10
Confirmatory MDS

If more is known about the proximities or the objects, then additional restrictions (or constraints) can be imposed on the MDS model. This usually means that the MDS solutions must satisfy certain additional properties of the points' coordinates or the distances. These properties are derived from substantive considerations. The advantage of enforcing such additional properties onto the MDS model is that one thus gets direct feedback about the validity of one's theory about the data. If the Stress of a confirmatory MDS solution is not much higher than the Stress of a standard ("unconstrained") MDS solution, the former is accepted as an adequate model. Several procedures that allow one to impose such external constraints are described and illustrated.

10.1 Blind Loss Functions

In most MDS applications discussed so far, we did not just represent the data geometrically and then interpret the solutions, but started by formulating certain predictions on the structure of the MDS configuration. For example, in Section 4.1, it was conjectured that the similarity scores on the colors would lead to a circular point arrangement in a plane (color circle). In Section 4.3, it was predicted that the similarity data on facial expressions could be explained by a 3D coordinate system with specified axes. However, these predictions had no influence on the MDS solution. Rather, structural hypotheses were dropped when the data were handed over to

an MDS computer program. The MDS program optimizes Stress, which is substantively blind; that is, it is not tailored to the particular questions that are being asked. The program mechanically grinds out "some" optimal distance representation under a few general restrictions such as the dimensionality m and the admissible transformations on the proximities.

Minimizing Stress gives a solution that is locally optimal. Yet, other local minimum solutions may exist with a similar Stress, or possibly even with lower Stress (see also Section 13.4). The question is which solution should be preferred. If a hypothesis for the data is available, then, of course, we would be particularly interested in the solution that most directly speaks to this hypothesis. This is obviously the solution that most closely satisfies the hypothesis, even if its Stress is somewhat higher than the Stress for other solutions.

Assume, for example, that we had not obtained the color circle in Figure 4.1 because the computer program succeeded in finding a solution with a lower Stress value. Assume further that the formally optimal but theoretically unintelligible solution had Stress .05, but the one matching our predictions had .06. Having had only the Stress-optimal solution, we probably would have concluded—incorrectly—that the predictions were wrong. Thus, what we want is a method that guarantees that the solution satisfies our expectations. Once we have it, we can decide whether this solution has an acceptable fit to the data.

10.2 Theory-Compatible MDS: An Example

Consider an example. Noma and Johnson (1977) asked subjects to assess the similarity of 16 ellipses having different shapes and sizes. The ellipses were constructed according to the design shown in Figure 10.1. The horizontal dimension of this design configuration is eccentricity ("shape"), and the vertical, area ("size").[1] The design shows, for example, that ellipse 4 is very flat and long, but 13 is more circular and also larger.[2] The subjects had to rate each pair of ellipses on a scale from 1 ("most similar") to 10 ("least similar or most different"). This rating was replicated three times, with the pairs presented in different random orders. Table 10.1 gives the aggregated scores for one individual.

---

[1]Eccentricity is defined as $[1-(h/n)^2]^{1/2}$ and area is $\pi/4 \cdot h \cdot n$, where h is the length of the ellipse's major axis and n is the length of its minor axis. Hence, eccentricity is a function of the ratio of h and n, and area depends on the product of h and n.

[2]Noma and Johnson (1977, p. 31) characterize the design as follows: "A factorial design with four equally spaced levels of area crossed with four equally spaced levels of eccentricity was employed in constructing the stimuli. The largest ellipse was in a 3:1 ratio to the smallest, and the most eccentric was in a 1.66:1 ratio to the least eccentric."

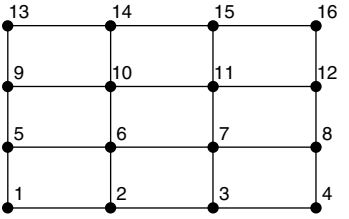

FIGURE 10.1. Design configuration for ellipses in Noma–Johnson study. X-axis is eccentricity ("shape"); Y-axis is area ("size").

TABLE 10.1. Dissimilarities for 16 ellipses; summed over three replications of subject DF (Noma & Johnson, 1977).

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	–															
2	18	–														
3	26	9	–													
4	23	14	12	–												
5	4	17	17	25	–											
6	15	6	13	13	18	–										
7	16	17	10	6	21	17	–									
8	24	20	13	10	24	17	12	–								
9	8	16	20	22	4	16	24	24	–							
10	16	9	16	14	21	8	11	10	14	–						
11	20	13	11	9	18	12	8	11	22	12	–					
12	22	18	17	12	21	21	12	6	23	13	11	–				
13	16	16	21	24	13	16	22	23	4	16	21	22	–			
14	17	14	16	19	20	9	14	17	13	4	14	17	17	–		
15	21	20	15	9	25	14	8	11	19	19	4	16	22	17	–	
16	26	19	14	15	24	16	11	12	22	16	10	6	30	17	9	–

From related research (see Section 17.4) it could be expected that an MDS configuration similar to the design configuration would result from the proximities. That is, the MDS configuration should form a rectangular grid as in Figure 10.1, although not necessarily with the same spacing of the vertical and horizontal lines. This would allow us to explain the similarity judgments by the dimensions "shape" and "area". Ordinal MDS of the data in Table 10.1 yields, however, a configuration (Figure 10.2) that is in definite disagreement with these predictions. But, then, a theory-conforming configuration does not necessarily have to have the *lowest-possible Stress*. Rather, it would be sufficient if it had an *acceptably low Stress*. Indeed, such a solution exists. It is shown in Figure 10.3. Its Stress is .185, as compared to .160 for the theory-incompatible solution in Figure 10.2. This example shows that there can be different MDS configurations that all represent a given set of data with roughly the same precision.

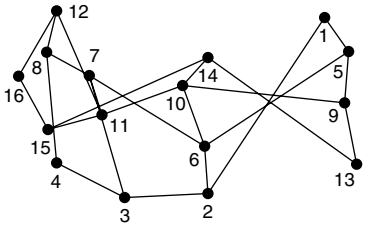

FIGURE 10.2. Minimal-Stress MDS representation for data in Table 10.1.

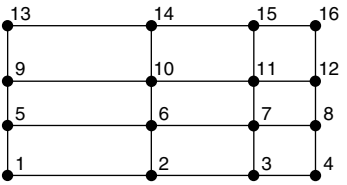

FIGURE 10.3. Minimal-Stress theory-compatible MDS.

