Chapter 6
Confirmatory MDS

Abstract Different forms of confirmatory MDS are introduced, from weak forms with external starting configurations to enforcing theoretical constraints onto the MDS point coordinates or onto certain regions of the MDS space.

Keywords Confirmatory MDS · External scales · Dimensional constraints
Shearing · Axial partition · Penalty function

In the MDS models discussed so far, the computer was free to move the points to any positions in space that would minimize the configuration's Stress. This is *exploratory* MDS. If one has clear hypotheses about the MDS configuration, one may be less interested in blindly minimizing Stress, but rather in finding an optimal theory-consistent MDS solution. This leads to *confirmatory* MDS.

6.1 Weak Confirmatory MDS

The least one can do when testing structural theories using MDS is running the MDS with an external initial configuration derived from theory rather than leaving it to the program to choose its own start. This can help finding good solutions in the vicinity of what is expected. One can also fit the MDS solutions thus obtained to theory-based *target* configurations. For example, in case of the rectangle study from Sect. 2.3, the design configuration of Fig. 2.4, appropriately stretched or compressed along its dimensions, can serve both as an initial configuration and also as a target in subsequent Procrustean transformations of the MDS configuration (see Sect. 7.6).

An external initial configuration can also help to make a set of different MDS solutions more similar. Consider a study by Dichtl et al. (1980). These authors analyzed consumer perceptions of various automobiles collected year after year over a period of five years. They first computed the MDS solution of the averaged data and then used this configuration as the initial configuration when scaling each of the

I. Borg et al., *Applied Multidimensional Scaling and Unfolding*,
SpringerBriefs in Statistics, https://doi.org/10.1007/978-3-319-73471-2_6

five yearly data sets. This makes it more likely that the various solutions are more similar, because the MDS algorithm always begins its optimization process with the same configuration.

6.2 External Side Constraints on the Dimensions

A *strict* confirmatory MDS approach *enforces* a solution that satisfies the external constraints while minimizing Stress. The simplest such model is to impose certain restrictions onto dimensions that span the MDS space.

As an application example, we use the rectangle study from Sect. 2.3. Exploratory MDS of these data leads to a solution that closely approximates a psychologically reasonable transformation of the design grid (Fig. 2.5). We now employ confirmatory MDS to enforce such a grid *perfectly* onto the solution and then check whether this leads to Stress values that are still acceptably low. This can be realized by the `smacofConstraint()` function. It allows the user to request that an $n \times m$ MDS solution $\mathbf{X}$ is generated by optimally scaling the column vectors of an external $n \times m$ matrix $\mathbf{Y}$. For $\mathbf{Y}$, we here take the coordinates of the points in the design grid, i.e., their width and height measurements (see Fig. 2.4 or simply activate these data by `data(rect_constr)`). The columns of $\mathbf{Y}$ are called *external scales*, and after *optimal re-scaling*, they become the *internal scales*, the columns of $\mathbf{X}$.

Re-scaling can mean different things:

- In the simplest case, it means dimensional weighting. That is, the data are approximated, as much as possible, by the distances computed on a configuration whose dimensions are the optimally weighted columns of $\mathbf{Y}$. Expressed formally $\mathbf{X} = \mathbf{YC}$, with $\mathbf{C}$ a diagonal matrix that minimizes the Stress of $\mathbf{X}$.
- If we drop the constraint that $\mathbf{C}$ is diagonal, then $\mathbf{C}$ becomes a *composite* transformation. It can be understood as a rotation/reflection followed by dimensional weighting and then rotated/reflected once more. Thus, expressed geometrically, the dimensional weighting can be done along a *rotated* set of dimensions.
- A third case is allowing for optimal monotone transformations of $\mathbf{Y}$'s columns or of the columns of a rotated $\mathbf{Y}$.

For the rectangle data, the third model is theoretically most convincing. We test it by first running exploratory MDS and then plotting this solution with its points connected as a grid. Then, we use this solution as the initial configuration in confirmatory MDS,[1] enforcing an ordinal rescaling of the unrotated design dimensions. Finally, we also allow for a rotation of the design configuration.

