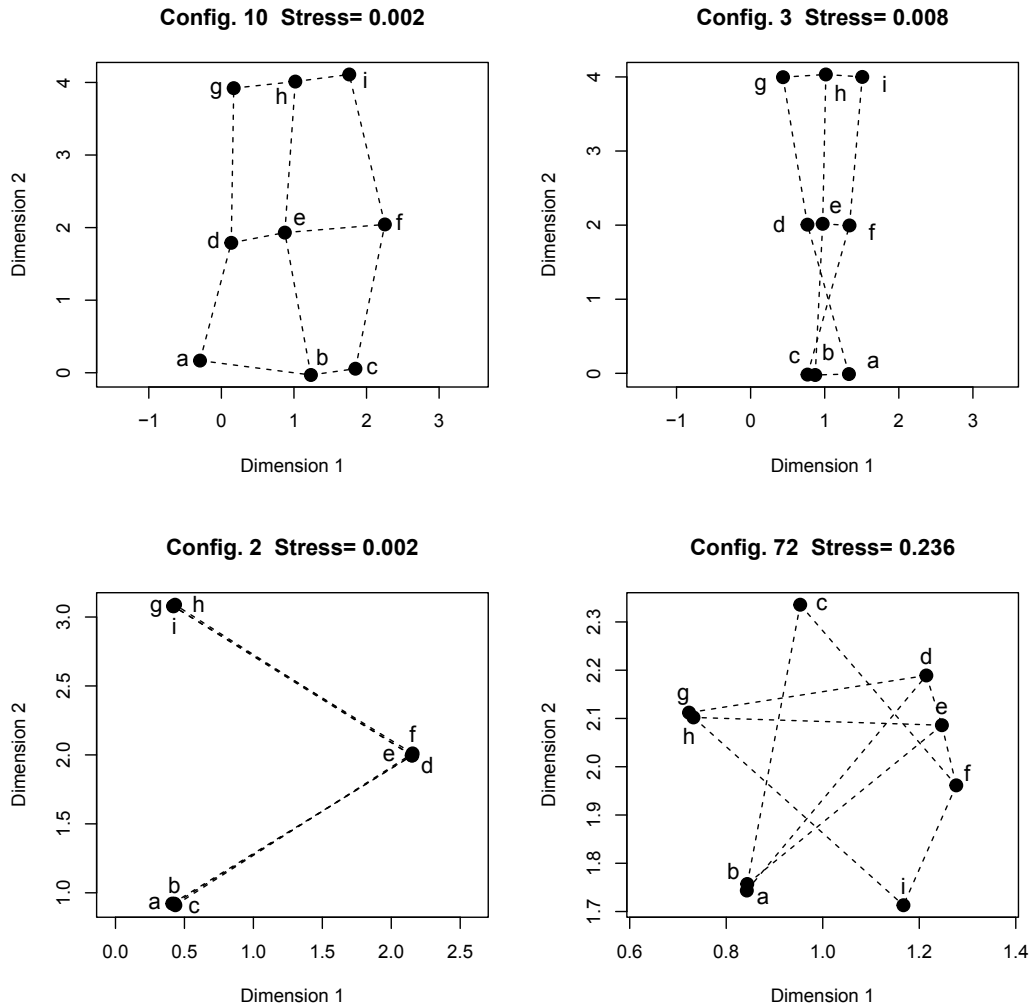


Figure 7: Configurations that represent four different clusters of the dendrogram in Figure 6.

they are all excellent representations of the data in the sense of the ordinal primary approach to ties MDS model, but neither the “collapsed horizontals” nor the “triangle” solutions recover the underlying configuration of Figure 1 very well. Rather, they are examples of degenerate solutions.

Ordinal MDS with the secondary approach to ties (“keep ties”) recovers the latent configuration of Figure 1 (and the slightly non-linear relation of dissimilarities and distances) perfectly when using the Torgerson IC or the best random IC. However, using other ICs, one also obtains many undesirable solutions. Some examples are shown in Figure 8. The configurations #5, #11, and #7 swap the points on one or two of the horizontal grid lines of the design configuration. Their Stress values are not zero, but they are all significant by the norms of Spence and Ogilvie (1973). We also find a probability of essentially zero in all cases using the permutation test provided by **smacof**. Hence, if the true configuration is not known, one would likely accept any one of these solutions as “the” MDS solution if there are theoretical reasons that speak for such a choice.