10.3 Imposing External Constraints on MDS Representations

We now show how a confirmatory MDS procedure can be constructed. We begin by considering the task of constructing an ordinal MDS representation of the facial expression data from Table 4.4 of Abelson and Sermat (1962) in 2D so that (a) the Stress is as low as possible, and (b) all points can be coordinated by dimensions that are a linear combination of the three external scales of Engen et al. (1958). Condition (b) is an additional requirement imposed on the MDS representation. It is called a side constraint or an *external constraint* to distinguish it from the *internal* constraints due to the data and the general representation function.

The restriction that is imposed on the configuration is

$$\mathbf{X} = \mathbf{YC},$$

where $\mathbf{Y}$ is the 13×3 matrix with the three external scales of Table 4.3, and $\mathbf{C}$ is a 3×2 matrix of unknown weights. This matrix equation is shown explicitly in Table 10.2. The mathematical problem to be solved is to minimize $\sigma_r^2(\mathbf{X})$ by an appropriate choice of $\mathbf{C}$, subject to the condition $\mathbf{X} = \mathbf{YC}$. Solutions for this problem were given by Bentler and Weeks (1978), Bloxom (1978), and De Leeuw and Heiser (1980). We follow the approach of De Leeuw and Heiser (1980), because they show that this and more general constrained MDS models can be handled easily within the majorization framework (see Chapter 8).

As shown in (8.27), raw Stress can be majorized by

$$\tau(\mathbf{X}, \mathbf{Z}) \quad = \quad \eta_\delta^2 + \text{tr } \mathbf{X}'\mathbf{V}\mathbf{X} - 2\text{tr } \mathbf{X}'\mathbf{B}(\mathbf{Z})\mathbf{Z}, \tag{10.1}$$

which is equal to $\sigma_r(\mathbf{X})$ if $\mathbf{Z} = \mathbf{X}$; that is, $\sigma_r(\mathbf{X}) = \tau(\mathbf{X}, \mathbf{X})$. Let $\overline{\mathbf{X}} = \mathbf{V}^+\mathbf{B}(\mathbf{Z})\mathbf{Z}$ be the Guttman transform (8.28) of $\mathbf{Z}$, where $\mathbf{Z}$ satisfies the imposed constraints. Then (10.1) equals

$$\begin{aligned} \tau(\mathbf{X}, \mathbf{Z}) \quad &= \quad \eta_\delta^2 + \text{tr } \mathbf{X}'\mathbf{V}\mathbf{X} - 2\text{tr } \mathbf{X}'\mathbf{V}\overline{\mathbf{X}} \\ &= \quad \eta_\delta^2 + \text{tr } (\mathbf{X} - \overline{\mathbf{X}})'\mathbf{V}(\mathbf{X} - \overline{\mathbf{X}}) - \text{tr } \overline{\mathbf{X}}'\mathbf{V}\overline{\mathbf{X}}. \end{aligned} \tag{10.2}$$

TABLE 10.2. Matrix equation $\mathbf{X} = \mathbf{YC}$ in (10.1), with $\mathbf{Y}$ taken from Table 4.3; $\mathbf{X}$ is the desired 13×2 MDS configuration; $\mathbf{C}$ is an unknown matrix of weights.

$$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \\ x_{51} & x_{52} \\ x_{61} & x_{62} \\ x_{71} & x_{72} \\ x_{81} & x_{82} \\ x_{91} & x_{92} \\ x_{10,1} & x_{10,2} \\ x_{11,1} & x_{11,2} \\ x_{12,1} & x_{12,2} \\ x_{13,1} & x_{13,2} \end{bmatrix} = \begin{bmatrix} 3.8 & 4.2 & 4.1 \\ 5.9 & 5.4 & 4.8 \\ 8.8 & 7.8 & 7.1 \\ 7.0 & 5.9 & 4.0 \\ 3.3 & 2.5 & 3.1 \\ 3.5 & 6.1 & 6.8 \\ 2.1 & 8.0 & 8.2 \\ 6.7 & 4.2 & 6.6 \\ 7.4 & 6.8 & 5.9 \\ 2.9 & 3.0 & 5.1 \\ 2.2 & 2.2 & 6.4 \\ 1.1 & 8.6 & 8.9 \\ 4.1 & 1.3 & 1.0 \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}$$

For a given configuration $\mathbf{Z}$, only the second term of $\tau(\mathbf{X}, \mathbf{Z})$ is dependent on $\mathbf{X}$, whereas the first and last terms are constant with respect to $\mathbf{X}$. Using $\sigma_r(\mathbf{X}) = \tau(\mathbf{X}, \mathbf{X})$, (10.2) shows that, for every configuration $\mathbf{X}$, raw Stress is the sum of *lack of model fit*, $\eta_\delta^2 - \text{tr } \overline{\mathbf{X}}'\mathbf{V}\overline{\mathbf{X}}$, and *lack of confirmation fit*, $\text{tr } (\mathbf{X} - \overline{\mathbf{X}})'\mathbf{V}(\mathbf{X} - \overline{\mathbf{X}})$. The latter is best expressed as a percentage of the total raw Stress. For example, if there are no constraints on $\mathbf{X}$, then the lack of confirmation fit is 0%.

Finding a constrained update amounts to minimizing

$$\text{tr } (\mathbf{X} - \overline{\mathbf{X}})'\mathbf{V}(\mathbf{X} - \overline{\mathbf{X}}), \tag{10.3}$$

subject to the restrictions on $\mathbf{X}$, in each iteration. The $\mathbf{X}$ that minimizes (10.3) *and* satisfies the constraints is used as the update. Thus, step 6 (computation of the Guttman transform) in the majorization algorithm of Chapter 9 (see also Figure 9.2) is followed by step 6a, in which (10.3) is minimized over $\mathbf{X}$, subject to the constraints on $\mathbf{X}$. De Leeuw and Heiser (1980) note that it is not necessary to find the global minimum of (10.3). The decrease of Stress is guaranteed as long as

$$\text{tr } (\mathbf{X}^u - \overline{\mathbf{X}})'\mathbf{V}(\mathbf{X}^u - \overline{\mathbf{X}}) \leq \text{tr } (\mathbf{Y} - \overline{\mathbf{X}})'\mathbf{V}(\mathbf{Y} - \overline{\mathbf{X}}) \tag{10.4}$$

holds for the update $\mathbf{X}^u$.