---

[1]If no external initial configuration is provided, the program will use a random start. In most applications, this will not lead to low Stress nor to a meaningful solution.

```
## MDS with theory-based initial configuration
fit.expl <- mds(rectangles, type = "ordinal", init = rect_constr)
## MDS enforcing an ordinally re-scaled design grid
fit.cfdiag <- smacofConstraint(rectangles, constraint = "diagonal",
                type = "ordinal", ties = "secondary",
                init = fit.expl$conf, external = rect_constr,
                constraint.type = "ordinal")
## Confirmatory MDS, also permitting a rotation of the design grid
fit.cflin  <- smacofConstraint(rectangles, constraint = "linear",
                type = "ordinal", ties = "secondary",
                init = fit.expl$conf, external = rect_constr,
                constraint.type = "ordinal")
```

Figure 6.1 shows the resulting configurations. The exploratory MDS solution (left panel) is already nearly theory-compatible except for some small dents of the grid. Its Stress is 0.089. The first confirmatory solution (middle panel) is theory-wise perfect, with a Stress of 0.115. Hence, the dents of the grid in the exploratory MDS solution do not explain the data "much" better. Rather, it seems that they essentially represent some of the data noise. So, one may decide not to reject the hypothesis that the observed judgments for the rectangles' similarity are generated by a composition rule that behaves just like the distance formula operating on the rectangles' design dimensions.

If we drop the diagonality constraint on **C**, we get the *sheared* grid in the right panel of Fig. 6.1. Its Stress is 0.103, slightly better than without the rotation. It suggests that not the original dimensions were rescaled but a slightly rotated (but theoretically obscure) dimension system. This causes the shearing of the grid. (In practice, such shearings can become extreme in this model which make the solutions difficult to interpret.)

If we set `constraint.type="interval"`, the transformations on the design grid are limited to stretchings of the external scales, i.e., to simple dimensional weightings (plus possible shearings). Under this condition, the successively smaller compressions of the grid along its dimensions generated by `constraint.type="ordinal"`

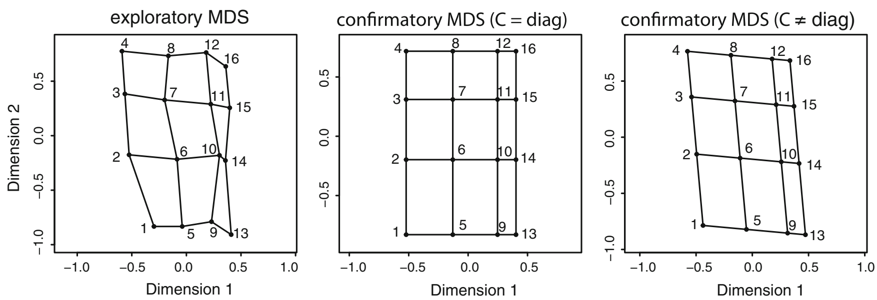

Fig. 6.1 Exploratory (left) MDS of rectangle data of Sect. 2.3; and confirmatory MDS of the same data with stretchings of the given dimensions (center panel) and with stretchings of rotated dimensions (right panel)

cannot occur anymore in the MDS solution. This would be undesirable here, because the Weber–Fechner law of perception is predicting such logarithmic shrinkage effects.

Returning to the model equation $\mathbf{X} = \mathbf{YC}$, we note that the matrix $\mathbf{C}$ represents a *linear transformation* of the configuration $\mathbf{Y}$. Any linear transformation can be decomposed into rotations and dimensional weightings of $\mathbf{Y}$. Algebraically, that means that $\mathbf{C}$ can be split by singular value decomposition into the product $\mathbf{PMQ}$, where $\mathbf{P}$ and $\mathbf{Q}$ represent rotations and $\mathbf{M}$ is a diagonal matrix of dimension weights. Thus, $\mathbf{C}$ first rotates the configuration $\mathbf{Y}$ in some way and then stretches and/or compresses this *rotated* configuration along its dimensions and finally rotates the result once more. If $\mathbf{C}$ is a diagonal matrix, then the column vectors of $\mathbf{Y}$ are weighted directly. If $\mathbf{C}$ is not diagonal, then $\mathbf{Y}$ is first rotated and then dimensionally weighted, and this is what causes the shearing.