Interval MDS leads to MDS solutions that are either unacceptable local minima with high Stress, or to configurations that closely correspond to the configurations #2, #5, and #7 in Figure 8. They have Stress values of .046, .113, and .113, respectively. All Stress values indicate a highly significant model fit according to **smacof**’s permutation test.

Finally, ratio MDS leads to solutions that are quite similar to those of interval MDS. Of course, the Stress values are higher, even in case of “grid” solutions, because the regression

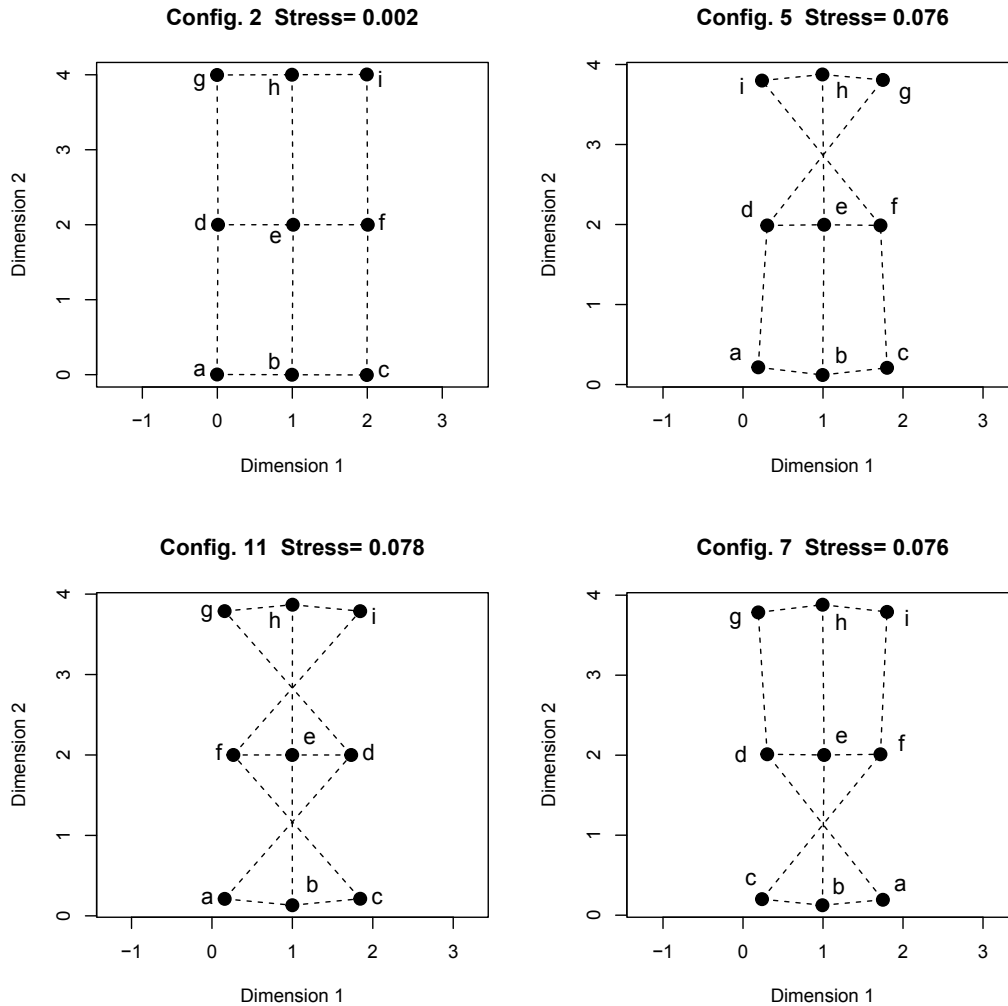


Figure 8: Four local minima solutions using ordinal MDS with the secondary approach to ties.

line in the Shepard diagram must be linear and must run through the origin in ratio MDS.

When adding some jitter to the design configuration, we find similar results in all cases, except when using ordinal MDS with the primary approach to ties. The noise that is added to the design configuration makes all ties in the dissimilarities go away. The dissimilarities are, therefore, all different and this makes it impossible for MDS to generate gross step functions as in Figure 4, for example. Hence, solutions in the form of a triangle are not observed anymore. Rather, the solutions are either grid-like, or they show some swapping of points on the lower and/or on the upper horizontal grid line (as in Figure 7).