For the faces data, we simply substitute $\mathbf{X}$ by $\mathbf{YC}$ in (10.3), which yields

$$\begin{aligned} L(\mathbf{C}) &= \text{tr } (\mathbf{YC} - \overline{\mathbf{X}})'\mathbf{V}(\mathbf{YC} - \overline{\mathbf{X}}) \\ &= \text{tr } \overline{\mathbf{X}}'\mathbf{V}\overline{\mathbf{X}} + \text{tr } \mathbf{C}'\mathbf{Y}'\mathbf{VYC} - 2\text{tr } \mathbf{C}'\mathbf{Y}'\mathbf{V}\overline{\mathbf{X}}. \end{aligned}$$

$L(\mathbf{C})$ needs to be minimized over $\mathbf{C}$, because $\mathbf{Y}$ is fixed (see also Section 8.3). Finding the optimal weights $\mathbf{C}^u$ is a simple regression problem that is solved by

$$\mathbf{C}^u = (\mathbf{Y}'\mathbf{VY})^{-1}\mathbf{Y}'\mathbf{V}\overline{\mathbf{X}},$$

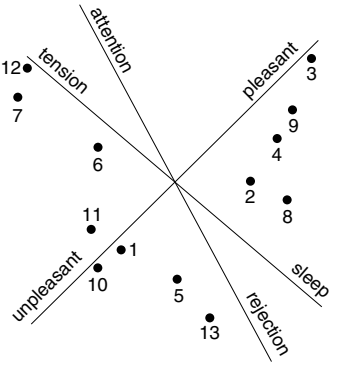

FIGURE 10.4. Theory-consistent solution by constrained MDS of the facial expression data of Abelson and Sermat (1962) with Stress .14.

so that

$$\mathbf{X}^u = \mathbf{Y}\mathbf{C}^u.$$

The unconstrained 2D solution of the faces data in Figure 4.9 has $\sigma_1 = .11$. The theory-compatible solution (where the dimensions are constrained to be linear combinations of three external scales) has $\sigma_1 = .14$ (see Figure 10.4). The unconstrained and constrained solutions only differ with respect to point 8. The raw Stress of the constrained solution is 0.0186 ($= \sigma_r$), of which .0182 (=97%) is the lack of model fit and .0004 (=3%) is Stress due to the constraints. Therefore, the theory-consistent solution can be accepted at the cost of a slightly higher Stress value. The optimal $\mathbf{C}$ equals

$$\mathbf{C} = \begin{bmatrix} .219 & .031 \\ -.035 & .137 \\ -.024 & .053 \end{bmatrix}.$$

What does $\mathbf{C}$ do to $\mathbf{Y}$? One way of interpreting $\mathbf{C}$ as an operator on $\mathbf{Y}$ is to decompose $\mathbf{C}$ into its singular value components:

$$\mathbf{C} = \mathbf{P}\mathbf{\Phi}\mathbf{Q}' = \begin{bmatrix} .985 & .168 \\ -.140 & .919 \\ -.102 & .357 \end{bmatrix} \begin{bmatrix} .224 & .000 \\ .000 & .150 \end{bmatrix} \begin{bmatrix} 1.000 & -.029 \\ .029 & 1.000 \end{bmatrix}.$$

$\mathbf{C}$ first rotates $\mathbf{Y}$ by $\mathbf{P}$, because $\mathbf{P}$ is orthonormal,[3] and takes out the third dimension. Then, $\mathbf{\Phi}$ multiplies the X-axis by .224 and the Y-axis by .150.

[3]In fact, the orthonormality of $\mathbf{P}$ only implies that $\mathbf{P}'\mathbf{P} = \mathbf{I}$, but not that $\mathbf{P}'\mathbf{P} = \mathbf{I}$, as required for a rotation matrix. However, $\mathbf{P}$ can be interpreted as a matrix that rotates $\mathbf{Y}$ to two dimensions.

Finally, $\mathbf{Q}'$ rotates the space somewhat, but only in the X–Y-plane. Note that the final rotation by $\mathbf{Q}'$ could as well be omitted, because it does not change the distances. Thus, $\mathbf{C}$ can be understood as a transformation of the 3D space of the external scales that rotates this space into a plane and also stretches this plane differentially along its coordinate axes so that the resulting configuration has minimal Stress.

The three external variables are plotted as lines such that the angles of the lines correspond to correlations between the variables and the X- and Y-axes, respectively (Figure 10.4). In comparison to Figure 4.9 (where the external variables were fitted *afterwards*, not simultaneously), variable pleasant/unpleasant has about the same direction, whereas the variables sleep/tension and attention/rejection have been interchanged. Because the latter two variables are highly intercorrelated, however, this interchange does not lead to a much different interpretation.

External Constraints and Optimal Scaling

Instead of the linear constraints used above, a multitude of other constraints can be used for which the least-squares solution of (10.4) can be computed, each constraint leading to a different model. De Leeuw and Heiser (1980) discuss many sorts of constraints, some of which are shown below. Apart from the general majorization result that Stress is reduced in every iteration, they also prove several other convergence results if the global minimum of (10.4) can be established. A more accessible overview of constrained MDS and its applications is given by Heiser and Meulman (1983b).

The facial expression data were analyzed by Heiser and Meulman (1983b) in a slightly different way. They used only the ordinal information of the three external variables of the faces data. Let $\mathbf{Y} = [\mathbf{y}_1\ \mathbf{y}_2\ \mathbf{y}_3]$ be the matrix of the three external variables. The constraints on the MDS solution are $\mathbf{X} = \widehat{\mathbf{Y}}\mathbf{C}$, where $\widehat{\mathbf{Y}} = [\widehat{\mathbf{y}}_1\ \widehat{\mathbf{y}}_2\ \widehat{\mathbf{y}}_3]$ and each column $\widehat{\mathbf{y}}_k$ can be *optimally scaled* (see, e.g., Young, 1981; Gifi, 1990). In optimal scaling of an ordinal variable, the original variable $\mathbf{y}_k$ is replaced by a different variable $\widehat{\mathbf{y}}_k$ that has the same order as the original variable *and* reduces the loss function, hence the name optimal scaling. Ordinal transformations on the external variables are computed using monotone regression, and implemented in the programs SMACOF-II (Meulman, Heiser, & De Leeuw, 1983) and PROXSCAL (see Appendix A). Thus, (10.4) was not only optimized over $\mathbf{C}$, but also over $\widehat{\mathbf{Y}}$, where every column of $\widehat{\mathbf{Y}}$ is constrained to have the order of the corresponding external variable $\mathbf{y}_k$. Apart from finding an optimal transformation of the proximities, an optimal transformation of the external variables is also found here. In their analysis of the facial expression data, Heiser and Meulman (1983b) conclude that pleasant–unpleasant is a more basic dimension than attention–rejection, which is a nonlinearly related ef-

fect. Optimal scaling of external variables allows interesting models to be specified, such as the one below.