A different approach to impose external constraints onto the MDS solution is to focus on the distances of the MDS configuration, not on its coordinates. If, for example, one requests for the rectangle data that $d(1, 6) = d(2, 5)$, $d(6, 11) = d(7, 10)$, and $d(11, 16) = d(12, 15)$ must hold in the MDS solution, shearings of the point grid are avoided. To guarantee that a grid is generated in the first place, one can additionally enforce that some of the horizontal grid distances be equally long, e.g., that $d(1, 5) = d(2, 6) = d(3, 7) = d(4, 8)$, $d(5, 9) = d(6, 10) = d(7, 11) = d(8, 12)$, and $d(9, 13) = d(10, 14) = d(11, 15) = d(12, 16)$. Restrictions like these can be imposed on the MDS configuration by the program CMDA (Borg and Lingoes 1980). CMDA is, unfortunately, an old Fortran program that is not easily accessible and difficult to use because it is not always easy to derive what a given theory implies for the distances among the points in MDS space.

6.3 Regional Axial Restrictions

One can use the methods discussed above to solve confirmatory MDS problems that arise quite frequently in applied research, that is, impose particular *axial partitions* onto the MDS solution. Here is an example. Rothkopf (1957) studied to what extent test persons confused different acoustic Morse signals. He used 36 different signals, the 26 letters of the alphabet, and the natural numbers from 0 to 9. The signal for A, for example, is “di” (a beep with a duration of 0.05 s), followed by a pause (0.05 s) and then by “da” (0.15 s). We code this as 1–2 or 12 for di-da.

The symmetrized confusion probabilities collected for these signals from hundreds of test persons can be represented quite well in a two-dimensional MDS configuration (Fig. 6.2). The partitioning lines were inserted by hand. They cut the plane in two ways, related to two facets: The nine solid lines discriminate the signals into classes of signals with the same total duration (from 0.05 to 0.95 s); the five dashed lines separate the signals on the basis of their composition (e.g., signals containing only long beeps are all on the right-hand side). The pattern of these partitioning lines is not very simple, though, but partially rather curvy and hard to

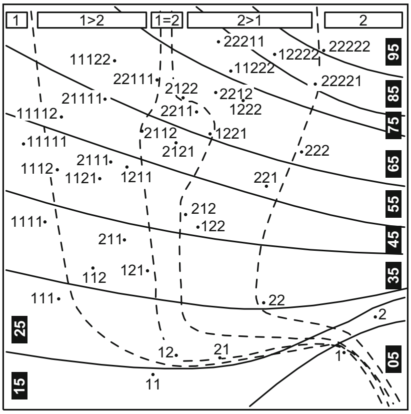

Fig. 6.2 Exploratory MDS representation for 36 Morse signals; lines correspond to two typologies for the signals

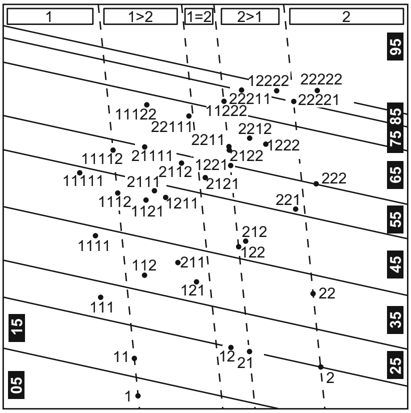

Fig. 6.3 Confirmatory MDS solution with two perfect axial partitioning lines

describe. Particularly, the dashed lines are so twisted that the pattern of the emerging regions does not exhibit a simple law of formation. Rather, the partitioning seems over-fitted. The substantive researcher, therefore, would probably not bet that it can be completely replicated with new data.