3.2. Results: real data applications

We now turn to three classic real data sets. For the data on the subjective similarity of different countries reported by Wish (1971), ordinal MDS starting with random configurations leads to different solutions. Many of them have unacceptably high Stress, but there are two types of solutions with the same minimal Stress of .185. The fit of these solutions is also significant ($p = .04$) using **smacof**'s permutation test. Figure 9 shows these solutions next to each other. They are rather similar but differ in two important details: In configuration #1, the positions of Japan and Israel are swapped in comparison to where they are in configuration #13; moreover, in configuration #1 India is positioned more in the center of the configuration. The crossed dashed lines in the plots represent two external scales that were optimally fitted

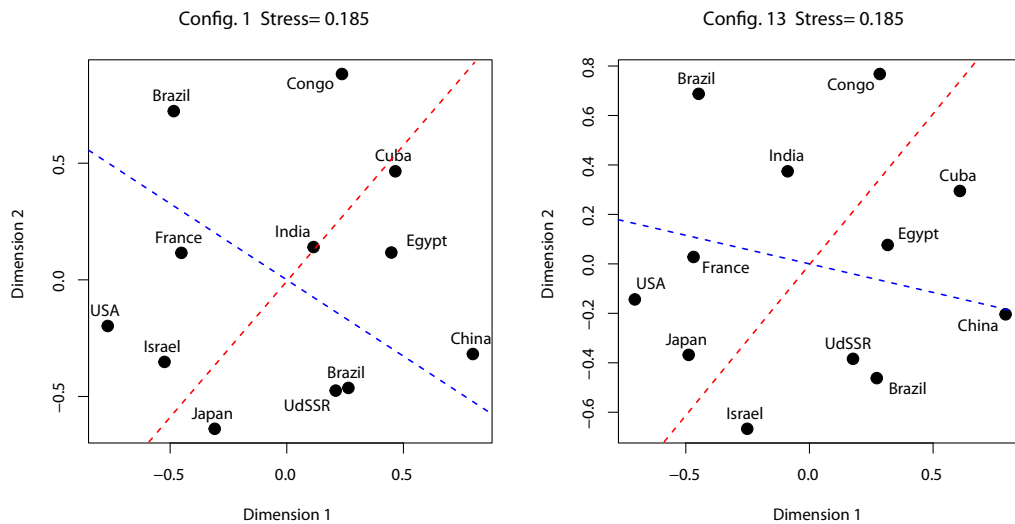


Figure 9: Two MDS solutions for ratings on the subjective similarity of different nations.

into these plots. The external scales show the economic development and the number of inhabitants, respectively, of the 12 countries in 1971 (Borg *et al.* 2013). In case of economic development (red lines), the external scales and the fitted scales correlate with $r = .936$ and $.925$ in the plots. In case of the number of inhabitants, the correlation is $r = .464$ in configuration #1 but only $r = .303$ in configuration #13. Hence, configuration #1 is the more meaningful MDS solution if one wants to follow Wish (1971) and Wish, Deutsch, and Biener (1972) in interpreting the configuration in terms of these dimensions. However, this solution may not be the one that is reported by the MDS program as the final solution.

A second application uses ratings of students on the importance of personal values as guiding principles in their life (Borg *et al.* 2016). What is scaled here are the inter-correlations of indexes for ten basic values. These values are predicted to form a circular scale, with the points ordered as PO(wer), AC(hievement), HE(donism), ST(imulation), S(elf-)D(irection), UN(iversalism), BE(nevolence), TR(adition), CO(nformity), SE(curity), and back to PO(wer).

Ordinal MDS (using `mds()` with its default settings) leads to the solution shown in the lower left-hand panel of Figure 10 with $Stress.1 = .128$. This solution obviously closely corresponds to the predicted value circle even though the circle is somewhat dented, with AC(hievement) moved towards the circle's center. Random initial configurations identify three more local minima. They all have almost the same Stress values. They are also all quite similar, but a closer inspection shows that configuration #2 exhibits a theory-incompatible swapping of CO(nformity) and TR(adition), while configuration #5 moves AC(hievement) somewhat towards the center of the circle (see Figure 10). Configuration #3 on the other hand supports the Schwartz theory perfectly.

As a third example, we study the color similarity data collected by Ekman (1954). For these data, MDS finds a circular configuration with points ordered in terms of the physical wavelengths of the colors that they represent. With both Torgerson or with random initial configurations, the usual ordinal MDS with the primary approach to ties almost always leads to the same solution, the color circle. Configurations that do not exhibit this circle are extremely rare when testing many different initial configurations. Moreover, they all have much higher Stress. This is useful information for the substantive researcher because it shows that these data have an essentially unique MDS representation.