If only two external variables are involved in a 2D MDS space, then the ordinal restrictions on the two external variables result in dimensional restrictions or in an axial partitioning of the space. For example, the hypothesized grid-like structure in Figure 10.3 was enforced onto the MDS configuration by the two external variables

$$\begin{aligned} \mathbf{y}_1' &= [1,2,3,4,1,2,3,4,1,2,3,4,1,2,3,4] \text{ and} \\ \mathbf{y}_2' &= [1,1,1,1,2,2,2,2,3,3,3,3,4,4,4,4]. \end{aligned}$$

Then, $\mathbf{X}$ is obtained by $\mathbf{X} = \widehat{\mathbf{Y}}\mathbf{C}$, with the secondary approach to ties (i.e., ties remain tied). The external variables $\mathbf{y}_1$ and $\mathbf{y}_2$ are derived from the design configuration in Figure 10.1. Thus, the 2-tuple $(f_1(y_{i1}), f_2(y_{i2}))$ contains the coordinates of each point i, where f_1 and f_2 are Stress-minimizing monotonic functions that use the secondary approach to ties. These requirements come from psychophysics. We do not expect that an ellipse that is twice as eccentric in terms of the ratio of its axes is also perceived as twice as eccentric, for example. Rather, we would expect by the Weber–Fechner law that perceived eccentricity should be related to "objective" eccentricity in a roughly logarithmic way, in general. Indeed, that is exactly what the data show in Figure 10.3. Note that we did not enforce a logarithmic spacing on the horizontal axis. Rather, this function was found by MDS as the best in terms of Stress.[4]

Regionally Constrained MDS

Optimal scaling of the external variables can also be used to impose regional constraints on the MDS solution. For example, we know for each point two facets and use MDS to separate the different classes of points by two sets of parallel lines, where each set corresponds to one facet. This constraint only works if the number of dimensions is equal to the number of external variables and only for axial partitioning of the MDS space. In addition, the facets should be ordered.

Consider the following example. The ordinal MDS solution (Figure 4.7) of the Morse code data of Rothkopf (1957) was interpreted using two physical properties of the signals. The two properties are signal length (varying from .05 to .95 seconds) and signal type (the ratio of long vs. short beeps). Figure 4.7 shows that the unconstrained MDS solution can be partitioned by these two facets. However, the dashed lines (partitioning the plane by signal type) have a rather irregular curvature. We now ask whether an MDS solution

[4] On the dimension "size", in contrast, this spacing is not obvious, which may be due to the size range of the ellipses. See also Figure 17.8 for a similar experiment with rectangles.

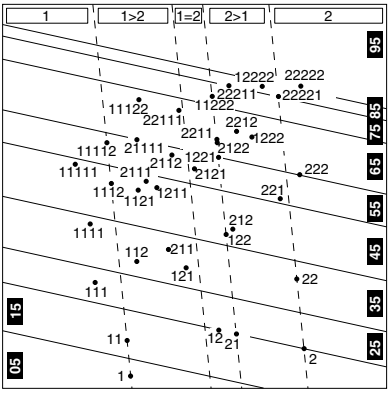

FIGURE 10.5. Theory-consistent solution by constrained MDS of the Morse code data of Rothkopf (1957) with Stress .21.

exists that can be partitioned by straight lines while still being acceptable in terms of Stress. We use PROXSCAL to answer this question.

The external variables for signal type $\mathbf{y}_1$ and the signal length $\mathbf{y}_2$ have a value for every Morse code. The 2D MDS space is constrained to be a linear combination of signal length and signal type; that is, $\mathbf{X} = \widehat{\mathbf{Y}}\mathbf{C}$, with $\widehat{\mathbf{Y}} = [\widehat{\mathbf{y}}_1\ \widehat{\mathbf{y}}_2]$ where $\widehat{\mathbf{y}}_1$ and $\widehat{\mathbf{y}}_2$ are monotonic transformations of $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively. In contrast to the example above, we now allow that tied coordinates can be untied (primary approach to ties). This combination of restrictions implies that there is a direction in the MDS space that represents $\widehat{\mathbf{y}}_1$, and all projections of the points onto this line satisfy the order of the signal lengths, so that perpendicular lines separate the space into regions with equal signal lengths. The same holds for $\widehat{\mathbf{y}}_2$, so that a separation of the space for signal types is obtained. Figure 10.5 shows the solution of the ordinal MDS analysis with the external constraints described above. This theory-consistent configuration has Stress .21 ($\sigma_r = .043$), and the unconstrained solution in Figure 4.7 has Stress .18. The raw Stress can be decomposed into model Stress (=.0429, .997%) and Stress due to the constraints (=.00011, .003%). Apart from theory-consistency, Figures 10.5 and 4.7 differ with respect to the location of the points representing the Morse codes E and T (labeled as 1 and 2, respectively, in both figures). These points are less well represented, which is reflected by their Stress per point (see Section 3.4) of .095 and .082, respectively, the largest contributions to overall Stress. In summary, however, the difference in Stress of the constrained and the unconstrained solutions is rather small, so that the theory-consistent solution seems acceptable.

Cluster Differences Scaling

A different type of constraint was suggested by Heiser (1993), who proposed a mixture of an MDS analysis and a cluster analysis. This method was called *cluster differences scaling*. Every object is assigned to a cluster, and every cluster is represented by a point in the space. In cluster differences scaling, Stress is optimized over the coordinates *and* over the cluster memberships. Groenen (1993) showed that this method can be interpreted as MDS with the restriction that the configuration is of the type $\mathbf{GX}$, where $\mathbf{G}$ is an $n \times k$ indicator matrix (which has a single one in each row, and all other values zero), and $\mathbf{X}$ is a $k \times m$ matrix of the k cluster coordinates. Heiser and Groenen (1997) elaborate on this model, give a decomposition of the dispersion (sum of squared dissimilarities), and present a convergent algorithm. The assignment of objects to clusters (by matrix $\mathbf{G}$) gives rise to many local minima. Groenen (1993) and Heiser and Groenen (1997) managed to avoid such local minima by repeatedly computing a fuzzy form of cluster differences scaling until the fuzzy form yields the same result as the crisp form.