We now want to straighten the two sets of partitioning lines. For that purpose, we again use the $\mathbf{X} = \mathbf{YC}$ restriction. To generate the internal scales in $\mathbf{X}$, we make use of two of the signal codes' properties, duration and type, as shown in Fig. 6.2 by the vertical black boxes (duration) and the boxes on top labeled as "1", "$1 > 2$", "$1 = 2$", "$2 < 1$", and "2" (type). Each Morse code is thus coded in terms of its duration into one of ten categories and in terms of type into one of five categories. This defines the external variables, $\mathbf{Y}$. They can be viewed by typing `data(morsescales); morsescales` in SMACOF.

With these constraints in an ordinal MDS, with ordinal external scales, and with the primary approach to ties, we find the solution in Fig. 6.3. This simple-to-interpret MDS solution has almost the same overall Stress as the exploratory MDS solution in Fig. 6.2 (0.21 vs. 0.18). Upon closer investigation one notes, however, that the confirmatory solution moved only very few points by more than a small amount. Particularly, point 1 (at the bottom, to the right) was moved a lot so that the substantive researcher may want to study this signal (and its relationship to other stimuli such as signal 2) more closely. Overall, though, the simpler and, probably, also more replicable solution in Fig. 6.3 appears to be the better springboard for further research.

6.4 Circular and Spherical MDS

Spherical MDS is an MDS model where all points lie on the surface of an m-dimensional sphere. There are data sets where it can be argued that spherical MDS is more relevant than the usual flat-geometry MDS, but the really interesting case is $m = 2$, i.e., the case where spherical MDS becomes *circular* MDS. Circular scales abound in psychology. Two prominent examples are color perception (see Sect. 5.4 on p. 60ff.) and the psychology of personal values (see p. 21ff., and Chap. 8).

For personal values, we used exploratory MDS to study the structure of the intercorrelations of value items. Figures 2.10 and 2.13 indicate that the value items and the value indexes form approximately circular configurations of points. We may ask how much the Stress goes up if the points of the configurations were forced onto perfect circles. An answer is found by using the `smacofSphere()` function: The Stress of the exploratory solution is 0.051; it goes up to 0.085 in the perfect-circle solution.

Enforcing a perfect circle for these data does, however, not really lead to new insights, since the exploratory configuration is already roughly circular. Moreover, a perfect circle is not needed for indexes that are based on real and therefore error-affected data. To see more dramatic or unexpected effects, let us therefore request a circular MDS configuration for the similarity of countries data represented in Fig. 2.2. Since there is no substantive reason to enforce a circle, we should expect that this constraint entails a substantial increment in Stress.

When running this type of analysis with `smacofSphere()`, we have a choice of two algorithms: The primal algorithm enforces a strict circle from the beginning, and the dual algorithm uses a penalty function that pushes the MDS solution in the direction of a perfect circle. The default penalty weight is 100, and when setting it to 22, say, the force that pulls the solution toward a perfect circle is mitigated. Let us try both specifications as follows:

```
diss <- sim2diss(wish, method=max(wish))
res1 <- smacofSphere(diss, type="ordinal")
res2 <- smacofSphere(diss, type="ordinal", algorithm="dual", penalty=22)
res3 <- mds(diss)
res1$stress; res2$stress; res3$stress ## gives Stress values of each solution
op <- par(mfrow = c(1,3))
plot(res1, main="Circular MDS (primal)")
plot(res2, main="Circular MDS 2 (dual)")
plot(res3, main="Exploratory MDS")
par(op)
```

The three results are shown in Fig. 6.4. As expected, the solution generated by the default algorithm (`algorithm="primal"`) has all country points on a perfect circle, while the solution computed by the dual algorithm and using `penality=22` only comes close to a perfect circle. When setting the penalty weight to 100 (i.e., the default value), then the circle is perfect too. So, we see that choosing smaller penalty weights is a way to avoid that the algorithm is pushing too hard toward a perfect circle.