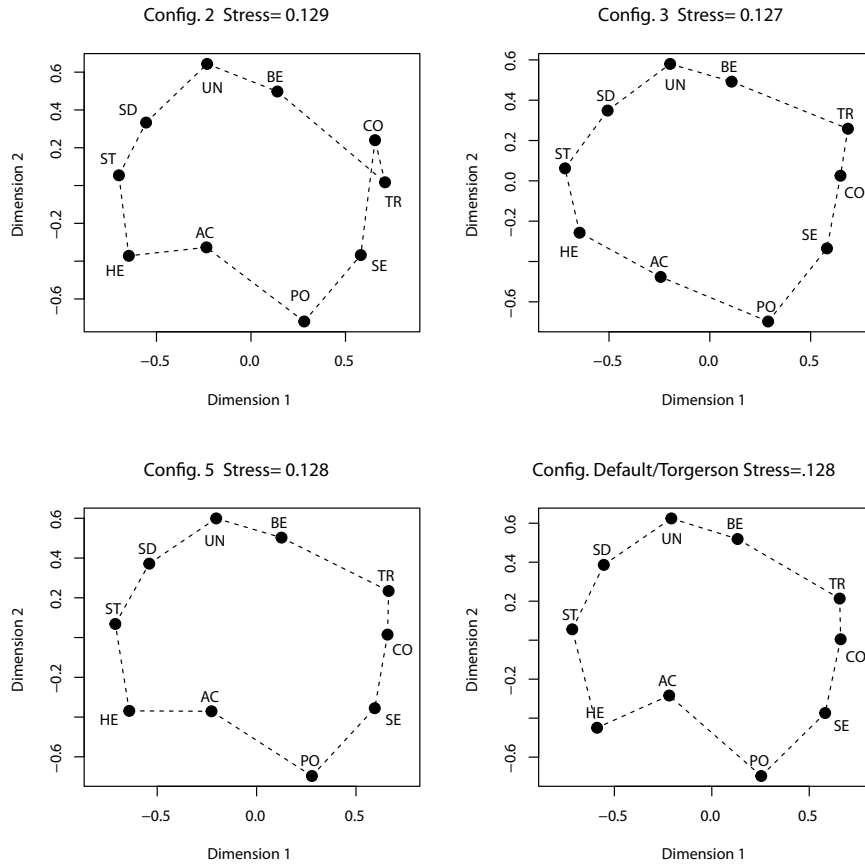


Figure 10: Local minima MDS solutions for intercorrelations of indexes on the importance of personal values.

4. Discussion

The above analyses demonstrate that it can be risky to use MDS without thoroughly studying the effects of different initial starting configurations. Not just in contrived simulations, but also with real data MDS can have many local minima with almost the same Stress but with different configurations, allowing different interpretations. The Torgerson IC proves to be a good rational IC, but (when adding error to the artificial data derived from the design grid in Figure 1), even it does not guarantee to always succeed recovering a known underlying configuration. When using stronger MDS models, particular types of solutions that result from non-smooth step functions (as in Figure 4) cannot occur anymore, but radically different solutions that all have acceptably low Stress values still exist, in particular if the data contain ties and the usual ordinal MDS model with the primary approach to ties is employed.

If MDS is used in an exploratory way to visualize the structure of proximity data and make them accessible to the researcher's eye, there is really nothing to recover. In this case, one should pick the local minimum solution that is best interpretable, ideally even suggesting a content-based law of formation. Of course, this solution should also have an acceptably low Stress value. In the applications discussed above, it turned out that the best-interpretable solutions always had a Stress value that was not worse than the Stress value of other solutions. However, this may not always be true in applied research, and so we recommend to at least take a look at other local minima solutions before accepting the solution offered by the MDS program.

This also holds if one does not use ordinal MDS but stronger MDS models. To make testing the effects of different initial configurations easy in practice, the supplementary materials provide corresponding R code. With this code, users can identify the configurations they want to look

at in more detail. The various plots (together with the fit statistics) are the basis for deciding what to pick as the best MDS representation for the given data. If computing time becomes an issue as in large scale MDS settings, R's facilities for parallelizing the random IC fits can be used (e.g. using the **parallel** package). Since these MDS fits are independent from each other, the job can be distributed easily on machines with multiple cores.