The Extended Euclidean Model

Suppose that the proximities are not very well explained by an MDS in low-dimensional space. One reason could be that some objects are quite unique. One could account for this and allow each object to retain its uniqueness in MDS by assigning a uniqueness dimension to each object, in addition to the low-dimensional space common to all objects (Bentler & Weeks, 1978). A uniqueness dimension $\mathbf{x}_i$ for object i has coordinates of zero for all objects, except for object i. Thus, the matrix of coordinates consists of the usual $n \times m$ matrix of coordinates $\mathbf{X}$ common to all objects, augmented by a diagonal $n \times n$ matrix $\mathbf{U}$ of uniqueness coordinates. The augmented coordinate matrix is denoted by $[\mathbf{X}|\mathbf{U}]$. The distance between objects i and j is

$$d_{ij}(\mathbf{X}|\mathbf{U}) = \left(\sum_{a=1}^{m} (x_{ia} - x_{ja})^2 + u_{ii}^2 + u_{jj}^2 \right)^{1/2},$$

which is called the *extended Euclidean* distance (Winsberg & Carroll, 1989). The distance consists of a common part and a part determined by the uniqueness of the objects i and j.

This special coordinate matrix also can be viewed as an example of a constrained configuration. The constraints simply consist of fixing the off-diagonal elements of $\mathbf{U}$ to zero while leaving the diagonal elements free (Bentler & Weeks, 1978).

10.4 Weakly Constrained MDS

We now consider a weaker form of constraining an MDS solution. It puts additional external restrictions on the configuration that are not strictly enforced. Rather, they may be violated, but any violation leads to higher Stress. Weakly constrained MDS attempts to minimize such violations.

Let us try to represent the color data from Table 4.1 ordinally by distances in a plane so that (a) the Stress is as low as possible, and (b) all points lie on a perfect circle. Condition (b) is the external constraint imposed on the MDS representation.

Figure 4.1 shows that the usual MDS result already satisfies condition (b) very closely, so we use this solution in the following as a starting configuration. The confirmatory scaling problem then consists of finding a projection of the points onto a circle such that the Stress value goes up as little as possible. If we pick a point somewhere close to the center of the color circle in Figure 4.1 and construct a circle around this point such that it encloses all points of the configuration, then an approximate solution to our scaling problem could be found simply by projecting all points radially towards the outside onto this circle. An optimal solution can be constructed in a similar fashion.

First, augment the proximity matrix in Table 4.1 with a "dummy" object Z. Z does not stand for an additional concrete stimulus, but serves an auxiliary purpose here and represents the circle center in the MDS configuration. The proximities between Z and any of the real stimuli $434, 445, \ldots, 674$ are defined as missing data. This leads to the 15×15 data matrix $\mathbf{P}_1$ in Table 10.3.

Second, define another 15×15 proximity matrix $\mathbf{P}_2$ which expresses the side constraints. No external constraints are to be imposed on the distances between any two color points. However, all should lie on a circle and so all must have the same distance to point Z. This gives the constraint pattern $\mathbf{P}_2$ shown in Table 10.4, where all elements except those in row Z and in column Z are missing values. All elements in row and column Z are set equal to 10, but any other number would do as well.

Third, use the configuration in Figure 4.1 as a starting configuration, after adding proper coordinates for one further point, Z. The coordinates of Z should be chosen so that Z lies somewhere in the center of the circular manifold in Figure 4.1. This can be done most easily by centering the MDS configuration in Figure 4.1, that is, shifting it so that the centroid of all 14 points coincides with the origin, or, computationally, by subtracting the mean from the values in each column of $\mathbf{X}$ in turn. Z then has the coordinates (0.0, 0.0).

Fourth, define a loss criterion for the scaling procedure. We choose

$$\sigma_T(\mathbf{X}; \mathbf{P}_1; \mathbf{P}_2) = \sigma_r(\mathbf{X}; \mathbf{P}_1) + a \cdot \sigma_r(\mathbf{X}; \mathbf{P}_2), \tag{10.5}$$

TABLE 10.3. Similarities for colors with wavelengths 434 to 674 nm (Ekman, 1954); Z is a dummy variable; – denotes a missing value; the matrix is called $\mathbf{P}_1$ in the text.

nm	434	445	465	472	490	504	537	555	584	600	610	628	651	674	Z
434	–	86	42	42	18	06	07	04	02	07	09	12	13	16	–
445	86	–	50	44	22	09	07	07	02	04	07	11	13	14	–
465	42	50	–	81	47	17	10	08	02	01	02	01	05	03	–
472	42	44	81	–	54	25	10	09	02	01	00	01	02	04	–
490	18	22	47	54	–	61	31	26	07	02	02	01	02	00	–
504	06	09	17	25	61	–	62	45	14	08	02	02	02	01	–
537	07	07	10	10	31	62	–	73	22	14	05	02	02	00	–
555	04	07	08	09	26	45	73	–	33	19	04	03	02	02	–
584	02	02	02	02	07	14	22	33	–	58	37	27	20	23	–
600	07	04	01	01	02	08	14	19	58	–	74	50	41	28	–
610	09	07	02	00	02	02	05	04	37	74	–	76	62	55	–
628	12	11	01	01	01	02	02	03	27	50	76	–	85	68	–
651	13	13	05	02	02	02	02	02	20	41	62	85	–	76	–
674	16	14	03	04	00	01	00	02	23	28	55	68	76	–	–
Z	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

TABLE 10.4. Restriction matrix for color data in Table 10.3; – denotes a missing value; the matrix is called $\mathbf{P}_2$ in the text.

nm	434	445	465	472	490	504	537	555	584	600	610	628	651	674	Z
434	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
445	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
465	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
472	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
490	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
504	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
537	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
555	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
584	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
600	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
610	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
628	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
651	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
674	–	–	–	–	–	–	–	–	–	–	–	–	–	–	10
Z	10	10	10	10	10	10	10	10	10	10	10	10	10	10	–

where $\sigma_r(\mathbf{X};\mathbf{P}_1)$ is the loss of configuration $\mathbf{X}$ relative to $\mathbf{P}_1$, $\sigma_r(\mathbf{X};\mathbf{P}_2)$ the loss relative to $\mathbf{P}_2$, and a is a nonnegative weight. This means that $\sigma_r(\mathbf{X};\mathbf{P}_1)$ is the Stress of a given configuration relative to the proximity matrix $\mathbf{P}_1$, and $\sigma_r(\mathbf{X};\mathbf{P}_2)$ is the Stress of this configuration relative to the constraint matrix $\mathbf{P}_2$. The second term of (10.5) is called a *penalty* term. It penalizes the solution for not satisfying the constraints. The strength of the penalty is determined by the size of the *penalty parameter* a. Of course, $\sigma_r(\mathbf{X};\mathbf{P}_1)$ and $\sigma_r(\mathbf{X};\mathbf{P}_2)$ are computed only over those elements of the matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ that are not defined to be missing data. Thus,

$$\sigma_r(\mathbf{X};\mathbf{P}) = \sum_{i<j}[\widehat{d}_{ij} - d_{ij}(\mathbf{X})]^2, \text{ for all defined } p_{ij},$$

where $\widehat{d}_{ij}$ (dependent on $\mathbf{P}$) is the target distance (disparity) of $d_{ij}(\mathbf{X})$ defined by the chosen MDS model. In the present example, we choose ordinal MDS and the secondary approach to ties on $\mathbf{P}_2$, because all tied data values in the restriction matrix $\mathbf{P}_2$ should be mapped into exactly the same distance. (With the primary approach to ties, $\sigma_r(\mathbf{X};\mathbf{P}_2) = 0$ for any $\mathbf{X}$, because all defined elements of $\mathbf{P}_2$ are equal.) But then the target distances in $\sigma_r(\mathbf{X};\mathbf{P}_2)$ obtained by monotone regression are all equal to the arithmetic mean of the distances from point Z to all other points of the configuration $\mathbf{X}$.