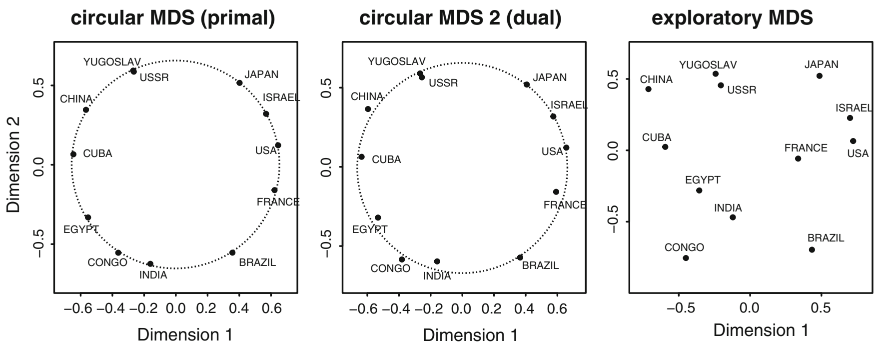

Fig. 6.4 Circular MDS using the primal and the dual algorithm, resp., and exploratory MDS (right panel) of Wish country similarity data

The Stress values for the solutions in Fig. 6.4 are 0.271, 0.266, and 0.225, respectively. The increment in Stress is not much higher than it is in case of the personal values example discussed above, even though now there are 12 points and not just 10 points. However, one should take into account that the Stress for the Wish data is quite high even without circular side constraints. Nevertheless, when studying the three plots more closely, one can indeed see that the exploratory MDS solution is not that far from being circular: Only France needs to be pulled somewhat to the outside and Congo more toward the center of the plot. Whether a circular configuration for the countries is substantively meaningful is, of course, another question.

When testing theories about real data, forcing the points onto a perfect circle in MDS space may seem exaggerated formalism. An approximate circle would be sufficient, but it is much harder to formulate this idea as a clear scaling target. Moreover, a perfect circle is, by itself, rarely ever a meaningful structural theory. It only becomes interesting if it is supplemented with additional notions such as a particular order of the points on the circle. In case of the data on personal values (see p. 25), the Theory on Universals in Values (Schwartz 1992) predicts such an order. The theory also claims that the point order is structured into four subsets of opposite higher-order personal values. This would split the circle into four arcs that lie in four different quadrants. If you have inter-correlations as data, circular scaling solutions with various additional constraints can be generated using the R package CIRCE (Grassi et al. 2010). This program implements the Guttman–Browne circumplex model for inter-correlations (Browne 1992). It assumes that an observed correlation r_{ij} corresponds to an angle between the vectors pointing to the points i and j on a unit circle. The method does not accept order constraints, but they can be approximated to some extent by restricting the points to lie in certain sectors of the circle. For example, with the personal values grouped into four higher-order values, and the PVQ40 data aggregated into ten indices as in the first 11 lines of the R script on p. 23, the R commands are:

```
require(CircE); R <- cor(PVQ40agg)
## CircE commands (with lots of default agruments):
lower1 <- c(0,0,0,270,270,180,180,180,90,90) ## lower bounds for point angles
upper1 <- c(90,90,90,360,360,270,270,270,180,180) ## upper bounds
res <- CircE.BFGS(R, v.names=colnames(R), m=1, N=10, upper=upper1, lower=
     lower1, equal.com=FALSE, equal.ang=FALSE)
CircE.Plot(res, ef=0.1)
```

`CircE` computes a circular configuration together with extensive output, including many fit indexes such as GIF, AGIF, RMSEA that are used in structural equation modeling. They test the hypothesis that the observed correlations match the correlations derived from the model. See (Grassi et al. 2010) for detailed examples. For the above PVQ40 data, the fit is highly significant, and the results are quite similar to what is shown in Fig. 2.13.