In general, no formal decision rule seems possible that tells the user what solutions satisfy the criterion of having a Stress that is "still acceptable". This decision always requires to consider a set of statistical and content criteria (see Mair, Borg, and Rusch 2016) such as the overall Stress; the composition of the Stress (Stress per point); the assumed error level of the data; the mapping requirements of the chosen MDS model; and the statistical significance, the robustness, the replicability, and the theoretical interpretability of the solution.

References

- Bentler PM, Weeks DG (1978). "Restricted Multidimensional Scaling Models." *Journal of Mathematical Psychology*, **17**, 138–151.
- Borg I, Bardi A, Schwartz SH (2016). "Does the Value Circle Exist within Persons or Only across Persons?" *Journal of Personality*. Forthcoming.
- Borg I, Groenen PJJ (2005). *Modern Multidimensional Scaling*. 2nd edition. Springer, New York.
- Borg I, Groenen PJJ, Mair P (2013). *Applied Multidimensional Scaling*. Springer, New York.
- Borg I, Leutner D (1985). "Measuring the Similarity of MDS Configurations." *Multivariate Behavioral Research*, **20**, 325–334.
- Borg I, Lingoes JC (1980). "A Model and Algorithm for Multidimensional Scaling with External Constraints on the Distances." *Psychometrika*, **45**, 25–38.
- Cox TF, Cox MAA (2000). *Multidimensional Scaling*. 2nd edition. Chapman & Hall, London.
- Davison ML (1978). *Multidimensional Scaling*. Wiley, New York.
- De Leeuw J, Heiser WJ (1980). "Multidimensional Scaling with Restrictions on the Configuration." In PR Krishnaiah (ed.), *Multivariate Analysis*, volume V, pp. 501–522. North-Holland, Amsterdam.
- De Leeuw J, Mair P (2009). "Multidimensional Scaling Using Majorization: SMACOF in R." *Journal of Statistical Software*, **31**(3), 1–30. URL <http://www.jstatsoft.org/v31/i03/>.
- Ekman G (1954). "Dimensions of Color Vision." *Journal of Psychology*, **38**, 467–474.
- Fellows I (2014). *wordcloud: Word Clouds*. R package version 2.5, URL <https://CRAN.R-project.org/package=wordcloud>.
- Groenen PJJ (1993). *The Majorization Approach to Multidimensional Scaling: Some Problems and Extensions*. Ph.D. thesis, University of Leiden.
- Kruskal JB (1964). "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis." *Psychometrika*, **29**, 1–27.
- Kruskal JB, Wish M (1978). *Multidimensional Scaling*. Sage, Beverly Hills, CA.
- Mair P, Borg I, Rusch T (2016). "Goodness-of-fit Assessment in Multidimensional Scaling and Unfolding." *Multivariate Behavioral Research*. Forthcoming.

- Mair P, De Leeuw J (2015). “Unidimensional Scaling.” In *Wiley StatsRef: Statistics Reference Online*. Wiley, New York.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Schwartz SH, Bilsky W (1987). “Toward a Psychological Structure of Human Values.” *Journal of Personality and Social Psychology*, pp. 550–562.
- Schwartz SH, Cieciuch J, Vecchione M, Davidov E, Fischer R, Beierlein C, Ramos A, Verkasalo M, Lönnquist JE, Demirutku K, Dirilen-Gumus O, Konty M (2014). “Refining the Theory of Basic Individual Values.” *Journal of Personality and Social Psychology*, **103**, 663–688.
- Shepard RN (1962). “Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function.” *Psychometrika*, **27**, 125–140.
- Spence I, Ogilvie JC (1973). “A Table of Expected Stress Values for Random Rankings in Nonmetric Multidimensional Scaling.” *Multivariate Behavioral Research*, **8**, 511–517.
- Torgerson WS (1958). *Theory and Methods of Scaling*. John Wiley & Sons, New York.
- Wish M (1971). “Individual Differences in Perceptions and Preferences among Nations.” In CW King, D Tigert (eds.), *Attitude Research Reaches New Heights*, pp. 312–328. American Marketing Association, Chicago, IL.
- Wish M, Deutsch M, Biener L (1972). “Differences in Perceived Similarity of Nations.” In AK Romney, RN Shepard, SB Nerlove (eds.), *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences*, pp. 289–313. Academic Press, New York.

Affiliation:

Ingwer Borg
Fachrichtung Psychologie
Westfälische-Wilhelms Universität
Fliednerstr. 21
48149 Münster, Germany