Fifth, find a method to minimize (10.5). This does not pose a new problem. We proceed as in Chapter 8, that is, using the majorizing approach to minimize Stress.

Sixth, given the initial configuration of the unconstrained solution in Figure 4.1, iterate to solve the MDS task. If we start with $a = 1$, the restrictions only slightly determine the final solution. As a is increased, the effect of the side constraints on the configuration is increased. If $a \to \infty$, then every solution strictly satisfies the circular constraint. Because the effect of the constraints on the solution is set by the penalty parameter a, the method in this section that minimizes (10.5) may be called *weakly constrained MDS* (after the weakly constrained regression of Ten Berge, 1991). Often, choosing $a = 100$ generates a theory-conforming solution.

If it is at all possible to impose the side constraints onto a configuration of n points in a space of fixed dimensionality, we should end up with $\sigma_r(\mathbf{X};\mathbf{P}_2) = 0$, provided the iterations do not get stuck in a local minimum. Of course, we can impose conditions that are impossible to satisfy (e.g., attempting to represent the distances among a cube's corners in a ratio MDS plane). The final $\sigma_r(\mathbf{X};\mathbf{P}_1)$ is an index for how well the theory-conforming solution represents the given data. However, the raw measures $\sigma_r(\mathbf{X};\mathbf{P}_1)$ and $\sigma_r(\mathbf{X};\mathbf{P}_2)$ are not very practical, so we express the fit of $\mathbf{X}$ relative to $\mathbf{P}_1$ and $\mathbf{P}_2$ by a more familiar index such as Stress.

A procedure similar to the one described above is the program CMDA (Borg & Lingoes, 1980). Weakly constrained MDS can also be computed

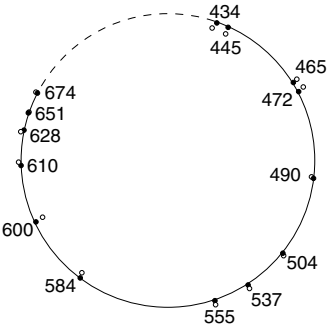

FIGURE 10.6. Stress-optimal (circles) and perfectly circular (points) MDS representation of color proximities in Table 4.1.

by programs that allow for data weights, such as KYST and PROXSCAL. Then data weights w_{ijk} are set to 0 for missing proximities, and $w_{ij1} = 1$ for nonmissing proximities of $\mathbf{P}_1$ and $w_{ij2} = a$ for nonmissing proximities of $\mathbf{P}_2$.

With the matrices given in Tables 10.3 and 10.4, and using ordinal MDS with the secondary approach to ties, we obtain the configuration of the solid points in Figure 10.6. To demonstrate how $\sigma_r(\mathbf{X}; \mathbf{P}_2)$ has affected the MDS solution, Figure 10.6 also shows the MDS configuration (open circles) obtained from a regular MDS analysis. The Stress of the weakly constrained MDS configuration relative to $\mathbf{P}_1$ is $\sigma_1 = 0.0374$, whereas it is $\sigma_1 = 0.0316$ for the unconstrained MDS configuration. The side constraints have led to an increment in Stress so small that both representations should be considered equally good, especially because we can assume that the data are not error-free. We therefore conclude that the color-circle theory is compatible with the given data.

A different approach was followed by Cox and Cox (1991). They forced the configuration onto the surface of a sphere by expressing the point coordinates not as Cartesian but as spherical coordinates, and then minimizing Stress only over the longitude and latitude angles that specify the points' positions in space.

Hubert, Arabie, and Meulman (1997) analyzed a related but different problem, namely to model the dissimilarities by distance between points *along the path* of the circle. This model is essentially the same as unidimensional scaling where the dimension is bent to be circular. Hubert et al. call their model circular unidimensional scaling. In Section 13.5, we show that unidimensional scaling suffers from many local minima and that a combinatorial approach is useful to find a global minimum. For this reason, Hubert et al. use combinatorial optimization together with iterative projection techniques to solve the circular unidimensional scaling problem.

TABLE 10.5. Correlations of eight intelligence tests (lower half); for structuples, see text; upper half contains hypothesized similarities.

Test	NA1	NA2	NI	GI1	GI2	GA1	GA2	GA3
NA1	–	5	4	3	3	4	4	4
NA2	.67	–	4	3	3	4	4	4
NI	.40	.50	–	4	4	3	3	3
GI1	.19	.26	.52	–	5	4	4	4
GI2	.12	.20	.39	.55	–	4	4	4
GA1	.25	.28	.31	.49	.46	–	5	5
GA2	.26	.26	.18	.25	.29	.42	–	5
GA3	.39	.38	.24	.22	.14	.38	.40	–

Enforcing Order Constraints onto MDS Distances

We now look at weakly constrained MDS where certain order relations are imposed onto the MDS distances. Consider the correlation matrix in Table 5.1, repeated for convenience in the lower half of Table 10.5 (Guttman, 1965). The variables here are eight intelligence test items, coded by the facets "language of communication" = {N = numerical, G = geometrical} and "requirement" = {I = inference, A = application}. For example, item 1 and item 2 both were classified as NA or numerical-application items.

One can predict how these items should be correlated among each other by invoking the *contiguity principle*. This principle is based on the (seemingly) plausible idea that "variables which are more similar in their facet structure will also be more related empirically" (Foa, 1965, p.264).[5] Similarity in facet structure is typically defined as the number of structs that two structuples have in common, whereas empirical similarity is assessed by some correlation between items (Foa, 1958). Hence, one predicts here, for example, that item 4 should be at least as similar to item 8 as to item 2, because the former share one definitional element (their language), whereas the latter differ on both facets. Predictions of this kind imply that the MDS configuration should be circular (Figure 10.7).