6.5 Challenges of Confirmatory MDS

The challenges of confirmatory MDS for the user are, most of all, how to formulate theoretical expectations so that they can be expressed in, say, a penalty function, a pseudo-data matrix, or a system of equations that can be solved by an existing confirmatory MDS program. Confirmatory MDS, therefore, is often much harder than exploratory MDS, because it requires the user to not only develop explicit theories but also translating them into a proper computational language. So far, the MDS programs accessible to the general user can handle only relatively simple confirmatory analyses. Dimensional restrictions are easy to test, while confirmatory MDS analyses with regional restrictions are typically difficult to set up and solve. Computer programs that allow all forms of restrictions (combined, in addition, with particular MDS models, certain missing data patterns, or distances other than Euclidean distances) do not exist yet. Rather, in such cases, a suitable MDS algorithm must be programmed ad hoc.

If the users succeed generating a confirmatory MDS solution, a number of additional challenges await them. They have to evaluate not only the absolute Stress values, but also the Stress increment resulting from adding the particular external constraints to the MDS analysis. Typically, such evaluations amount to deciding whether the Stress increment is substantial or not, given the number of points, the dimensionality of the MDS space, the MDS model, the distance function, and the quality of the data (error level). These and further criteria are summarized by Lingoes and Borg (1983) in a quasi-statistical decision procedure.

An important additional criterion is the strength of the external constraints. These constraints may be easy to satisfy for a given number of points in a given dimensionality, but they may also be quite demanding. An approach for evaluating this issue is described in Borg et al. (2011). They use data from a survey where a sample of employees assessed 54 organizational culture themes (e.g., “being competitive,” “working long hours,” and “being careful”) in terms of how important they are for them personally. The correlations of these importance ratings are represented in a theory-compatible MDS solution, where the 54 points are forced into the quadrants of a 2d coordinate system on the basis of a priori codings of the items in terms of the TUV theory. The strength of the external constraints is assessed by studying the Stress values that result from running 1,000 different confirmatory MDS analyses, each one using a random permutation of these TUV codings. It is found that the theory-based assignment of codes to the 54 items does indeed lead to a Stress value that is smaller than any of the Stress values that are found if random permutations of the codings are enforced onto the MDS solution. Hence, the codings are *not trivial* in the sense that random assignments of the codings would lead to equally good MDS solutions when enforced onto the configuration.

6.6 Summary

MDS is mostly used in an exploratory way, where the MDS configuration is chosen so that the Stress is minimal. Confirmatory MDS enforces additional structure onto the MDS space, or it at least tries to push the solution toward a theoretically expected structure. Confirmatory MDS configurations may be very different from exploratory MDS solutions. Often, their Stress is higher, but sometimes it is not. Without running confirmatory MDS, one would not know. A weak way to push an MDS solution toward a theoretical structure is using a theory-derived initial configuration. Harder confirmatory requirements need special MDS programs such as PROXSCAL or `smacofConstraint`. With such programs, one can enforce certain dimensional requirements and strict axial partitionings. Circular configurations require spherical MDS programs such as `smacofSphere()` or `CircE`.

References

Borg, I., Groenen, P. J. F., Jehn, K. A., Bilsky, W., & Schwartz, S. H. (2011). Embedding the organizational culture profile into Schwartz's theory of universals in values. *Journal of Personnel Psychology*, *10*, 1–12.

Borg, I., & Lingoes, J. C. (1980). A model and algorithm for multidimensional scaling with external constraints on the distances. *Psychometrika*, *45*, 25–38.

Browne, M. W. (1992). Circumplex models for correlation matrices. *Psychometrika*, *57*, 469–497.

Dichtl, E., Bauer, H. H., & Schobert, R. (1980). Die Dynamisierung mehrdimensionaler Marktmodelle am Beispiel des deutschen Automobilmarkts. *Marketing*, *3*, 163–177.

Grassi, M., Luccio, R., & Di Blas, L. (2010). CircE: An R implementation of Browne's circular stochastic process model. *Behavior Research Methods*, *42*, 55–73.

Lingoes, J. C., & Borg, I. (1983). A quasi-statistical model for choosing between alternative configurations derived from ordinally constrained data. *British Journal of Mathematical and Statistical Psychology*, *36*, 36–53.

Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning. *Journal of Experimental Psychology*, *53*, 94–101.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, *25*, 1–65.