To test this prediction, we have to set up a restriction matrix $\mathbf{P}_2$ that enforces certain order relations onto the MDS distances. Because $\mathbf{P}_1$ (lower half in Table 10.5) contains similarity coefficients, we choose $\mathbf{P}_2$'s values correspondingly. A $\mathbf{P}_2$ that confirms the theory of Figure 10.7 is given in the upper half of Table 10.5. It is built as follows. The proximities for items with the same structuples, such as p(NA1,NA2) and p(GA1,GA3), all are set to the value 5. The proximities that correspond to the immediate

[5]Upon closer inspection, the contiguity makes sense only if all facets are ordered in the sense of the observational range (see Borg & Shye, 1995). However, we do not study these conditions here but simply use the example to demonstrate how certain constraints can be set up.

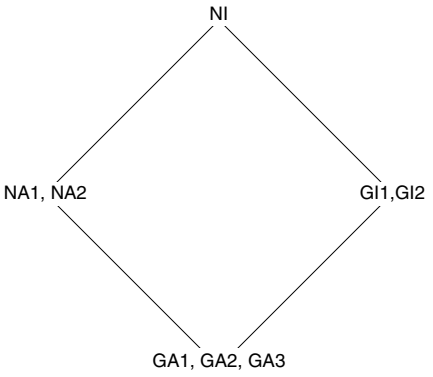

FIGURE 10.7. Hypothesized configuration of points representing intelligence tests with different facet compositions (G = geometrical, N = numerical; A = application, I = inference).

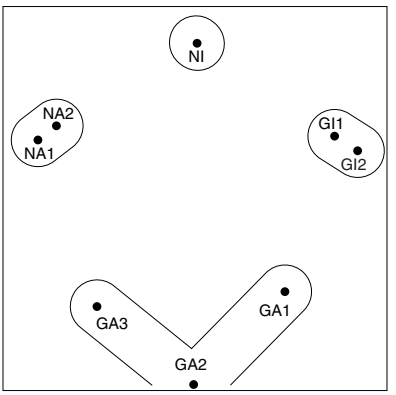

FIGURE 10.8. Best MDS representation that perfectly satisfies regional hypotheses in Figure 10.7.

neighborhood relations (shown in Figure 10.7 by the line segments) are set to the value 4, since none of these distances should be larger than any distance between definitionally equivalent items. Finally, what remain are the large distances between the groups NI, GA and the groups NA, GI, which are set to 3 in $\mathbf{P}_2$. Because we are doing ordinal MDS on similarities, the values 3, 4, and 5 are immaterial and may be replaced by any numbers that have the same order. By using the primary approach to ties on $\mathbf{P}_2$, all distances associated with a, say, 4 in $\mathbf{P}_2$, should not be larger than those associated with a 3 in $\mathbf{P}_2$. However, the distances within either class are not required to be equal.

The weakly constrained MDS representation (with $a = 100$) is shown in Figure 10.8. It satisfies the side constraints perfectly, with an acceptably small overall Stress ($\sigma_1 = .0404$). What remains, though, is some scatter among the items with the same structuples (notably GA and GI), so more and/or modified facets might be considered.

10.5 General Comments on Confirmatory MDS

Confirmatory MDS offers models that are open for theory-driven ad hoc specifications. Standard MDS models, in contrast, are more like closed systems that allow the user only some global specifications, such as choosing the dimensionality of the solution space or the Minkowski metric parameter (see Chapter 17). The purpose of confirmatory MDS is to enforce certain expected relations on an otherwise optimal data representation in order to see how compatible these relations are with the data.

The varieties of confirmatory MDS are, in principle, without bounds. New theories may require new confirmatory MDS procedures. Dimensional theories are best served by the existing procedures, and regional theories worst. We have seen that it is rather easy to enforce certain axial partitionings onto an MDS solution. It is also not too difficult to enforce cluster structures onto the MDS configuration, for example, for strict clustering by cluster differences scaling (Heiser & Groenen, 1997) and for weak clustering by setting up appropriate order constraints on the distances (Borg & Lingoes, 1987). However, with the MDS programs available today, it is difficult to enforce a more intricate regional pattern such as a radex, for example, onto an MDS solution. It is even more difficult, or even impossible, to formulate constraints on *general partitionability* relative to a facet design for the points, as discussed in Chapter 5.[6]

Apart from such problems of enforcing particular types of constraints onto an MDS solution, the general question of how to evaluate such methods and their results within cumulative scientific research remains to be answered. The more theoretically guided researcher may be tempted to always force his or her theory onto an MDS solution and then evaluate the resulting Stress. Unless there is a good estimate for the random noise component in the data (e.g., reliability measures), this is a dangerous strategy, because it does not allow one to separate *errors of approximation* from *errors of estimation*. The latter are due to sampling errors, and Stress in standard (unconstrained) MDS essentially reflects, as we saw in Chapter 3, such random errors in the data. Thus, high Stress values may be discarded as "technical" information only. Errors of approximation, however, would not go away even if the data were perfectly reliable. They simply express the misfit of the model. To separate these two error sources, one should always compute a standard MDS solution and then compare its Stress to the Stress obtained under additional restrictions. What is important, then, is the Stress increment. If strict constraints are used (as opposed to weakly constrained MDS), one should compare the ratio of Stress due to model misfit and Stress due to violation of the constraints. If the latter term is relatively small, then the theory-confirming solution can be accepted.

As a rule of thumb, it holds that if a standard MDS solution is similar to what was predicted theoretically, enforcing the theory onto the MDS solution does not make much difference in terms of Stress. However, if standard MDS does not yield the expected configuration, then it is impossible to say whether confirmatory MDS will make much difference. That depends on the loss function and its local minima.

---

[6] Guttman (1976) suggested combining MDS with *multidimensional scalogram analysis*, using MSA's notions of contiguity. See also Borg and Groenen (1998). These ideas have not yet been studied systematically, however.

Stress increments must be evaluated in any case, and this is a complex matter (Lingoes & Borg, 1983). What must be taken into account here is: (a) the number of points, n, because enforcing external constraints on few points or distances is generally easier than dealing with a large n; (b) the dimensionality of the MDS solution, m, for reasons similar to those for n; (c) the error component in the proximities, because with very noisy data, further constraints have less effect on Stress; (d) the similarity between the standard solution and the confirmatory solution: minor corrections of the standard solution should have little effect on Stress; and (e) the increased theoretical intelligibility of the confirmatory solution over the standard solution: if the latter makes little sense, one may be willing to accept larger increments in Stress, because a theoretically justified solution promises to be more stable over replications, and there is no reason to predict stability for structures that are not understood.

10.6 Exercises

Exercise 10.1 Consider the matrix below. In its lower half it shows similarity coefficients for tonal stimuli (Levelt, Geer, & Plomp, 1966). Each stimulus consisted of two simultaneously heard tones with a fixed ratio between their frequencies. Fifteen stimuli were used: the twelve musical intervals within the octave; and in addition two wider intervals (4:9 and 2:5) and one narrow interval between minor and major second (11:12). To control for pitch effects, the mean frequency for each interval was held constant at 500 Hz. Previous analyses by Levelt et al. (1966) and by Shepard (1974) had shown that the subjective similarities for these tone intervals form a horseshoe-like structure in the two-dimensional MDS plane.

Freq. Ratio	No.	15	13	12	7	6	3	9	2	10	5	8	14	1	11	4
15:16	15	–	14	13	12	11	10	9	8	7	6	5	4	3	2	1
11:12	13	32	–	14	13	12	11	10	9	8	7	6	5	4	3	2
8:9	12	29	32	–	14	13	12	11	10	9	8	7	6	5	4	3
5:6	7	19	22	28	–	14	13	12	11	10	9	8	7	6	5	4
4:5	6	14	17	23	28	–	14	13	12	11	10	9	8	7	6	5
3:4	3	15	10	13	22	25	–	14	13	12	11	10	9	8	7	6
5:7	9	8	8	14	25	24	27	–	14	13	12	11	10	9	8	7
1:2	2	9	10	14	13	18	21	22	–	14	13	12	11	10	9	8
5:8	10	6	7	13	20	21	17	22	27	–	14	13	12	11	10	9
3:5	5	12	11	12	15	20	24	13	25	24	–	14	13	12	11	10
4:7	8	7	11	15	14	18	14	16	13	27	30	–	14	13	12	11
8:15	14	7	10	7	10	17	10	19	18	18	18	26	–	14	13	12
1:2	1	8	3	9	9	14	15	9	14	12	13	22	26	–	14	13
4:9	11	9	14	6	8	8	12	9	10	10	17	20	13	29	–	14
2:5	4	9	14	6	10	7	12	10	16	18	18	14	13	25	30	–

TABLE 10.6. Correlation coefficients (decimal points omitted) for the 30 forms of protest acts described in Table 1.2.

Act	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2	41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
3	53	45	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	20	30	26	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
5	20	37	33	46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	18	27	28	41	54	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
7	19	29	29	45	58	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
8	25	30	37	36	47	46	56	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
9	03	14	06	31	34	40	41	33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10	03	09	04	20	25	31	31	27	52	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
11	53	15	31	10	08	07	08	12	-03	05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12	37	64	39	26	31	24	25	27	10	07	29	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13	38	26	57	17	18	19	16	23	03	05	49	39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
14	19	26	24	58	37	32	34	29	21	15	16	37	24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
15	21	30	29	31	52	34	37	36	18	12	12	40	24	45	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
16	13	16	21	23	33	49	32	30	23	15	15	23	25	31	43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
17	17	23	23	26	36	36	46	39	21	20	06	36	21	40	60	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.
18	20	22	25	23	27	32	33	51	17	15	15	32	29	36	50	44	62	.	.	.	.	.	.	.	.	.	.	.	.	.
19	02	08	05	16	18	23	22	20	29	20	00	15	08	31	34	38	43	43	.	.	.	.	.	.	.	.	.	.	.	.
20	00	06	03	12	14	20	18	17	21	37	00	16	07	25	31	31	39	38	70	.	.	.	.	.	.	.	.	.	.	.
21	57	34	44	18	22	18	20	25	-01	02	44	34	36	16	24	14	21	23	04	06	.	.	.	.	.	.	.	.	.	.
22	33	68	37	22	28	21	25	29	09	06	16	63	25	22	28	15	26	27	08	08	45	.	.	.	.	.	.	.	.	.
23	41	35	60	21	29	26	27	34	04	05	30	34	48	20	27	18	24	28	04	04	55	46	.	.	.	.	.	.	.	.
24	23	26	26	54	38	32	39	33	15	17	14	27	14	52	31	21	32	25	16	14	29	30	34	.	.	.	.	.	.	.
25	22	28	30	32	61	38	42	40	20	15	10	27	18	29	51	27	37	31	17	13	28	33	41	45	.	.	.	.	.	.
26	17	21	25	29	36	59	37	34	26	20	09	22	18	28	31	48	35	31	22	20	22	27	32	35	47	.	.	.	.	.
27	20	26	27	31	43	42	52	42	23	22	11	27	18	27	34	28	47	31	17	15	24	33	34	47	54	54	.	.	.	.
28	24	25	32	30	36	36	42	58	20	16	15	27	24	28	33	26	38	46	17	16	31	36	45	44	49	45	58	.	.	.
29	02	08	03	18	20	23	20	17	32	24	03	09	05	18	15	16	17	13	24	15	07	15	09	22	22	30	32	32	.	.
30	02	06	01	16	18	19	16	14	25	33	03	09	04	15	10	12	15	12	17	28	09	12	09	22	22	26	31	30	57	.

(a) Replicate Shepard's MDS analysis and verify that the order of the stimuli on the horseshoe corresponds to the order of the entries of the table.

(b) Use confirmatory ordinal MDS and try to unbend the horseshoe such that it does not bend back upon itself. The upper half of the table indicates a pattern of values that satifies such a simple structure. You may use these pseudodata to impose the "unbending" constraints, but note that a simplex is a biconditional order structure. Impose only minimal constraints.

(c) Compare the constrained MDS solution with the one that does not use external constraints and discuss the findings.

Exercise 10.2 Consider the lower-half matrix in Table 17.7. Its unconstrained city-block MDS representation is shown in Figure 17.8. Try to force a perfect "rectangular" structure onto this solution so that, for example, points 1, 2, 3, and 4 lie on a straight vertical line, and points 1, 5, 9, and 13 lie on a straight horizontal line (see dashed grid in Figure 17.8).

Exercise 10.3 Table 10.6 shows the intercorrelations of the 30 forms of protest behavior (Levy, 1983) analyzed before in Section 1.2 (West German data of early 1974, $N = 2307$).

(a) Use confirmatory MDS to enforce a solution where the points are perfectly separated in 3D space in the sense of their design facets that are shown in Table 1.2.

(b) Compare the solution to an "unconstrained" solution as discussed in Section 1.2.

(c) Discuss any amount of additional Stress due to the external constraints.